UC San Diego
UC San Diego Electronic Theses and Dissertations

Title
The Ethics of Distribution /

Permalink
https://escholarship.org/uc/item/8w4732hp

Author
Pummer, Theron

Publication Date
2013

UNIVERSITY OF CALIFORNIA, SAN DIEGO

The Ethics of Distribution

A dissertation submitted in partial satisfaction of the
requirements for the degree of Doctor of Philosophy

in

Philosophy

by

Theron Pummer

Committee in charge:

      Professor Richard Arneson, Co-Chair
      Professor David Brink, Co-Chair
      Professor Dana Nelkin
      Professor David Schkade
      Professor Joel Sobel

2013

The Dissertation of Theron Pummer is approved, and it is acceptable in quality and

form for publication on microfilm and electronically:

_____

_____

_____

_____
                                                                    Co-Chair

_____
                                                                    Co-Chair

University of California, San Diego

2013

DEDICATION

This dissertation has benefited tremendously from feedback I have received from many friends, colleagues, and mentors. I therefore dedicate it to the many kind people listed in the Acknowledgements.

EPIGRAPH

…[W]hen we are choosing between two acts or policies, one relevant fact is how great the resulting benefits will be. For Utilitarians, that is all that matters. On their view, we should always aim for the greatest sum of benefits. But, for Egalitarians, it also matters how well off the beneficiaries would be. We should sometimes choose a smaller sum of benefits, for the sake of a better distribution.

How can we make a distribution better? Some say: by aiming for equality between different people. Others say: by giving priority to those who are worse off. As we shall see, these are different ideas.

Should we accept these ideas? Does equality matter? If so, when and why? What kind of priority, if any, should we give to those who are worse off?

These are difficult questions, but their subject matter is, in a way, simple. It is enough to consider different possible states of affairs, or outcomes, each involving the same set of people. We imagine that we know how well off, in these outcomes, these people would be. We then ask whether either outcome would be better, or would be the outcome we ought to bring about. This subject we can call *the ethics of distribution*.

Derek Parfit

TABLE OF CONTENTS

LIST OF FIGURES

ACKNOWLEDGEMENTS

interact with each of them. In person and over e-mail, each of these philosophers has very kindly and substantially helped me.

I am grateful to David DeGrazia for introducing me to ethics as an undergraduate, and I am grateful to David Brink, Dana Nelkin, Dick Arneson, and Jerry Doppelt for together reintroducing me to ethics as a graduate student.

Finally, and on a more personal note, I would like to thank my parents, Mari and Doug, my parents-in-law, Jerry and Julie, my grandparents, Ray and Lorraine, all my siblings, my very good friend, Caley, and my wonderful wife, Sarah. I am thanking them for too many different things to even seriously attempt to list here, but suffice it to say that there is no way this dissertation would exist without their kindness.

If you are reading this and have helped me, but are not on this list, I sincerely apologize. I promise I have not intentionally omitted you – my memory could be better.

VITA

2006	Bachelor of Science, George Washington University
2006-2013	Teaching Assistant, Research Assistant, and Instructor, University of California, San Diego
2010	Master of Arts, University of California, San Diego
2011	Recognized Student, University of Oxford
2013	Doctor of Philosophy, University of California, San Diego

PUBLICATIONS

"Intuitions about Large Number Cases" *Analysis* 73(1) (2013): 37-46

"Does Division Multiply Desert?" *Philosophical Review* 123(1) (2014) (forthcoming)

ABSTRACT OF THE DISSERTATION

The Ethics of Distribution

by

Theron Pummer

Doctor of Philosophy in Philosophy

University of California, San Diego, 2013

Professor Richard Arneson, Co-Chair
Professor David Brink, Co-Chair

For any badly off person we consider, there could be someone who is much worse off. For example, suppose that Jane is the currently worst off person. Jane is so miserable that her life is worth not living. Nonetheless, one would be much worse off than Jane if one were as miserable as her at each time, but lived for much longer. Since there could be such persons who would live *indefinitely* longer than Jane, there could be persons who would be indefinitely worse off than Jane. I call such extremely badly off persons *Priority Monsters*.

The possibility of Priority Monsters raises new and important challenges in the ethics of distribution, an area of ethics which addresses how we ought to distribute benefits and burdens across separate persons. The principal challenge I focus on is how to avoid the conclusion that, if such Priority Monsters existed, we would be morally required to benefit them, no matter how little we provide them, at the expense of doing much more good for others, like Jane, who are themselves significantly badly off. Most of us find this conclusion very hard to accept – surely if it were between sparing Jane of many years of misery and sparing a Priority Monster of a mild headache, we should do the former.

Utilitarian principles of distribution say that we ought to maximize the overall sum of benefits. Egalitarian and prioritarian principles place special moral weight on achieving benefits for persons who are worse off. Many contemporary philosophers contend that the latter principles are more plausible than utilitarian principles. I argue that while utilitarian principles easily and naturally avoid implausible conclusions about Priority Monsters, egalitarian and prioritarian cannot plausibly avoid such conclusions. Utilitarian principles thus avoid a significant difficulty that these other principles face, and they might therefore be more plausible, overall, than their non-utilitarian rivals.

Chapter 1

The Ethics of Distribution

CHAPTER OUTLINE

1.1   The Distribution of Well-Being

Derek Parfit's words provide a clear and compact characterization of the

subject matter of this dissertation:

> …[W]hen we are choosing between two acts or policies, one relevant
> fact is how great the resulting benefits will be. For Utilitarians, that is all that
> matters. On their view, we should always aim for the greatest sum of benefits.
> But, for Egalitarians, it also matters how well off the beneficiaries would be.
> We should sometimes choose a smaller sum of benefits, for the sake of a better
> distribution.
>   How can we make a distribution better? Some say: by aiming for
> equality between different people. Others say: by giving priority to those who
> are worse off. As we shall see, these are different ideas.
>   Should we accept these ideas? Does equality matter? If so, when and
> why? What kind of priority, if any, should we give to those who are worse
> off?
>   These are difficult questions, but their subject matter is, in a way,
> simple. It is enough to consider different possible states of affairs, or
> outcomes, each involving the same set of people. We imagine that we know
> how well off, in these outcomes, these people would be. We then ask whether

either outcome would be better, or would be the outcome we ought to bring about. This subject we can call *the ethics of distribution*.[1]

Let me add to Parfit's words, to further introduce the content, scope, and importance of our subject.

1.1.1 Resources and Benefits

The first thing to notice about this subject is that its concern is the distribution of *benefits*, rather than *resources*. Benefits are increases in *well-being*, or whatever makes lives go well for the individuals living these lives. Resources are *means* to benefits, or to other ends. Money, for example, is a resource. If we donate an equal amount of money to everyone, the effect may not be an equal benefit for everyone. This is in part for the familiar reason that money, and most other resources, exhibit diminishing marginal well-being. Beyond some point, gaining more and more money will benefit a person less and less. Giving a Kenyan child $10,000 would probably result in a greater benefit than giving it to a Hollywood millionaire. This, it seems, gives us more reason to give the money to the poor child, and to make the overall distribution of money somewhat less unequal.

Does the fact that resources generally exhibit diminishing marginal well-being imply that, insofar as we should promote the most total well-being, we should distribute resources, such as money, more equally? Not necessarily. Sometimes the

---

[1] Parfit 1991, 81-2.

worse off are inefficient converters of resources into well-being. Consider a disabled American child, who is just as badly off in terms of well-being as the Kenyan child. $10,000 might only pay for a tiny fraction of her medical supplies, and so only marginally benefit her. And sometimes the very well off are highly efficient converters of resources into well-being. Consider a particular Hollywood millionaire who gets *a lot* of pleasure from each additional $10,000 she receives. In a two-person society consisting of this disabled child and this millionaire, the aim of promoting the most total well-being would be best served through an extremely unequal distribution of money. These are familiar points about the relation between resources and well-being.[2] I mention them because, while the distribution of resources is clearly relevant to the distribution of well-being, the subject of this dissertation is the latter, and isn't to be confused with the former. And when I talk of people being better or worse off, I mean better or worse off in terms of well-being, not resources.

1.1.2 Well-Being

Benefits are increases in well-being. Well-being is whatever it is that makes lives go better for the individuals living them. But what *is it* that makes lives go better for the individuals living them? That is a very good question, to which there are

---

[2] For example, see Sen 1979, 203-4 and Broome 1991, 175-7.

several competing answers.[3] On one broad type of account, what makes an individual's life go better for her simply consists in her getting her desires, or preferences, satisfied. On another very broad type of account, things can make an individual better or worse off, independently of her desires or preferences. For example, many people think that pleasure is one thing that, other things equal, makes lives go better, and that pain is one thing that, other things equal, makes lives go worse, independently of whether or not these things are desired.

Though it matters a great deal what well-being consists in, we can, I believe, make progress in the ethics of distribution without settling this question. We can understand the sorts of questions Parfit listed in the above quote, and feel strongly inclined to answer them in one way or another, prior to filling in all the details about well-being. This is not to deny the possibility that certain answers to these structural distributive questions might seem more or less plausible, depending on which views about the content of well-being we have in mind.[4] But this is to say that such structural issues can arise independently of questions about the content of well-being, and that these potential complications concerning the interaction between the ethics of distribution and substantive theories of well-being can be temporarily set aside. I am happy to go back through the critical steps and arguments made in this dissertation, through the lens of various particular theories of well-being, to see what effect, if any,

---

[3] Parfit 1994, Appendix I, Griffin 1986, Sumner 1996, and Crisp 2006a provide excellent discussions of competing accounts of well-being, or what makes lives go well for the individuals living them.

[4] Perhaps, for example, "average" views will seem more plausible than "total" views, when focusing on accomplishments rather than pleasures. One can find an excellent discussion of such issues in Hurka 1993, chapter 6.

doing so would have on the plausibility of these steps and arguments. But, for now, I can only do so much. I hope that what I will go on to say is at least consistent with the most plausible theories of well-being. However, I should add that I do believe that, other things equal, pain makes lives go worse. Most people agree with this. It will sometimes be useful for me, for the sake of concreteness, to cite pain relief as a benefit.

1.1.3 Better Than

The ethics of distribution is concerned with what makes a given distribution of well-being across persons *better than* another such distribution. In cases where the only facts that could make one outcome better than another are facts about well-being and its distribution, our question becomes which *outcome* is better.

Some people find it problematic to ask whether an outcome could be *better* than another. They insist that, rather than "better," we must instead mean "better *for*" or some other notion, in order to make sense.[5] But I am not sure I understand what is supposed to be problematic here. Suppose that in one outcome, everyone that exists suffers horribly from birth until death. In a second outcome, consisting of persons different from those in the first, everyone that exists lives a superb life from birth until death. The second outcome is clearly *better* than the first, even though it is *better for*

---

[5] For example, see Thomson 2001, 17-9, and 2008.

no one (the second outcome cannot be better for the people in it than the first, because

they do not exist in the first outcome). The second outcome is all things considered

better in a general and agent-neutral sense, or, if you like, it is better "from the point of

view of the universe,"[6] or, if you like, it is *better, period*.[7] I do not believe that this

concept can be analyzed much further, but there is one further thing I can add which

might be helpful. For the sense of "better than" which here concerns us, it is

necessarily true that if A is better than B, and if either could be brought about, then

there is more *reason* to bring about A than there is to bring about B, other things being

equal.[8] This is not what "better than" *means*, but it is a truth about "better than" which

might help us grasp the notion more clearly.

But it turns out that, even if it is nonsense to say that it is better that everyone

that exists lives a superb life from birth until death than that everyone that exists

suffers horribly from birth until death, nearly everything I argue in this dissertation

can easily be translated into *reasons*-talk or *ought*-talk. For instance, claims like

"benefiting person A would produce an all things considered better outcome than the

outcome that would be produced by benefiting person B," could instead often be read,

"we have more reason, all things considered, to benefit person A than to benefit

person B." Systematically making such translations throughout the whole dissertation

---

[6] Sidgwick 1907, 382.

[7] Arneson 2010.

[8] This claim, which I call the *Reason-Giving Thesis*, will become important in chapter 6.

would not, I believe, have any effect on the plausibility of my main claims and

arguments.

I just considered the worry that it is nonsense to say that one outcome is better

than another. A separate worry is that "better than" might apply very broadly – so

broadly, in fact, that it becomes more difficult to distinguish between

Consequentialism and its rivals, and to define the scope of the ethics of distribution.

According to:

> *Consequentialism*: an act is permissible if and only if it promotes an outcome
>
> at least as good as that which could be promoted by any alternative act.

Non-Consequentialists are those who deny Consequentialism.[9] The outcome

in which one person dies is better than the outcome in which five people die, other

things equal. Thus, Consequentialists claim that it is permissible to kill one person if

it is the only way to save five, whereas most Non-Consequentialists deny this. The

latter need not deny that the act of killing the one person promotes a better outcome,

but they do deny that it is a permissible act.

But it is possible to deny that the act of killing the one person promotes a better

outcome, and to claim, for example, that acts of killing make the outcome far worse

than do letting people die. If we claimed this, we could say that killing the one person

does not result in the better of two possible outcomes. Note that our claim here would

---

[9] For defenses of Non-Consequentialism, see Thomson 1992 and Foot 2002. For a defense of
Consequentialism, see Kagan 1989. Also see Unger 1996 for Consequentialism-supporting arguments.

not be that killing *causes* worse outcomes than does letting die (e.g., by somehow, in the long run, resulting in more deaths). Instead, we would be claiming that killing is partly *constitutive* of the overall goodness of an outcome. We would be claiming that the outcome in which one person is killed is worse than the outcome in which five people are allowed to die, other things being equal.

We might, in this way, attempt to defend Consequentialism from various possible objections. Possible objection: "here is an act which promotes a better outcome, but seems impermissible." Proposed reply: "you are conceiving of *better* too narrowly – this act actually promotes a worse outcome, and is for that reason impermissible." But even if it were plausible to expand the notion of "better than" very broadly in this way, doing so would not support traditional Consequentialists in any important way. That is because, traditionally, participants in the debate between Consequentialism and Non-Consequentialism took it for granted that, for example, killing one to save five results in a better outcome. If the only way traditional Non-Consequentialists could claim that it is impermissible to kill one to save five is by becoming newfangled Consequentialists, claiming that the outcome in which the one is killed is worse, then their debate with traditional Consequentialists would not disappear. It would only be relocated; the new debate would become: *is* the outcome in which the one is killed really worse?

It is widely recognized that how people are faring in some outcome is a factor relevant to how good that outcome is. If it were the only such factor, the outcome in which five die is worse than the outcome in which one dies. But we might claim that

it is also a relevant factor, whether an agent kills someone or instead lets someone die. Traditionally, Consequentialists have denied that this factor has any relevance, and, traditionally, Non-Consequentialists have claimed that it is factor relevant to the *permissibility* of acts, but not to the goodness of outcomes. When we instead claim that such a factor is relevant to permissibility (only) *because* it is relevant to the goodness of outcomes, we *consequentialize* it.[10] There are many other such factors which we might consequentialize in this way. Consider: promise-keeping, punishing the guilty, and respecting autonomy. We might claim, for instance, that when it is impermissible not to punish the guilty, this is (only) because outcomes in which the guilty are not punished are made worse by the fact that the guilty are not punished.

I will not here take a stand on which factors it is plausible to consequentialize, or how to formulate the distinction or debate between Consequentialists and Non-Consequentialists. Instead, I will simply indicate below which factors fall within the scope of the ethics of distribution, and repeat that it does not matter, for most of my purposes here, whether the discussion is formulated in terms of *better*, *reasons*, or *ought*.

1.1.4 The Ethics of Distribution: Relevant Factors

---

[10] See Dreier 1993, Schroeder 2007, and Portmore 2007.

The defining question of the ethics of distribution is what makes a given

distribution of well-being across persons better than another such distribution. Thus,

we are here abstracting from factors besides well-being and its distribution. Thus, in

the following examples, one outcome is better than another if and only if it contains a

better distribution of well-being.[11] It seems that, at least when other things are equal,

we can make a distribution better by adding more well-being to it. We can do this by

benefiting people, or increasing their level of well-being. There are two broad sorts of

issues I will explore: first, the *weight* that benefits should receive, and second, how to

aggregate size and number of benefits. On the first sort of issue, we might claim that

a distribution in which a badly off person receives a benefit is better than an otherwise

similar distribution in which a well off person receives a slightly larger benefit. We

might claim this because we believe that, though it is a *smaller* benefit, the former

benefit is more morally important, or makes a greater contribution to the overall

goodness of the distribution, in virtue of the fact that its recipient is worse off. And

we might claim this because we believe that, other things equal, a more *equal*

distribution of well-being across persons is better. If we claimed this, we would be

denying what I will call the:

---

[11] We might accept *Welfarism*: outcomes are better only insofar as they contain a better distribution of well-being. If we did, then we would not need to *abstract from* factors other than well-being and its distribution for it to be true that, in the examples discussed below, outcomes are better only insofar as they contain a better distribution of well-being. For, according to Welfarism, such other factors are irrelevant to how good outcomes are.

Equal Weight View: when it comes to distributing benefits, only their size and

number matters, facts about the possible recipients of these benefits do not

ultimately matter.

I will have more to say about this view in chapters 2, 3, 4, and 7. On the

second, aggregating, sort of issue, we might claim that it is, other things equal, better

to bring about a sufficiently large benefit than it is to bring about any number of very

small benefits. If we claimed this, we would be denying what I will call the:

Total View: other things equal, the outcome with more total (weighted) well-

being is better.

Just as the Equal Weight View is purely a claim about weighting benefits, and

takes no stand on how to aggregate them, the Total View is purely a claim about

aggregating benefits, and takes no stand on how to weight them. I will have more to

say about the Total View, and the distinction between aggregating and weighting, in

chapters 4 and 7. My main focus in this dissertation, however, will be on weighting,

i.e., on the relative plausibility of the Equal Weight View and its rivals.

If we deny the Equal Weight View, we claim that there are facts about the

recipients of benefits that can give benefits more or less importance, in addition to

their size and number. In the ethics of distribution, we abstract from factors besides

well-being and its distribution. Thus, the rivals to the Equal Weight View I will here

discuss claim that it is facts *about how well off* various possible recipients of benefits are, or would be, that affects the importance of benefiting them. There are further facts about possible recipients of benefits which some claim have an effect on the importance of benefits they might receive. For instance, there are views which place greater weight on achieving benefits for persons over sentient non-persons (e.g., cats, chickens, and cows),[12] and there are views which place greater weight on achieving benefits for the more virtuous over the less virtuous, or the vicious.[13] But, as the basis of weighting in these latter two cases is not how well off the recipients are, I will set such views aside. If it helps, we might imagine that the individuals in the examples below are equally virtuous persons.

A very important question, which Parfit and many others have taken up, is what we should do in situations where we can affect the identities and number of future people.[14] For the sake of simplicity, I will here largely limit myself to *Same People Choices*: choices which affect neither the identities nor the number of future people.[15] But I believe that much of the discussion in this dissertation can and should be extended to choices which affect the identities and number of future people.

Lastly, I should note that, while the ethics of distribution has a fairly limited focus, this does not mean that it is of correspondingly limited relevance. On the

---

[12] See Singer 1975, Steinbock 1978, DeGrazia 1996, and Arneson 1999.

[13] See Feldman1997, Kagan 1999 and 2012, and Arneson 2007.

[14] Parfit1984, part four. For a useful collection of essays, see Ryberg and Tännsjö 2004. And for a recent exploration of the topic, see Arrhenius 2012.

[15] Parfit 1984, 356.

contrary, the ethics of distribution seems relevant to nearly all versions of Consequentialism and Non-Consequentialism alike. Most of us correctly believe that at least *one* thing that matters, both to the permissibility of acts and to the justice and moral desirability of various policies, is the goodness of the outcomes these acts or policies (expectably) lead to. As John Rawls wrote, "All ethical doctrines worth our attention take consequences into account in judging rightness. One which did not would simply be irrational, crazy."[16] I agree, and I believe it would be equally irrational not to take the distribution of well-being into account, in judging whether a consequence or outcome is better than another. What makes one distribution of well-being better than another is, on any non-crazy view, an important question.

One might agree with me and Rawls on the above points, but nonetheless worry that the views within the ethics of distribution which seem most plausible when we are abstracting from factors other than well-being and its distribution might seem less plausible than rival such views when other "external" factors are brought into play (e.g., considerations of desert, autonomy, or deontological constraints). For example, perhaps the Equal Weight View seems less plausible than its rivals when abstracting from factors other than well-being and distribution, but seems more plausible than its alternatives when considerations of deontological constraints are brought into play. I do not deny the possibility of such contextual interactions between factors internal and external to the ethics of distribution.[17] But, for now, I can only do so much. I am

---

[16] Rawls 1971, 30.

[17] On such contextual interactions in ethical theory, see Kamm 1983 and Kagan 1988.

happy to go back through the critical steps and arguments made in this dissertation, bringing various such external factors into play, to see what effect, if any, doing so would have on the plausibility of these steps and arguments. Moreover, I should reiterate that there seem to be many actual and possible situations in which the only relevant factors in play are well-being and distribution. And it seems a worthwhile question, what is it better, or ought we, to do in such situations.

1.1.5 Practical Importance

The ethics of distribution is not merely important in theory, but also in a wide range of real-world contexts. Let me give just three quick examples.

First, the ethics of distribution matters to a variety of policy decisions in healthcare. This is true in many countries, including the US; the UK's *National Health Service* (NHS) provides a clear example. The NHS measures well-being in terms of "quality-adjusted life years," which are the product of time and quality of life, measured as a fraction of full health. It currently decides which medical treatments to fund on the basis of the total unweighted well-being they are likely to bring about. It thus adopts both the Equal Weight View and the Total View. But both views are highly controversial in contemporary ethical theory.

Second, the ethics of distribution matters to philanthropists whose aim is to do the most good with their donations. Such philanthropists not only need to know what

the effects of their donations will likely be, but how good those effects are. Perhaps

one charity is highly effective at preventing deaths from malaria, whereas another

charity, for the same amount of money, could prevent a much larger number of people

from suffering from schistosomiasis. This might be a situation where it would be

useful to have a defensible philosophical theory for comparing fewer large benefits

with many smaller benefits. Similarly, some of these charities might be able to help

moderately badly off people to a greater extent than they can help the extremely badly

off.[18]

      Third, insofar as the ethics of distribution can be applied to choices which

affect the identities and number of future people (and I believe it can), it is relevant to

measuring the costs of climate change, and more broadly to the question of how to

weigh the interests of future generations against those of our own. For instance, it

seems implausible to put less weight on the interests or well-being of future persons

merely because of their location in time.[19] That seems just as implausible as claiming

that spatial location, as such, makes a morally relevant difference. However, we might

be able to make reasonable predictions about how well off future persons will be.

Insofar as they will be much better off than we are, and insofar as a more equal

distribution of well-being across persons is better (or insofar as benefiting the worse

---

[18] For two independent and rigorous charity effectiveness evaluators, see *Give Well* (givewell.org) and *Giving What We Can* (givingwhatwecan.org). For further distributive issues that arise in this context, see Hare 2012.

[19] See Rawls 1971, 293-298, and Parfit 1984, Appendix F.

off matters more), there might be a plausible case for putting less weight on the interests of future persons.[20]

There are, I believe, many other ways in which the ethics of distribution is practically important. In general, this area of ethics has the potential to be *extremely* important. Consider a distributive policy which is only somewhat evaluatively inaccurate. Perhaps it evaluates outcomes as better, the more total *weighted* well-being they contain. The policy gives somewhat greater weight to benefits to the worse off. But suppose, for illustrative purposes, that the truth is that outcomes are better, the more total *un*weighted well-being they contain. Then each time we follow the inaccurate policy, we bring about a somewhat worse outcome than we could have if we had followed the accurate one. But perhaps this inaccurate policy, over the course of several years, affects billions of people. The *total* effect of adopting this policy, then, might be that we bring about a *much* worse outcome than we could have if we had instead adopted the accurate one. Since such distributive policies might affect very many people, it seems they should be informed by the most careful studies in the ethics of distribution.

1.2   Impartiality, Equal Weight, and Intuitionism

---

[20] On the future generations and giving priority to the worse off, see Adler 2009 and Holtug 2010. On climate change and weighing well-being, see Broome 1992 and 2012.

In the ethics of distribution, we are concerned with what makes some outcomes (distributions of well-being) better than others. We are concerned with what makes them better *period*, or better from the point of view of the universe. It thus seems a precondition on any minimally plausible theory within the ethics of distribution that it be appropriately *impartial*.[21]

1.2.1 Impartiality and Equal Weight

It would seem to be an extremely implausible view of what makes one distribution of well-being better is that it is the distribution in which Anton Chigurh is better off. This would be the better distribution from Anton Chigurh's point of view, but how would that make it the better distribution from the point of view of the universe, or the better distribution, period? Certainly it would not. Similarly, Rawls maintains that "[Ethical] principles should be general. That is, it must be possible to formulate them without the use of what would be intuitively recognized as proper names, or rigged definite descriptions."[22]

Suppose we instead adopted the view that it is always better to achieve a benefit, however small, for someone with a higher IQ, rather than achieve a benefit, however large, for someone with a (slightly) lower IQ. This view avoids making any

---

[21] For a compact and useful discussion of some different senses of impartiality, see Hooker 2000, 23-9.

[22] Rawls 1971, 131.

essential reference to proper names and definite descriptions, yet it is clearly not appropriately impartial. But *how* does this view fail to be appropriately impartial?

One answer is that this view fails to be appropriately impartial because it makes essential reference to facts about the potential recipients of benefits, i.e., their IQ. If this answer is correct, then it seems that any view in the ethics of distribution other than the Equal Weight View fails to be appropriately impartial. This is because the Equal Weight View is the only such view which makes no essential reference to facts about the potential recipients of benefits; it claims that all that matters is the size and number of benefits, and that facts about the potential recipients of these benefits are ultimately irrelevant. We might think that this is therefore a (non-decisive) point in favor of the Equal Weight View.[23]

But it might be that the view which gives absolute priority to benefiting people with higher IQs fails to be appropriately impartial for a different reason. It does not so fail because it makes essential reference to *facts about potential recipients* of benefits; rather, it so fails because it makes essential reference to *irrelevant* facts about potential recipients. We might believe, on the other hand, that certain facts about how well off potential recipients are (relative to others) *are* relevant, in determining which distributions of well-being are better, period. And we might believe that *not* taking such facts about individuals into consideration is a failure to be appropriately impartial.

---

[23] Hooker 2000, 59-65 suggests this.

The next obvious question here is: how can we tell which factors are relevant, which it is (not) appropriately impartial to consider, and which should be taken into account in determining which distributions of well-being are better? The answer, I believe, is that we can tell by reflecting carefully about the full range of possible views in the ethics of distribution, and the intuitive plausibility of the implications these views have in a variety of actual and hypothetical cases. I shall now say a bit more about the intuitionist methodology I here assume.

1.2.2   Normative Intuitionism

I accept a view in moral epistemology which plays a substantial role in this dissertation. It, or something close to it, is fairly widely accepted in contemporary ethical theory.[24] It is:

Normative Intuitionism: if we have the intuition that some normative

proposition is true, we thereby have epistemic reason to believe this

proposition, absent sufficiently strong defeaters.[25]

---

[24] Though Normative Intuitionism is often paired with a foundationalist epistemology (which I myself am attracted to), I believe that what I say below in the main text can be given a coherentist spin. For example, the main thoughts offered below can be embraced by fans of *Reflective Equilibrium* (see Rawls 1971). And certainly the main substantive ideas and arguments offered in this dissertation do not depend on adopting foundationalism over coherentism; Normative Intuitionism in a fairly wide and ecumenical form is sufficient for my purposes here.

[25] For further explication of this view (and the terminology offered below), see Huemer 2005. Also see McMahan 2000, Audi 2004, and Crisp 2006b, chapter 3.

I cannot argue for Normative Intuitionism here, but there are some terms that need to be defined in order to understand it:

- *Normative propositions* are propositions which have normative or evaluative content, e.g., "A is *better* than B," "It is *wrong* for P to do A," or "We have *reason* to promote O." We can contrast them with non-normative propositions, e.g., "It is sunny outside," or "There are no beards," or "That's a pink elephant."

- *Intuitions* are mental states about the truth values of propositions. An intuition that proposition P is true (or "that P," for short) is a mental state of its *seeming* to be the case that P. The intuition that P is distinct from the *belief* that P. For example, it might seem or appear to me that it is false that I ought to push the large man in front of the trolley to save the five, but at the same time it might also seem or appear to me, in another and possibly stronger way, that it is true that I ought to push the large man to save the five, and I might well *believe* this.

- When we have more epistemic reason overall to believe a normative proposition than not, we are *justified* in believing that proposition.

- There are two ways that the intuition that P can fail to justify the belief that P. This intuition could face either a,

- *Rebutting defeater*: an epistemic reason not to believe P, or an,

- *Undercutting defeater*: an epistemic reason to believe that the intuition that P gives no epistemic reason to believe P.[26]

It may help to give some examples, illustrating rebutting defeaters and undercutting defeaters, respectively. Suppose it seems to you that, even if the only way to save five people were to push a large man in front of a runaway trolley, it would be wrong to do so. But suppose it also seems to you that it is always wrong to fail to do what would make the world go best. And suppose you believe that it would make the world go best if you pushed the large man. Your intuitions are in conflict. They rebut each other. If, however, one of these intuitions is sufficiently stronger, as a seeming state, than the other (e.g., it *really* seems wrong not to make the world go best), then you would be justified in believing its content, rather than that of the other, countervailing intuition.

To illustrate undercutting defeaters, let us consider a non-normative example. It seems to Maude that it's sunny outside. Nonetheless, it also seems to Maude that a reckless neuroscientist is tampering with her brain so as to cause her to have the seeming that it is sunny. The seeming that there is tampering going on threatens to undercut the seeming that it is sunny. It's not that it seems to Maude that it is *not* sunny (if it did, then that would be a rebutting defeater). It's just that it seems to Maude that her intuition that it is sunny does not give her reason to believe that it is sunny.

---

[26] On the distinction between rebutting and undercutting defeaters, see Pollock and Cruz 1999, 195-6.

Most defenders of Normative Intuitionism are willing to grant that intuitions do not always provide reasons for beliefs. With John Rawls, most would agree that we should distinguish between intuitions rendered "under conditions favorable for deliberation and judgment in general"[27] and those rendered under unfavorable conditions. For instance, Rawls notes that our capacity for rational moral judgment tends to be distorted or unreliable in circumstances in which we are hesitant, frightened, or excessively self-interested. Intuitions formed under such circumstances would seem to face undercutting defeaters. Those not formed under such particularly poor circumstances, and which we have consistently and firmly, we can refer to as *considered* intuitions. In general, the intuitions cited throughout this dissertation (and there are many of them!) are considered intuitions. They are not rendered under the particularly poor circumstances Rawls mentioned, and many of us have them consistently and firmly.

But that, by itself, does not mean that the intuitions here cited do not face *any* undercutting defeaters. Many of the intuitions here appealed to are about particular cases, in addition to those about more general or abstract principles. Some authors believe that we should put little or no epistemic weight on intuitions about particular, concrete cases.[28] While perhaps some such intuitions face undercutting defeaters (and for the very reasons these authors mention), I believe that it is a mistake to put less

---

[27] Rawls 1971, 48.

[28] For example, see Hare 1981 and Huemer 2008a.

epistemic weight on case-specific intuitions merely in virtue of the level of generality of their content.

Others might agree that there is no problem *in principle* with relying on intuitions about particular cases, but argue that intuitions about the *particular* particular cases I here rely on are unreliable. For example, many of the particular cases here discussed involve large numbers or durations. Some philosophers worry that our intuitions about these sorts of large number cases are not reliable. I will address this worry in 4.3.

Others still might worry about our ability to correctly imagine some of the most basic sorts of cases in the ethics of distribution. For instance, here is one such case: we can either bring about a benefit for a badly off person, or a slightly larger benefit for a well off person. Many of us find it intuitively plausible that it is better to achieve the smaller benefit for the worse off person. But Joshua Greene and Jonathan Baron have conducted an empirical study suggesting that many people have this intuition merely because they are incorrectly imagining the case; they incorrectly imagine that the worse off person receives a bigger benefit, despite the fact that this conflicts with the very description of the case.[29] Normally, the worse off benefit more from receiving *resources* (recall that resources generally exhibit diminishing marginal well-being), and apparently this tricks some of us to erroneously treat *well-being* as if it exhibited diminishing marginal well-being.

---

[29] Greene and Baron 2001. Peter Singer tells a similar debunking story about such intuitions in an interview with Alex Voorhoeve. See Voorhoeve 2009, 51-2.

We should take care not to confuse resources and well-being, and consequently

claim that it is better to benefit the worse off person for a confused reason. However,

not only can we distinguish the two, but at least those who are engaged in the ethics of

distribution are typically careful to do so. For example, Parfit writes, "For each extra

unit [of well-being] to be an equal benefit, however well off the recipient is, these

units cannot be thought of as quantities of resources. The same increase in resources

usually brings greater benefits to those who are worse off."[30] Parfit and many others

are, I believe, correctly imagining the most basic cases in ethics of distribution.

Insofar as they have the intuition that it is better to provide a slightly smaller benefit

for the worse off person, it is not because they are erroneously treating well-being as if

it exhibited diminishing marginal well-being. It is because they have the intuition that

well-being exhibits diminishing marginal *goodness* or *moral importance*. It is dubious

that that this sort of intuition is, in general, undercut. But what, if any, *rebutting*

defeaters this intuition faces, and how strong they are, will be a major question

throughout the dissertation.

1.3   Outlandish Cases

Extremely unrealistic, unlikely, and outlandish cases can supply powerful

counterexamples to ethical principles. They can do so as follows:

---

[30] Parfit 1991, 83.

(1) If principle P is true, then X is true in outlandish case C.

(2) X is false in outlandish case C.

So,   (3) Principle P is false. (1, 2, modus tollens)

Here is a familiar example of this sort of argument. Suppose someone puts

forth the ethical principle that it is never permissible to kill an innocent person. This

implies that it would be impermissible to kill an innocent person even if this were the

only way to prevent Paris from being destroyed by a nuclear weapon. But it seems

false that it would be impermissible to kill an innocent person if this were the only

way to prevent Paris from being destroyed by a nuclear weapon (incidentally,

Consequentialists and Non-Consequentialists tend to agree here).[31] Therefore, it

seems false that it is never permissible to kill an innocent person.

Here is a different example.[32] Suppose someone puts forth the principle that

an individual's right to life is violated only if that individual does not persist as the

same biological organism. This principle implies that, if an evil neuroscientist

reprogrammed an individual's brain such that all of her beliefs, desires, memories,

attitudes, personality traits, etc. were removed and replaced with completely different

---

[31] There are some Absolutist Non-Consequentialists, who would maintain that it is impermissible to kill an innocent person, *whatever* the consequences. Probably most contemporary Non-Consequentialists, however, are Moderates, claiming that while it would be impermissible to kill an innocent person to save *five*, it would be permissible to kill an innocent person to save *millions*. For discussions of Moderate versus Absolutist Non-Consequentialism, see (for example) Thomson 1992 and Kagan 1998, chapter 3.

[32] This example comes from Tooley 1972, 46. Thomson 1971 is a treasure trove of effective outlandish cases.

ones, the neuroscientist would not have violated this individual's right to life, assuming the same biological organism exists. *But* it seems false that the neuroscientist would not have violated this individual's right to life in this reprogramming case. Therefore, it seems false that an individual's right to life is violated only if that individual does not persist as the same biological organism. This right can also be violated if an individual's brain is reprogrammed in a way that completely destroys her personality.

There are two premises in these arguments appealing to outlandish cases: (1) if principle P is true, then X is true in outlandish case C, and (2) X is false in outlandish case C. Since the inference from (1) and (2) to (3) (principle P is false) is valid, we can protect P only by denying either (1) or (2). We could deny (2) by claiming that our intuitions about outlandish cases are unreliable, or anyway less reliable than our intuitions about ordinary cases. But it is unclear why the fact that an intuition is about an outlandish case would, by itself, render the intuition unreliable or less reliable. (*Star Trek* is one big outlandish case, but this alone does not prevent me from thinking coherent thoughts about the plot or characters). As long as we can grasp or understand the relevant details of the outlandish cases in question, it seems that our intuitions about them should not receive any less epistemic weight. Alternatively, we could deny (1). Sometimes, it is plausible to deny (1). But often it is not. For example, if P has a certain kind of implication in ordinary cases, it seems perfectly reasonable to ask why it would not have a similar implication in outlandish

cases (even very outlandish ones, which are merely logically possible). If there is no plausible answer to this question, denying (1) will be correspondingly implausible.

I include these remarks about outlandish cases because I make liberal use of such cases throughout the dissertation. It is easy to dismiss such cases as silly, bizarre, or irrelevant to real life. This is very often a mistake; an unjustified prejudice in favor of the familiar. It does not matter if we will never actually face such outlandish cases. They are nonetheless relevant to ordinary cases we often do and will face, insofar as they reveal why various principles which apply to these ordinary cases are false or implausible.

Chapter 2

Equal Weight, Equality, or Priority

CHAPTER OUTLINE

2.1   The Equality View

2.2   The Levelling-Down Objection and the Priority View

2.1   The Equality View

There are several competing views about what makes a distribution of well-being better than another. I will begin by considering three competing views about the *weighting* of benefits: the Equal Weight View, the Equality View, and the Priority View. In this chapter, I will illustrate some of the counterintuitive implications of the Equal Weight View, but show that its alternatives face a variety of challenges too.

2.1.1   Equality versus Equal Weight

According to the *Equal Weight View*, when it comes to distributing benefits, only the size and number of benefits matters. It does not ultimately matter who

receives them.[33] The distributive pattern of well-being across persons is of no

fundamental concern. The Equal Weight View is a central component of

Utilitarianism, though it could be accepted by Non-Utilitarians.

There are plenty of Non-Utilitarians, however, who object to Utilitarianism at

least in part *in virtue of* its espousal of the Equal Weight View. The Equal Weight

View shows no fundamental concern for how well-being is distributed across separate

lives. And yet, many people believe that it is more important to benefit the worse off

than it is to comparably benefit the better off. More precisely, many believe that it is

fundamentally bad that there is inequality in the distribution of well-being.

Defenders of the Equal Weight View need not deny that there are *non-*

fundamental ways in which inequality is bad. For example, inequality of *resources* is

bad when it results in less well-being than there might have been. Since resources

generally exhibit diminishing marginal well-being, there are Utilitarian reasons to

distribute resources equally. They are, however, highly contingent and limited; there

are myriad cases in which the Equal Weight View *would* deem a (very) unequal

distribution of resources to be better than an equal distribution of resources.[34]

There are various ways in which inequality of *well-being* can have bad effects.

For example, worse off individuals could become envious of the better off, or acquire

---

[33] As Parfit 1991, 101 puts it: "we should ... give equal weight to equal benefits, whoever receives them. Benefits to the worse off should [not] be given more weight." Utilitarians from Jeremy Bentham to Peter Singer have accepted the Equal Weight View. Note Bentham's Dictum, "everybody to count for one, nobody for more than one" discussed in chapter V, paragraph 36 of John Stuart Mill's *Utilitarianism* 1861/1998, and Singer's "Principle of Equal Consideration of Interests" in Singer 2011. But one could accept the Equal Weight View without being a Utilitarian, or even a Consequentialist.

[34] See, e.g., Broome 1991, 175-177.

servile or fatalistic inferiority complexes. Inequality can exacerbate or be indicative of

disutile social structures, which have weak or no bonds of inclusion, fraternity, and

sympathy. It is often indicative of corruption, manipulation, and exploitation. We do

not have to believe that inequality is *fundamentally* bad in order to regard these causes

and effects of inequality as bad. According to one explanation, these things are bad

simply because of their ill effects on overall well-being. But some people believe that

inequality is itself fundamentally bad, in addition to whatever bad effects it might

have. For this reason, these people may reject the Equal Weight View. Consider the

following case from Thomas Nagel:

> *City or Suburbs*.[35] Suppose I have two children, one of which is normal and
>
> quite happy, and the other of which suffers from a painful handicap. Call them
>
> respectively the first child and the second child. I am about to change jobs.
>
> Suppose that I must decide between moving to an expensive city where the
>
> second child can receive special medical treatment and schooling, but where
>
> the family's standard of living will be lower and the neighborhood will be
>
> unpleasant and dangerous for the first child – or else moving to a pleasant
>
> semi-rural suburb where the first child, who has a special interest in sports and
>
> nature, can have a free and agreeable life. This is a difficult choice on any
>
> view. To make it a test for the value of equality, I want to suppose that the
>
> case has the following feature: the gain to the first child of moving to the

---

[35] Quoted from Nagel 1979, 124-4.

suburb is substantially greater than the gain to the second child of moving to the city. After all, the second child will also suffer from the family's reduced standard of living and the disagreeable environment. And the educational and therapeutic benefits will not make him happy but only less miserable. For the first child, on the other hand, the choice is between a happy life and a disagreeable one. Let me add as a feature of the case that there is no way to compensate either child significantly for its loss if the choice favoring the other child is made. The family's resources are stretched, and neither child has anything else to give up that could be converted into something of significant value to the other.

In this case, the Equal Weight View would imply that it is better to move to the suburbs, because that would induce a larger benefit. However, many judge that it would instead be better to move to the city, thereby achieving a smaller benefit for the worse off child. Consider another case, this one from Parfit.

Divided World.[36] The two halves of the world's population are unaware of each other's existence. Each half lives on a separate continent, and each half is unable to cross the ocean separating them. Now there are two possible outcomes, (1) and (2). (See below). Each box represents one half of the population, living on a single continent. As indicated by the dotted line, the

---

[36] Quoted and paraphrased from Parfit 1991, 87.

average utility in (1) is slightly greater than the average utility in (2).

However, utility is distributed unequally in (1), and equally in (2). But, since

these halves are divided, the inequality in (1) has no bad effects on the worse-

off group, and the equality in (2) produces no desirable fraternal relations

between the two groups. Consider a diagram:
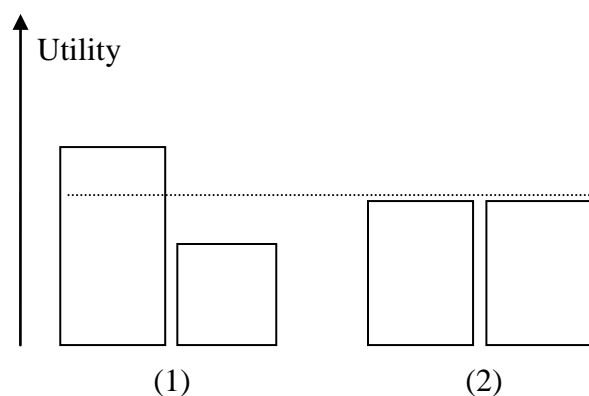
Figure 1: Divided World

Many believe that (2) is better than (1). Some might offer the further claim

that the *reason* (2) is better than (1) is that inequality is fundamentally or non-

derivatively bad, apart from its effects. It is possible, but fairly implausible, to claim

that the inequality in (1) is not fundamentally bad, *but* that inequality *within* groups or

communities is fundamentally bad. This would be fairly implausible since within

communities the badness of inequality can be explained in non-fundamental terms,

e.g., in terms of its bad effects or in terms of the social injustice it involves. It would

be a strange coincidence if inequality were only fundamentally bad when it occurred

within communities.[37] Divided World thus seems to be a good test case for whether inequality is *fundamentally* bad.

Defenders of the Equal Weight View agree that inequality is often bad in many non-fundamental ways. But people who believe that inequality is fundamentally bad will reject the Equal Weight View. They will claim that the size and number of benefits is not all that matters, and that it also matters whether benefits are distributed equally.

2.1.2   Equality is a Complex Ideal

No one seriously believes that equality is *all* that matters. If equality were all that mattered, it would be better that everyone suffer equally in hell than that everyone minus one person live in an equally great ecstasy, where the one person experiences a slightly greater ecstasy. But this is absurd.

It is very hard to deny that the Equal Weight View at least gets things partly right, that the size and number of benefits do matter. I will thus formulate the *Equality View* as follows:

In addition to the size and number of benefits, the degree to which benefits are distributed equally across lives also matters.

---

[37] Ibid., 88.

The Equal Weight View gives equal weight to benefits, no matter who receives them. In contrast, the Equality View places greater weight on benefits to those who are worse off, and less weight on benefits to those who are better off.

But equality is a complex moral ideal.[38] There are many particular versions of the Equality View that one might adopt. Let me quickly mention just a few of the kinds of questions that would need to be answered to arrive at a fully articulated version of the Equality View.

First, there is the question of *how much* weight to place on equality, relative to the size and number of benefits. For example, in Divided World, how much more well off than the worse off group in (1) would the better off group have to be for it to be the case that (1) is better than (2), despite the fact that (1) contains much more inequality than (2)? Or, in City or Suburbs, how much smaller would the benefit to the worse off child have to be to render it better to move to the suburbs rather than the city?

Second, there is disagreement about the units over which equality should operate.[39] As I formulated the Equality View, the relevant units are lives. But is it equality of well-being across *whole* lives, or equality across certain *parts* or *times* of lives that matters? Or do both matter? This raises the more general question: insofar

---

[38] For evidence that this is so, see Temkin 1993.

[39] See Nagel 1979, 121, McKerlie 1989 and 1997, and chapter 8 of Temkin 1993, for discussion of different possibilities regarding the appropriate temporal units. For related discussions in connection with personal identity, see chapter 15 of Parfit 1984 and Parfit 1986a.

as we are concerned about the distribution of well-being, is it the distribution across lives, parts of lives, groups of people, or what, that matters? In discussing various weighting principles throughout this dissertation, I sometimes assume (as many do) that the relevant units of distributive concern are *lives*. But this is largely just for convenience. In many cases, my arguments for or against these weighting principles do not crucially depend on what the distributive units are.

Third, there are different views about how to measure the size of inequality. One proposal is to find the average level of utility, calculate the difference between each person's utility and the average, and then sum these differences. Notice that, in Divided World, this will imply what is obvious: that there is more inequality in (1) than in (2). But there might be other situations in which this measure of inequality is not obviously correct. Moreover, we might distinguish between the *size* of inequality and the *badness* of inequality. For example, it might be that the inequality within a poor world, though the same size as the inequality within a rich world, is worse.[40]

Finally, what is the *scope* of the Equality View? I have assumed it a requirement on accepting the Equality View that one regard inequality as fundamentally bad. But among *which* people is inequality fundamentally bad? It is *possible* to claim that inequality is only fundamentally bad within communities, and not between unrelated communities who are unaware of each other's existence. Again, this seems implausible and strangely coincidental (since we could explain the badness of inequality within communities in non-fundamental terms). But just as it

---

[40] Again see Temkin 1993, especially chapters 5 and 6.

would seem strange if inequality mattered within but not between communities, it too would seem strange if inequality mattered among isolated groups of people that exist at the same time (as in Divided World) but not among people that exist at different times. If spatial distance should not make a difference, why should temporal distance? But the natural conclusion of this line of reasoning is that those people among whom inequality is fundamentally bad are *everyone* who ever lives.[41] But this conclusion has puzzling implications. It is puzzling that, in deciding how to distribute benefits now, it might be important that we carefully research how well off people in ancient societies, or those that may have lived a long time ago in a galaxy far far away, were.[42] If we found this implication puzzling, and we also thought that *if* inequality were fundamentally bad *then* the only plausible group among which it would be bad is everyone who ever lives, then we may find the Equality View itself to be puzzling or perhaps implausible. But I will not press this worry further here. There is another, more powerful objection to the Equality View.

2.2   The Levelling-Down Objection and the Priority View

---

[41] A further question is whether "everyone" should or should not include sentient non-persons. For fascinating discussions on the question of whether egalitarian principles (or, more broadly, distributive principles) should apply to non-persons, see: McMahan 1996, Arneson 1999, and Vallentyne 2005.

[42] Temporally neutral versions of Average Utilitarianism also have this puzzling implication (Parfit 1984, 420). Things get even more puzzling if there has been *infinite* well-being or suffering in the past.

The Equality View claims that inequality of well-being is fundamentally bad. It claims that, in addition to the size and number of benefits, the degree to which benefits are distributed equally across lives matters. The view may place enough weight on equality so that, in Divided World, it implies that (2) is better than (1), and that, in City or Suburbs, it is better to move to the city.

But it seems that, however little weight is placed on equality, the Equality View may yield implausible results in certain cases. Consider the following.

Levelling-Down. There are two same-sized populations, A and B. There are two possible outcomes, (3) and (4). In (3), A and B are not very well off, but they are equally well off. In (4), A is not very well off, but better off than in (3), and B is *extremely* well off. We are again abstracting from the bad effects that inequality might have.
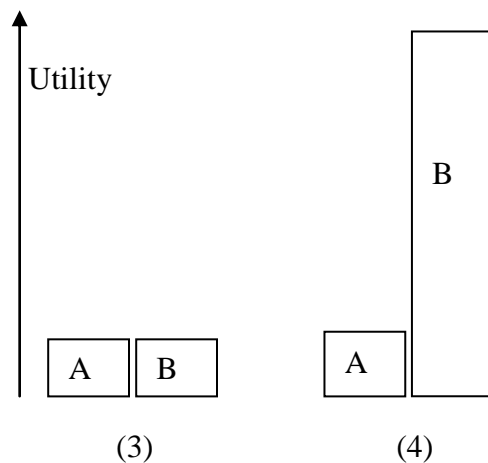
Figure 2: **Levelling-Down**

There is *enormous* inequality in (4). This may lead some versions of the Equality View to imply that (3) is better than (4). This, or so it seems to me, is implausible. (4) seems better, despite its enormous inequality.

Though some defenders of the Equality View would be not be bothered by the previous example, many defenders of the Equality View would agree that (4) seems better. They would claim that, because the benefits for B are *so* large, this outweighs the large inequality in (4). And they could point out that if the benefits for B were not large – but perhaps only big enough to make B slightly better off than A – then the inequality between A and B would be correspondingly small, and thus the small benefits for B would outweigh the small inequality in (4). It is important to notice the difference between this modest Equality View and the extremely modest view that any gain in utility, however small, would outweigh any inequality, however large.[43] The claim is rather that, in cases like Levelling-Down where the size of utility gains are proportional to the size of inequality, large utility gains always outweigh large inequalities, and small utility gains always outweigh small inequalities.

Of course, this modest Equality View is still committed to the claim that inequality is fundamentally bad. Thus, while the modest Equality View could weight inequality and overall well-being such that it avoids the *very* implausible implication that (3) is *all things considered* better than (4), it would still have to face the implausible implication that (3) is *in one respect* better than (4). In particular, (3) is better than (4) in the respect that it contains less (no) inequality. We can now state:

---

[43] Parfit 1991, 112-3.

The Levelling-Down Objection: the Equality View implies that there is a respect in which it is better to level down overall well-being. However, levelling-down overall well-being is never better, in *any* respect.[44] Hence, this view is implausible.

Some authors have argued that phrases like "better in any respect" are ambiguous, and that on any sensible interpretation of such phrases the Levelling-Down Objection either proves too much or too little.[45] I am not sure they are right. But to keep the discussion on track, we can simply grant that they are. Even if we could not appeal to *betterness in respects*, and even if the Equality View never had any *all things considered* implications which resulted in levelling-down, we might still find the *reasons* that it endorses to be implausible. Consider:

The Levelling-Down Objection*: the Equality View implies that there is *a reason* to level down overall well-being. However, there is *no* reason to level down overall well-being. Hence, this view is implausible.

The Equality View implies there is a reason to level down, since it regards inequality as fundamentally bad, and levelling-down from (4) to (3) would reduce

---

[44] Again, assuming we are holding other things, e.g., considerations of desert, constant.

[45] Broome 2002 and Brown 2005, sections 4.3 and 4.4.

inequality. But a move from (4) to (3) would be *worse for everyone*. Many of us find the Levelling-Down Objection* to be devastating because we cannot believe that we have *any* reason to bring about an outcome which is worse for everyone – at least, we cannot accept this for cases where we cannot affect the identities or number of people.[46] Many of us plausibly believe, *at least* when the only consideration in play is how well off various people could end up, that a reason in favor of bringing about an outcome must be grounded in some increase in well-being.[47] If we do think this, we may reject the Equality View, which necessarily entails that there *are* some reasons not so grounded in well-being.

Those of us who judged it better to move to the city in City or Suburbs, or that (2) is better than (1) in Divided World, might initially find it quite troubling that the Equality View is implausible. We might have thought that we have to appeal to the Equality View to accommodate these judgments. This is not so. We can instead appeal to the *Priority View*. This view, as formulated by Parfit, claims that "benefiting people matters more the worse off these people are."[48] That is, in addition to the size and number of benefits, it also matters how well off a potential recipient of a benefit would be without it. The worse off she would be without it, the more it matters that she receive it.

---

[46] Again, for discussion of the Non-Identity Problem in this and related cases, see chapter 16 of Parfit 1984.

[47] Cf. Roger Crisp's *Welfarist Restriction* from Crisp 2006b, 149.

[48] Parfit 1991, 101.

The Priority View and the Equality View will coincide in their implications in many cases. However, the Priority View does not claim that inequality is fundamentally bad. It thus entails *no* reason whatsoever in favor of levelling-down.

Next, it is important to understand the relationship between the Equality View, the Priority View, and two claims I call *Relative Matters* and *Absolute Matters*. The Priority View could deny:

> *Relative Matters* (henceforth *RM*): in addition to the size and number of benefits, it also matters how well off some are relative to others.

The Equality View entails RM, but RM does not entail the Equality View. One could claim that benefits to those who are worse off relative to others matter more, and that inequality *per se* does not matter at all.[49] So we need not deny RM in order to avoid the Levelling-Down Objection*. The Priority View, as formulated by Parfit, both denies RM and affirms:

> *Absolute Matters* (henceforth *AM*): in addition to the size and number of benefits, it also matters how well off people are in absolute terms.

---

[49] See Otsuka and Voorhoeve 2009, 182-4. There they discuss a view which espouses RM and denies that inequality is non-derivatively bad.

Insofar as it is bad that some people are worse off, it does indeed seem bad "that they are worse off than *they* might have been."[50] Parfit elaborates:

> [I]f I am worse off than you, benefits to me are more important. Is this because I am worse off than you? In one sense, yes. But this has nothing to do with my relation to you. It may help to use this analogy. People at higher altitudes find it harder to breathe. Is this because they are higher up than other people? In one sense, yes. But they would find it just as hard to breathe even if there were no other people who were lower down. In the same way, on the Priority View, benefits to the worse off matter more, but that is only because these people are at a lower *absolute* level. It is irrelevant that these people are worse off *than others*. Benefits to them would matter just as much even if there *were* no others who were better off.[51]

Just as Equality is a complex moral ideal, so too is Priority. We would need to answer analogs to several of the questions asked in 2.1.2 in order to arrive at a fully articulated version of the Priority View. For example, how much weight is it appropriate to place on benefits for the worse off, relative to the size and number of benefits? Many believe that we should at least give enough priority to benefiting the worse off to accommodate the judgments that it is better to move to the city in City or Suburbs, and that (2) is better than (1) in Divided World. It seems bad that people end up at lower absolute levels of well-being than they might have; benefits to such people matter more.

AM is consistent with RM. We could adopt a version of the Priority View that accepts either or both of these claims. There are several interesting and important

---

[50] Ibid., 104.

[51] Ibid., 104.

complications, concerning AM and RM, that have been relegated to Appendices A through C (hence the relative shortness of Chapter 2). These complications have been relegated to Appendices in order to avoid disturbing the main narrative and argumentative trajectory of the dissertation.

In Appendix A, I critically examine an argument for RM that does not appeal to the badness of inequality.

In Appendix B, I argue that it may be difficult to maintain RM without also accepting AM.

In Appendix C, I argue that, insofar as we accept the Equality View or the Priority View (and AM or RM), we may be forced to deny a plausible claim, called the Rational Prudence Constraint.

We began chapter 2 with a simple criticism of a simple view, the Equal Weight View. Recall that this view claims that only the size and number of benefits matter. Against this view, many believe that it *also* matters how well off potential recipients of these benefits are, or would be. But finding a plausible view which explains or captures this belief may turn out not to be a simple or easy task. I hope that Appendices A through C will illustrate this. Ultimately, to defend views which give priority to the worse off, we may have to deny plausible claims (the Rational Prudence Constraint, discussed in these Appendices). Some people may find this to be troubling. These people might claim that *the cure is worse than the disease*; that is, they might think that the proposed rivals to the Equal Weight View are, on balance, *more implausible* than the Equal Weight View itself. But I do not think anything I

have said in this chapter or its corresponding Appendices shows that this is true. At most, I have shown that these rivals to the Equal Weight View are less plausible than the Equal Weight View *in some respects*. Defenders of the Priority View or the Equality View could defensibly claim that, though their cure has some less than ideal side effects, it is not worse than the disease. In chapter 3, however, I argue that the rivals to the Equal Weight View here discussed have significantly worse side effects than any of the ones mentioned so far.

Chapter 3

Priority Monsters

CHAPTER OUTLINE

3.1   Introduction to the Argument

Other things being equal, is it better to relieve a badly off person of a serious illness or to treat the mild sore throat of a much *much* worse off person? As much sympathy as we no doubt have for the second person, it is plausible that it is better to provide the considerably larger benefit for the first, less badly off, person. I here argue that it is difficult for the Priority View to avoid the implausible conclusion that it is better to provide piddling benefits to *extremely* badly off people ("*priority*

monsters")[52] than it is to provide very large benefits for other people who are themselves significantly badly off. It can only avoid this conclusion at the cost of rejecting other claims which are independently plausible.

3.1.1   The Scope and Aims of the Priority Monster Spectrum Argument

The Priority Monster Spectrum Argument is the argument I will give in 3.2. There are some preliminaries to be discussed before then. First, I want to remind the reader that I am here holding constant potentially morally relevant factors other than the size and number of benefits, and how well off potential recipients of these benefits are prior to receiving them (for example, I am bracketing considerations of desert, autonomy, and special relationships).

Recall that the Equal Weight View claims that *only* the size and number of benefits matters; the Priority View and the Equality View deny this, and claim that it *also* matters how well off the recipients of these benefits are (in absolute or in relative terms). The Equal Weight View has implausible implications. Consider:

---

[52] Priority monsters were inspired by Robert Nozick's "utility monster" from Nozick 1974, 41. The utility monster is a hypothetical being that eats people – but in eating them, the monster gains vastly more well-being (or utility) than its victims could have ever achieved. I discuss this case further in section 4.1.

Lollipop. We could achieve a very large benefit for a badly off person, or a same-sized benefit *plus* a piddling benefit – say, the pleasure of one lollipop lick – for a very well off person.

It seems better to do the former. Both the Equality View and the Priority View capture this judgment; the Equal Weight View does not. The Equality View and Priority View thus have considerable intuitive appeal. But for all that, I will argue that they have implausible implications which are serious. Thus I will argue that the Equal Weight View is at least in one important respect, if not overall, more plausible than the Equality View and the Priority View.   (As I explain in 3.6.1, my argument targets the Equality View no less than it targets the Priority View; but for the sake of convenience, I will largely present and discuss it as an argument against the Priority View).

Next, it is important to distinguish briefly the topic of this chapter from the familiar topic of how to combine, or *aggregate*, the size and number dimensions of benefits. Aggregative principles answer questions like, "How many benefits of some particular smaller size would be (roughly) as good as some number of benefits of some particular larger size?" Some philosophers believe that we should resist certain kinds of aggregation. For example, they believe that there is no number of trivial benefits, such as lollipop licks, that could ever be as good as just one big benefit, such as the delight or personal achievement of playing a brilliant game of chess. They thus find standard Utilitarian views about aggregation, which entail that very large benefits

or burdens can always be outweighed by *enough* trivial benefits or burdens, spread across separate lives, to be implausible.

But questions of weighting benefits are different from questions of aggregating benefits. It would be a mistake, for example, to reject the Equal Weight View because one finds standard Utilitarian views about aggregation to be implausible. The Equal Weight View does not by itself take a stand on the question of aggregation. It says that *however* size and number of benefits should be aggregated (intrapersonally or interpersonally), *only* size and number of benefits matters. The Priority View, as I here understand it, takes the same neutral stance on aggregation, but says that in addition to size and number of benefits, it also matters how well off their recipients are or would be.

In chapter 4, I will return to the topic of aggregation, as well as the further important topic of how weighting issues might *interact* with aggregation issues. For now, however, the concern is principally with weighting benefits for prioritarian (or egalitarian) reasons, and not with aggregating benefits in the sense mentioned above.

3.1.2 The Maximin View

Is it better to relieve a significantly badly off person of a great amount of suffering or to relieve a much *much* worse off person of a very tiny amount of suffering? As I said, it is no doubt appropriate to have greater sympathy for the

second person. But if we reflect carefully on this question, we find a plausible answer:
it is better to provide the considerably larger benefit for the first, less badly off,
person. It is quite difficult to deny this claim.

The purpose of this subsection is to acknowledge that there is a widely known
view which is inconsistent with this crucial claim of this chapter. This view is:

Maximin: The best outcome, of some set of possible outcomes, is the one
whose worst-off people are better off than the worst-off people of all the other
possible outcomes in this set.[53]

As I noted in chapter 2, it is a question for every version of the Priority View,
"How much priority should we attach to benefiting the worse off?" According to
Maximin, we should attach *absolute* priority to benefiting the *worst* off. Maximin
would therefore imply that it would be better to make a very minimal improvement in
the quality of life of an extremely badly off person than to make a very large
improvement in the quality of life of another significantly, though not extremely,
badly off person (assuming the former person is the worst off). But, it seems to me, so
much the worse for Maximin.

Maximin faces bigger problems still. It has the following *absurd* implication:

---

[53] Parfit 1984, 422. See the Appendix of Parfit 1991 for a useful comparison of Maximin (as an
implausibly extreme version of the Priority View) with John Rawls's *Difference Principle*, as defended
in Rawls 1971.

It would be better to achieve an arbitrarily small benefit for the worst off

person than to achieve arbitrarily large benefits for an arbitrarily large number

of people who are just *slightly* better off than the worst off person.

Maximin is therefore unacceptable.[54] Again, the purpose of this subsection is

merely to acknowledge that there is a widely known view which is inconsistent with

the claim that it is better to provide large benefits for significantly badly off people

than to provide very small benefits for *extremely* badly off people. But this view is

unacceptable, and it is very hard to deny this claim. We can proceed undeterred by the

fact that one of the crucial claims of this chapter is opposed by Maximin.

3.2   The Priority Monster Spectrum Argument

Consider a putative counterexample to the Priority View:

Priority Monster. There could be a person so badly off that, if she existed, it

would be better to achieve an arbitrarily small benefit for her than to achieve

an arbitrarily large benefit for the currently worst off person.

---

[54] Fortunately, most people agree that it is unacceptable. See Temkin 1993, 103-5, Nagel 1979, 125, and Brink 1993, 265-270.

If the Priority View implied Priority Monster, it would seem to be an implausible view. Indeed, many defenders of the Priority View would rather avoid this implication. Can they avoid it? I will argue that they can, but only at a significant cost.

The rest of this section contains, in a very dense nutshell, my main argument against the Priority View. I will discuss some importantly distinct arguments against the Priority View in chapter 4; but by and large the remainder of the dissertation will clarify the following argument, defend it against a variety of possible objections, compare it with other structurally similar arguments, and determine its overall philosophical significance.

Consider the following finite *Series* of possible benefits:[55]

(1): an arbitrarily large benefit for the currently worst off person.

(2): a benefit slightly smaller than the one in (1) for a person much worse off than the one in (1).

(3): a benefit slightly smaller than the one in (2) for a person much worse off than the one in (2).

And so on… all the way up to,

---

[55] Note that, given the way the Series is set up, it is not possible to object that the below argument merely raises a version of Zeno's Paradox. Cf. Voorhoeve and Binmore 2003. Also see Appendix D.

(*n*): an arbitrarily small benefit for an arbitrarily badly off person.

There are four variables that can be adjusted, in arriving at a precise formulation of the Series: the largeness of (1), the tininess of (*n*), what counts as "slightly smaller," and what counts as "much worse off." Just as (1) can be arbitrarily large and (*n*) can be arbitrarily small, the size difference between adjacent benefits can be arbitrarily small and each person can be arbitrarily worse off than her predecessor. Next consider the following *Premises*:

P1. Benefit (2) is better than benefit (1).

P2. Benefit (3) is better than benefit (2).

And so on… all the way up to,

P*n*. Benefit (*n*) is better than benefit (*n*-1).

And consider:

> *Transitivity*: If A is better than B, and B is better than C, then A is better than C.[56]

---

[56] Again, "better than" is here taken to mean "*all things considered* better than."

Assuming Transitivity, the Premises together entail the conclusion that benefit
(*n*), an arbitrarily small benefit for an arbitrarily badly off person (i.e., a "priority
monster"), is better than (1), an arbitrarily large benefit for the currently worst off
person. But again, Priority Monster seems implausible. To avoid Priority Monster,
the Priority View will have to deny at least one of the Premises, no matter *how* the
four variables are adjusted or *what* the precise formulation of the Series is. Can the
Priority View do this? Yes, but only at a significant cost. Consider the following:

- *Tradeoffs*: the Priority View is, at a minimum, committed to at least some
  claims of the form **Ω**: *it is better to achieve a smaller benefit for a worse off*
  person than to achieve a larger benefit for a better off person. It is committed
  to such claims on the basis of their intuitive plausibility. (Recall what it is
  intuitively plausible to say in cases like Lollipop).

- *Parity*: the size difference between adjacent benefits can be *sufficiently* small,
  and each person can be *sufficiently* worse off than her predecessor, such that
  each of the Premises would seem, considered independently, at least as
  intuitively plausible as any other claim of form **Ω**.[57]

---

[57] This Parity claim is not to be confused with the evaluative relation Chang 2002 calls parity, or "being
on a par with."

- The *Conditional*: if we accept claim C1 on the basis of its intuitive

  plausibility, and if claim C2 is at least as intuitively plausible as C1, and if

  other things are equal, then we should also accept C2.

If these claims are true, then the Priority View implies the Premises (at least on

some formulations of the Series). The problem is that the Premises together imply

Priority Monster, and both Parity and the Conditional seem plausible. The Equal

Weight View can avoid Priority Monster without denying Parity or the Conditional,

since it accepts no claim of form Ω. It therefore seems in one way more plausible than

the Priority View.

The Premises and Transitivity together have the structure of what we can call a

Spectrum Argument.[58] These arguments have the following structure:

P1.   B is better than A

P2.   C is better than B

P3.   D is better than C

…   …

P(n-1).   Y is better than X

Pn.   Z is better than Y

---

[58] I borrow this term from Larry Temkin. In Temkin 1996, Temkin referred to these arguments as Continuum Arguments. In Temkin 2012, he calls them Spectrum Arguments. Temkin's Spectrum Argument, involving a spectrum of painful experiences, initially came from Stuart Rachels, whose version of the argument was later published in Rachels 1998. I believe the very first Spectrum Arguments of any sort appeared in Part IV of Parfit 1984, though Parfit did not call them by this name.

P*T*.   Transitivity

∴   Z is better than A

This overlapping chain of premises (P1 through P*n*) and Transitivity (P*T*) together imply that Z is better than A. Although some people reject Transitivity, for the time being I will accept it as an undefended assumption. I will defend Transitivity, or at least something close enough to it, in chapter 6. A perhaps more common objection to Spectrum Arguments is that they are, or are relevantly like, *Sorites Arguments*, which are notoriously unsound. This objection will be addressed in chapter 5. For the remainder of this chapter, I will focus on issues and objections which are relevant to the Priority Monster Spectrum Argument in particular, rather than to Spectrum Arguments in general.

3.3   Tradeoffs, Parity, and the Conditional

It may help to repeat briefly the alleged problem for the Priority View. According to Tradeoffs, defenders of the Priority View accept, on the basis of their intuitive plausibility, some claims of form **Ω**, e.g., in reaction to Lollipop. (A claim of form **Ω** says that it is better to achieve a smaller benefit for a worse off person than to achieve a larger benefit for a better off person). Parity says that each of the Premises can be made as intuitively plausible as any claim of form **Ω**. Given Parity, the

Conditional implies that if one accepts claims of form Ω on the basis of their intuitive plausibility, then one should accept each of the Premises. These claims entail that defenders of the Priority View should accept each of the Premises. But the Premises imply Priority Monster, which is implausible. To avoid Priority Monster, one must either reject the Priority View, or maintain the Priority View while rejecting Tradeoffs, Parity, or the Conditional. But Tradeoffs, Parity, and the Conditional are each plausible. So defenders of the Priority View can avoid Priority Monster only by denying a claim which is independently plausible.

It is possible to underestimate the number of views the Priority Monster Spectrum Argument puts pressure on. It would be a mistake, for example, to think that it only narrowly targets standard formulations of the Priority View, or formulations that have a simple shape. The Argument targets *all* versions of the Priority View in that it forces them *either* to accept Priority Monster, *or* to deny some independently plausible claim (e.g., Parity or the Conditional) in order to avoid it. It also targets other, similar, views (3.4.3) as well as the Equality View (3.6.1).

I will now defend Tradeoffs, Parity, and the Conditional.

3.3.1 Tradeoffs

Tradeoffs has two parts: first, that the Priority View accepts claims of form Ω, and second, that it accepts these claims on the basis of their intuitive plausibility.

I am simply assuming that the first part of Tradeoffs is true. That is, I am assuming that the Priority View goes beyond the trivially modest view that considerations of priority for the worse off are to function merely as "tie breakers." Such a view would deem it better to benefit a worse off person than to benefit a better off person only if the benefit to the worse off person were of equal or greater size than the benefit to the better off person.[59] But because this tie breaker view only trivially differs from the Equal Weight View, I will ignore it, and assume that the Priority View entails that it is at least sometimes better to trade off a little well-being for the sake of giving priority to the worse off. That is, that it entails some claims of form Ω.[60]

What about the second part of Tradeoffs? Defenders of the Priority View almost always defend their view by appealing to intuitions about cases like Lollipop, City or Suburbs, Two Children, and Divided World, or by appealing to the intuitiveness of claims of form Ω, considered more abstractly or generally. But this last sentence only points out some contingent historical facts. One might go farther, and claim that the Priority View *could* not be defended without appealing to the

---

[59] Cf. Henry Sidgwick's (1907, 416-7) claim that to determine "which mode of distributing a given quantum of happiness is better" we should choose the outcome in which the sum of happiness is more equally distributed across individuals. For Sidgwick, the Equality View serves as a tie breaker, and total well-being (according to him, total pleasure minus total pain) is never to be sacrificed for the sake of reducing inequality.

[60] As I have already indicated, I am here principally concerned with all things considered betterness. A tie breaker view of the sort mentioned above might converge with the Equal Weight View on the claim that it is never *all things considered* better to trade off well-being for the sake of giving priority to the worse off, but insist that it is nonetheless at least sometimes *in one respect* better to do so. If we count this tie breaker view as a version of the Priority View, and the Equal Weight View is incompatible with the claim that it is sometimes better in one respect (but never all things considered) to trade off well-being for the sake of giving priority to the worse off, then I concede that there is a version of the Priority View which my arguments do not clearly show to be less plausible (in any significant respect) than the Equal Weight View.

intuitiveness of some claims of form **Ω**. But while it seems true that to defend any substantive claim in ethics one must appeal to intuition at some point, a defender of the Priority View could argue that she accepts claims of form **Ω**, *not* because of *their* intuitiveness, but because they follow from other claims which are intuitively plausible.

For example, one might argue for the Priority View by appealing to what it would be rational to agree on from behind a veil of ignorance.[61] Instead of claims of form **Ω**, one might find claims of the form **ψ**, *it is rational for one to prefer a slightly smaller benefit in a possible scenario in which one is worse off to a slightly larger benefit in a possible scenario in which one is better off*, to be intuitively plausible. Since while in the hypothetical contractual situation (while behind the veil of ignorance) one does not know how well off one will be outside the contractual situation, given claims of form **ψ**, it is rational to agree on principles of distribution which give some priority to benefiting the worse off. This is roughly how the intuitiveness of claims of form **ψ** could, through a veil of ignorance style justification, support claims of form **Ω**.

As I discuss in Appendix C, claims of form **ψ** are inconsistent with a plausible measure of the size of benefits (the Rational von Neumann-Morgenstern measure), which cleanly avoids what I called the Ambiguity Problem. But never mind this problem for now. A potentially greater one is that, if we find claims of form **ψ** to be

---

[61] For the most famous statement and discussion of what we can call *Veil of Ignorance Contractualism*, see John Rawls's *A Theory of Justice*. For an argument for Average Utilitarianism based on such Contractualism, see Harsanyi 1955.

intuitively plausible, we can re-raise an analog of the Priority Monster problem which was outlined in 3.2. That is, we can construct a series of premises, each of which is no less intuitively plausible than any claim of form ψ, which together imply that it is rational for one to prefer an arbitrarily small benefit in a possible scenario in which one is arbitrarily badly off to an arbitrarily large benefit in a possible scenario in which one is as badly off as the currently worst off person. This sort of *Intrapersonal Priority Monster* is itself implausible. But insofar as claims of form ψ are meant to imply via a veil of ignorance style justification, claims of form Ω, the Intrapersonal Priority Monster will end up implying Priority Monster. (The Intrapersonal Priority Monster is spelled out more fully in 3.6.2).

I believe that the point just made generalizes. While it is true that a defender of the Priority View *could* argue that she accepts claims of form Ω not because of their intuitiveness but because they follow from claims of form *X* (which she regards as intuitively plausible), I suspect I will always be able to re-raise my puzzle for claims form *X*, and argue that insofar as they do imply claims of form Ω (the Priority View), they also threaten to imply Priority Monster in the way outlined in 3.2. At any rate, much of this is moot insofar as most defenders of the Priority View simply accept claims of form Ω on the basis of their intuitive plausibility. For the sake of convenience, this is what I shall assume from here onward.

This concludes my defense of both parts of Tradeoffs.

3.3.2 Parity

Of course, Tradeoffs alone is insufficient to establish that the defender of the Priority View should accept the Premises. To establish that defenders of the Priority View should accept the Premises, we would also have to establish Parity and the Conditional.

Recall that, according to Parity, the size difference between adjacent benefits can be sufficiently small and each person can be sufficiently worse off than her predecessor, such that each of the Premises would seem, considered independently, at least as intuitively plausible as any other claim of form Ω.

I do not think I can prove Parity. To do this might require me to list infinitely many precise formulations of the Premises, and to then go through and test our intuitions about each of them.[62] However, I think there is good presumptive evidence for Parity, based on various *representative samples* of the Premises. And to disprove Parity, we should not require one to point to any particular Premise which is not as intuitively plausible as others, but only to point to a vague region of the Premises where the Premises are less intuitive.

Consider one of the most plausible claims of form Ω: it is better to achieve a large benefit for the currently worst off person than to achieve a slightly larger benefit for the currently best off person. We can set up the Series such that the degree to which each person is worse off than her predecessor dwarfs the difference between the

_____

[62] In Carlson 2000, 251 comes close to demanding this (unreasonably high) level of proof in response to a different Spectrum Argument offered by Norcross 1997.

well-being levels of the currently best off person and the currently worst off person.

For example, P1 would say (something like): rather than achieving a very large

benefit for the currently worst off person, it is better to achieve a slightly smaller

benefit for a person who is *arbitrarily* worse off than the currently worst off person. I

find this no less intuitively plausible than the most intuitively plausible claims of form

Ω, like the one above. (If it helps one get a better intuitive grip, one might imagine

more concrete claims and examples involving pain and pain relief. For example, first

imagine Alfred, who has a life not worth living, in virtue of being in intense pain each

day. Now imagine Abby, who is much worse off than Alfred in virtue of living a

much longer life of pain as intense as his. Suppose we could achieve a large benefit

for either: we could spare Alfred of 365 days of pain, or we could spare Abby of 364

days and 23.9 hours of pain. It seems intuitively plausible that it is better to help

Abby; indeed, no less intuitively plausible than the most intuitively plausible claims of

form Ω).

What about those Premises involving small benefits, e.g., Pn? Pn would say

(something like): rather than achieving a very small benefit for an arbitrarily badly off

person, it is better to achieve a slightly smaller benefit for a person who is arbitrarily

worse off. This too seems to me no less intuitively plausible than the most plausible

claims of form Ω, at least as long as the size of the (n) benefit is only a small fraction

of the size of the (n-1) benefit. (Again, we might get a better intuitive grip by filling

out the details in the way I did above). And indeed, we can arrange the Series such

that the size difference between adjacent benefits is exactly the same as whatever this

small difference between these last two benefits would need to be in order to render P*n*

sufficiently intuitively plausible.[63] One might raise the following objection:

> P*n* would be more intuitively plausible if the (*n*) benefit and the (*n*-1) benefit
>
> were large, rather than small. Therefore, P*n* cannot be made as intuitively
>
> plausible as any claim of form **Ω**.

But this objection potentially confuses the *importance* of bringing about (*n*)

rather than (*n*-1) with the *intuitive plausibility* of the claim that it is better to do so.

Similarly, consider the distinction between importance and intuitiveness in the context

of two scenarios in which the only relevant factor is the size of the benefits that can be

achieved. In the first scenario, I can bring about a tiny benefit, or a somewhat greater

benefit. In the second scenario, I can bring about a tiny benefit, or an enormous

benefit. It is just as intuitively plausible that I bring about the larger benefit in the first

scenario as it is that I do so in the second scenario. But it is more important that I

bring about the larger benefit in the second scenario. This distinction also applies to

claims of form **Ω**; while some such claims are much more important than those

contained in Premises, Parity only claims that those contained in Premises can be

made at least as intuitive as any claim of form **Ω**. It may be useful to keep this

---

[63] Although it is not necessary to assume that each benefit in the sequence is smaller than its predecessor by a constant amount or percentage, making this assumption ensures that the sequence of benefits will inevitably move from a large benefit to a small benefit in a finite number of steps. And it thus allows us to avoid possible worries about Zeno's Paradox. Cf. Carlson 2005. Also see Appendix D.

distinction between importance and intuitiveness in mind in 3.4, where I discuss some further objections to Parity.

So far, I have claimed that various representative samples of the Premises suggest that the *location* of the Premise, i.e., whether it is closer to P1 or Pn, does not by itself have an effect on its intuitive plausibility – at least, when considering the best or most relevant formulations of the Series.

Next, we should observe that the fact that the Premises collectively imply Priority Monster, which is counterintuitive, does not tell in favor of the intuitive plausibility of some Premises over others – since it is equally true of each one of the Premises that if it were not accepted we could thereby avoid Priority Monster. This fact gives us reason to reject at least one of the Premises, but it does not give us any reason to reject some Premises *over* others. Having observed this, we can now see how one might have been tempted to think that Premises near Pn are less plausible than those near P1. If we went through intuitively judging the Premises *starting* and P1 and moving toward Pn, we might get progressively less and less confident about accepting individual Premises given our belief that if we accept P1 through Pn Transitivity will commit us to Priority Monster. This might bias us against Premises located closer to Pn. We should check whether the bias works both ways. If, when we went through intuitively judging the Premises starting at Pn and moving toward P1, we became progressively less and less confident about accepting individual Premises given our belief that if we accept Pn through P1 Transitivity will commit us to Priority Monster, then it might not be the location of the Premises as such that is

driving our intuitions. We should guard against being misled in this way, and thereby

avoid rejecting Parity for erroneous reasons.

In sum, there seems to be good evidence – at least, good presumptive evidence

– in favor of Parity.

3.3.3   The Conditional

Recall that the Conditional states: if we accept claim C1 on the basis of its

intuitive plausibility, and if claim C2 is at least as intuitively plausible as C1, and if

other things are equal, then we should also accept C2.

The Conditional requires us to be consistent about our acceptance standards.

We should treat like cases alike. Considered in the abstract, this seems highly

reasonable. The Conditional can be elucidated further while responding to a possible

objection.

One might object that the Conditional implies much more than that defenders

of the Priority View should accept the Premises. Doesn't the Conditional also imply

that anyone who accepts *any* claim C1 – not just a claim of form Ω – on the basis of

its intuitive plausibility should accept the Premises, as long as the Premises are each at

least as intuitive as C1?

The answer is "not necessarily." This is because the Conditional contains the

clause, "and if other things are equal." If C1 is a relevantly different *kind* of claim

than, say, P*n*-3 of the Premises, then it is not necessarily a failure to treat like cases alike if one accepts C1 for its intuitive plausibility, but denies P*n*-3 (though it is itself no less intuitive than C1). Perhaps, for example, C1 is not the kind of claim that, when taken collectively with other claims of its kind, has implausible implications. On the other hand, P*n*-3 *is* the kind of claim that, when taken collectively with other claims its kind (i.e., of form Ω), has an implausible implication: Priority Monster. But if we do accept claims of form Ω on the basis of their intuitive plausibility, then it does seem a failure to treat like cases alike not to accept other claims of this very same form which are no less intuitively plausible. If the former claims meet our standards of acceptance, then the latter should too. This is just what the Conditional says.

Indeed, since defenders of the Equal Weight View do not accept any claim of form Ω, the Conditional does not imply that they should accept the Premises. It is not a failure to treat like cases alike to, e.g., accept the claim that other things equal bigger benefits are better on the basis of its intuitiveness, but not to accept claims of form Ω on the basis of their intuitiveness.

Again, according to the Conditional, if we accept claim C1 on the basis of its intuitive plausibility, and if claim C2 is at least as intuitively plausible as C1 (and if other things are equal), then we should also accept C2. One might claim that, in cases where we *know* that the conjunction of C1 and C2 is false, the Conditional must also be false. This would be a mistake. *If* we accept C1, then we should accept C2, since the evidence for each is equally good, *but* in cases where we should not accept both C1 and C2, what we should do is accept *neither* C1 nor C2. Indeed, defenders of the

Equal Weight View might claim that this is the very sort of position we are in with respect to the Priority View: we know that Priority Monster is false, and so we know that the conjunction of P1 through Pn is false, and given that these Premises are equally intuitively plausible and as plausible as any claims of form Ω (Parity), we should reject claims of form Ω.

(In Appendix E, I discuss the question of whether we can plausibly accept *some* claims of form Ω without accepting any of the Premises, on the grounds that these claims of form Ω which are distinct from the Premises do not collectively imply Priority Monster).

In sum: Tradeoffs, Parity, and the Conditional together imply that the Priority View should accept the Premises and thereby imply Priority Monster. I have now said all I plan to say in defense of Tradeoffs (the claim that the Priority View is, at a minimum, committed to at least some claims of the form Ω, *it is better to achieve a smaller benefit for a worse off person than to achieve a larger benefit for a better off person*, and that it is committed to such claims on the basis of their intuitive plausibility). There is still further defense of Parity and elaboration of the Conditional to be given.

3.4   Well-Being, Lives, and Critical Thresholds

Claims of form **Ω** say that it is better to achieve a smaller benefit for a worse off person than to achieve a larger benefit for a better off person; I have argued that it is the acceptance of such claims on the basis of their intuitiveness which arguably gets the Priority View into trouble. However, one might worry that focusing on the form of these claims, or focusing on them in this abstract way, potentially glosses over certain important factors to which our intuitions might be sensitive. When we focus on the form of such claims, we might be led to pay attention solely to the size of benefits and how well off people are in a sort of comparative way, i.e., *smaller* benefits for *worse* off people. And we might thereby ignore important facts having to do with the absolute or non-comparative size of benefits or with the absolute levels of well-being at which lives can be lived. A possible worry is that, once we pay heed to at least some of these important facts, we will find it impossible to construct the Priority Monster Spectrum Argument, or we will find Parity to be implausible, or we will believe that the Conditional only has limited force against the Priority View.

3.4.1 Is There a Worst Possible Life?

I have been assuming that we can make P1 through P*n* at least as intuitively plausible as any claim of form **Ω** by arranging the Series such that the size difference between adjacent benefits is sufficiently small and such that the difference in quality of life between adjacent lives is sufficiently large. But we might think that there is a

limit on how great this latter difference can be, and consequently that we cannot arrange the Series such that it will always be sufficiently great so as to render each of the Premises plausible. If there were a worst possible life, we could not make each life arbitrarily worse off than its predecessor.

Some believe that there is a limit on how *well* lives could go. Their claim is not that there is a limit on how much pleasure, satisfaction, or accomplishment a life could actually contain. Rather, their claim is that once a life has achieved a certain level of well-being, adding more pleasure, satisfaction, or accomplishment (or whatever it is that makes lives go well) to it will fail to make the life go any *better*. I am not sure there is such a limit.

However, it seems that there is no limit on how *badly* lives could go. First, it is always possible to add more suffering to a life; one's stay in Hell can always be extended by another day, where the suffering experienced each day is no less intense than the previous. And second, other things equal, adding more suffering to any life will make it worse. It is hard to deny that for any life consisting of n days in Hell, a life consisting of $n+1$ days in Hell would be worse still.[64]

3.4.2 Lives and Times

---

[64] One might argue that there must be symmetry about these two sorts of limits: either there are such limits *both* on how well lives can go *and* on how badly lives can go, or there is *neither* sort of limit. I am not sure whether this is a symmetry that must be maintained, but if it must be, then I believe that the overwhelming plausibility of the claim that there is no limit on how badly lives can go would force us to accept that there is no limit on how well lives can go.

Many defenders of the Priority View accept:

Lives: all that matters for how much priority weight one should receive is how badly off one would be over the course of one's whole life.[65]

Some others accept:

Lives and Times: it matters, for how much priority weight one should receive, *both* how badly off one is now or during some other period of time, *and* how badly off one would be over the course of one's whole life.[66]

Finally, some might accept:

Not Lives but Times: while it does matter, for how much priority weight one should receive, how badly off one is now or during some other period of time, it does not matter how badly off one would be over the course of one's whole life.

Those defenders of the Priority View who accept Not Lives but Times might then object to the Priority Monster Spectrum Argument as follows: even if there were

---

[65] This is the version of the Priority View favored by Arneson 2000 and Crisp 2006b, chapter 6, among others.

[66] See McKerlie 1989 and 1997.

no limit on how badly *lives* could go, there is a limit on how badly *temporal segments* of lives could go; particularly, they cannot be arbitrarily worse than the worst temporal segment of human life. Therefore, the Priority Monster Spectrum Argument fails to target versions of the Priority View which accept Not Lives but Times (just as – as discussed in 3.4.1 – this Argument would fail to target the Priority View if there were a worst possible life).

There are two replies to this objection. First, we could assume, for the sake of argument, that Not Lives but Times is true, but argue that even versions of the Priority View which accept Not Lives but Times still face the Priority Monster Spectrum Argument outlined in 3.2. After all, if Satan could make the priority monster's life arbitrarily bad by making her sentence in Hell arbitrarily *long*, why couldn't he make her life *at a time* arbitrarily bad by making her suffering arbitrarily *intense*? There may be biological limits on how intensely a human being could suffer, but presumably Satan would have the power inflict superhuman levels of suffering; and it seems dubious that there is a metaphysical limit on the intensity of suffering as such.

Second, we could argue that Not Lives but Times is independently implausible. In particular, it seems unacceptably temporally biased. In claiming this, we need not claim that all sorts of temporal biases are unacceptable. For example, Parfit and others have argued that there is a normatively significant asymmetry between past and future suffering.[67] Perhaps some are attracted to what we might call *the Temporally Asymmetric Priority View*: it is only one's present and future suffering, or more

---

[67] See chapter 8 of Parfit 1984 and Hare 2008. For criticisms of past-future temporal bias, see Brink 2011, especially §7.

generally one's present and future well-being, which affects the amount of priority weight that one should receive. Though I am not convinced that the Temporally Asymmetric Priority View is plausible, it would nonetheless fail to avoid the Priority Monster problem as long as we specified that the priority monster's suffering were in the future.

Another temporally biased version of the Priority View accepts Lives and Times. Here too, while I am not convinced that this version of the Priority View is plausible, it would fail to avoid the Priority Monster problem as long as the priority monster could have an arbitrarily bad life (and I claimed in 3.4.1 that it can).

My claim is that, whatever we make of these other temporally biased versions of the Priority View, versions which espouse Not Lives but Times are unacceptably temporally biased. If facts about how well off people are at various times can ground priority for those worse off at various times, then it seems implausible that facts about how well off these same people are at others times would have *no* impact on considerations of priority *merely* because they are facts about how well off they are at *other* times. We can remain officially neutral both about the Temporally Asymmetric Priority View and about Lives and Times, but insist that it seems that if we should give priority weight on the basis of how well off people are at some times, we should, *other things equal*, give priority weight on the basis of how well off they are at all times (at least, all present and future times).

One might argue that there are plenty of cases in which the "other things equal" clause in the previous sentence is not satisfied. For instance, perhaps while we

should *in ordinary cases* give priority to benefiting people according to how well their lives as a whole go, we should not do so when dealing with *very* long lives, which last for millions or billions of years. Perhaps such extraordinarily long lives would not be psychologically integrated or unified in the way that lives ordinarily are, and so perhaps they ought to be treated differently by the Priority View.[68] We might adopt:

> *Not Lives but Psychological Units*: rather than how well off one would be over the course of one's whole life, what matters for how much priority weight one should receive is how well off one is during periods of one's life which are relevantly psychologically unified. (Often, these periods correspond to whole lives, but they need not).

If we adopted Not Lives but Psychological Units, we might treat very long lives as a series of many different lives, or "priority-relevant units," individuated on the basis of their degree of psychological unity. We can take lives which have a negative level of well-being at each time,[69] and make them arbitrarily worse off by making them last an arbitrarily long time. This is one way to make a priority monster. But we cannot make priority-relevant units which have a negative level of well-being at each time arbitrarily worse off by making them last an arbitrarily long time. This is

---

[68] For a seminal discussion of the relevance of psychological continuity and connectedness to rationality and to morality, see chapters 14 and 15 (respectively) of Parfit 1984.

[69] These lives are, because of their low quality, worth not living – or worse than nothing. Such lives might only be marginally worse than nothing at each time.

because priority-relevant units only last as long as they remain relevantly psychologically unified, and after some amount of time this psychological unity will break down.

Though Not Lives but Psychological Units seems to be an interesting variant of the Priority View, it does not provide a promising strategy for avoiding Priority Monster. First, we could repeat the point that Satan could make these priority-relevant units arbitrarily worse off by increasing the intensity of their suffering arbitrarily. Second, there are familiar difficulties with formulating views like Not Lives but Psychological Units. If, for example, the relevant sort of psychological unity breaks down after 137 years, then how many priority-relevant units exist on the 300th birthday of an individual who lives to be 969 years old? Given the fact that the various 137-year-old units *overlap*, the answer is arguably that there are infinitely many.[70] But would this then imply that, since it would benefit infinitely many priority-relevant units, giving this individual a piece of cake at his 300th birthday party is infinitely morally important? This cannot be right. Third, even if we could find a satisfactory way of formulating this view, it would pose no insurmountable difficulty for the Priority Monster Spectrum Argument. We could imagine that, in Priority Monster, Satan has engineered the priority monster to be sufficiently relevantly psychologically integrated and unified over the course of her very long and miserable life. Normally, our beliefs, desires, intentions, and memories (roughly, our "personality") gradually evolve over time. Given enough time, such a gradual process of personality evolution

---

[70] See Lewis 1976, 29-31, and Brink 1997, 113-4.

will result in a very large change in personality. But there seems to be no metaphysically necessary connection between the passage of time and personality evolution. Indeed, it seems that Satan could make it so that the priority monster's personality never changed at all: from the moment she came into existence, Satan gave her the all the relevant beliefs about her life, as well as the relevant "pre-memories" about her entire future. Her beliefs and memories thus never change as she continues on living. Nor do her desires, intentions, etc. The *only* facts about her that change as she ages, are (obviously enough!) her age, and the total amount of suffering that she has in fact endured over the course of her existence.

To repeat, if there were a worst possible life or worst possible "priority-relevant unit," one could offer a good defense of the Priority View against the Priority Monster Spectrum Argument in 3.2. In particular, if there were a worst possible life, then not all the lives in the Series could be made arbitrarily worse off than their predecessors, and so perhaps some of the Premises could not be made sufficiently intuitive (in order to satisfy Parity).

But since it is plausible that there is no worst possible life or "priority-relevant unit," we can arrange the Series such that each of the Premises can be made at least as intuitive as any claim of form Ω. So Parity still seems plausible. Adopting Not Lives but Times or Not Lives but Psychological Units will not help answer the Priority Monster Spectrum Argument, but we might think that there remain some promising ways of modifying Lives or Lives and Times which may enable us to find at least a partial answer to this Argument.

In particular, we can ask what is the relevant sense in which lives are badly off that earns them greater priority weight. There are different answers to this question, which would serve to further specify Lives or Lives and Times. I have been assuming:

Total Lifetime Well-Being: the lower one's total lifetime well-being score, the more priority weight one should receive.

But one could instead adopt:

Average Lifetime Well-Being: the lower one's average lifetime well-being score, the more priority weight one should receive.

A putative advantage of Average Lifetime Well-Being over Total Lifetime Well-Being is that it blocks one way to make a priority monster which the latter readily allows – that is, by taking a life which has a negative level of well-being at each time and making it arbitrarily worse off by making it last an arbitrarily long time. Doing this would not, according to Average Lifetime Well-Being, create a priority monster. That is because this view implies that making a bad life worse by making it longer would fail to make it worse in the way relevant to priority weight. Average Lifetime Well-Being would only offer a partial answer to the Priority Monster

Spectrum Argument, since it would not block the other way to make a priority monster – that is, by making the suffering contained in it arbitrarily intense.

But Average Lifetime Well-Being might have other advantages over Total Lifetime Well-Being. Take an example. For any person, Alpha, who suffers at an arbitrarily great intensity and lives for an arbitrarily long time, there is a person, Beta, who suffers at a very low intensity (e.g., has the sensation of a very mildly annoying headache) for an arbitrarily longer time, who has a lower total lifetime well-being score. But we might think it implausible to give greater priority to benefiting Beta than to benefiting Alpha.

Nonetheless, there are familiar problems which plague various sorts of average views (e.g., Average Utilitarianism), and the Average Lifetime Well-Being view is no exception. Compare a person who has suffered in Hell for 10,000 years with a person who has suffered in a slightly more intense Hell for only 1 year. It is hard to believe, as the Average Lifetime Well-Being view implies, that benefiting the second person matters more than benefiting the first person.

Defenders of Lives or Lives and Times might thus be inclined to reject both Total Lifetime Well-Being and Average Lifetime Well-Being in favor of some third view. Perhaps this third view would be sensitive to total lifetime well-being and average lifetime well-being in a way that would avoid the two problems outlined in the previous two paragraphs, or perhaps it would take into account the overall pattern

or shape of well-being across time in other ways.[71] It seems to me that, whatever view we choose here, as long as we are defenders of Lives or Lives and Times, we should be constrained by:

Duration Matters: if one has a life with a sufficiently low well-being level at each time, or a life with a sufficiently low average well-being level, then one should receive more priority weight the longer one's life is, if other things are equal.

Having more of a significantly bad thing seems to be an important respect in which one can be worse off, and it seems just as plausible or legitimate a basis for greater priority weight as any. If we did not amend Duration Matters with the words "sufficiently low," then we would potentially face the problem – involving Alpha and Beta – which Total Lifetime Well-Being faced. In particular, we could create priority monsters by taking lives which only have a marginally bad average (or are, on average, just slightly worse than nothing at each time) and making them arbitrarily long. Duration Matters does not imply this, since it need not count a marginally low well-being level as a sufficiently low well-being level. But any view should count a life filled with suffering as intense as any ever endured by humans as being at a sufficiently low well-being level at each time. And so Duration Matters would imply

---

[71] For discussions of what we can call *Global Views* of lifetime well-being, see Appendix I Parfit 1984, and see Griffin 1986, and Velleman 1991.

that we could make priority monsters by taking lives with this level of well-being at each time and making them arbitrarily long.

To summarize: we have not found, in this subsection or the previous one, a good reason for thinking that we cannot arrange the Series in 3.2 such that each of the Premises is as intuitively plausible as any claim of form **Ω**. The principal sort of objection I have been considering is that, contrary to what I had initially assumed, there is a worst possible life or a worst possible "priority-relevant unit." I have argued for the conclusion that this is false. However, what I have said here may give rise to a different but related kind of objection to my 3.2 Argument: that it relies on intuitions which are not trustworthy, because they involve unimaginably large numbers, durations, or intensities. This objection will be addressed in chapter 4.

3.4.3   Sufficiency Views

Some people who reject the Equal Weight View adopt neither the Equality View nor the Priority View per se. Instead, they accept a *Sufficiency View*. There are many possible such views,[72] which I will not here exhaustively list; to give a rough sense of the options, I will sketch three:

---

[72] See Frankfurt 1987 and Crisp 2003 for two different Sufficiency Views.

Sufficiency 1: Benefiting people *only* matters when these people would otherwise fall below a particular well-being threshold. Benefits to persons above the threshold do not matter at all.

Sufficiency 2: Benefiting people matters more, the farther below a particular well-being threshold they would otherwise fall, and takes *absolute* priority over benefiting people above the threshold. Benefits to persons above the threshold are treated the way the Equal Weight View treats all benefits.

Sufficiency 3: Benefiting people matters more, the farther below a particular well-being threshold they would otherwise fall. Benefits to persons above the threshold are treated the way the Equal Weight View treats all benefits.

We can call the particular well-being threshold relevant to Sufficiency Views the *sufficiency threshold*. I assume that there can be significant vagueness or indeterminacy about where the sufficiency threshold lies, and do not regard this sort of vagueness as a problem for Sufficiency Views. Indeed, the problems with these views I am about to identify arise independently both of what the sufficiency threshold is as well as whether or not it is only vaguely specifiable.

We do not need a Priority Monster Spectrum Argument to see that Sufficiency 1 and 2 are implausible. Against Sufficiency 1: if we could achieve a colossal benefit for an extremely well off person (far above the sufficiency threshold), or do nothing,

and other things are equal, it seems better to do the former. Against Sufficiency 2: if we could achieve a piddling benefit for someone just non-vaguely below the sufficiency threshold or achieve colossal benefits for one million others just non-vaguely above it, and other things are equal, it seems better to do the latter. Sufficiency 2 thus seems to be implausible in roughly the way that Maximin is (see 3.1.2).

At any rate, Sufficiency 2 and 3 are equivalent to the Priority View, when applied to persons below the sufficiency threshold. And my Priority Monster Spectrum Argument can be constructed so that it deals strictly with lives that fall below any possibly relevant threshold. So, it seems that my Priority Monster Spectrum Argument targets these Sufficiency Views in precisely the way that it targets the Priority View.

3.4.4 Nontrivial Benefits

One might agree with the Priority View that benefiting people matters more, the worse off they are, but add the following proviso: a large benefit for a badly off person always matters more than a piddling benefit for another person, no matter how badly off.[73]

---

[73] One might add: there is *no limit* on how much more it matters to benefit people, the worse off they are, but it is better to achieve a large benefit for a badly off person than to achieve a piddling benefit for another person, no matter how badly off. This would be a kind of *Lexical View* (cf. Parfit 1984, 414).

At a first glance, this proviso might appear to reconcile the Priority View with the denial of Priority Monster. But this will not work. This proviso, by itself, does nothing to rule out P1 through P*n*, since each Premise merely says it is better to achieve a *little* less well-being for a much worse off person. Unless the proviso is combined with the rejection of at least one of these Premises, or the rejection of Transitivity, it implies a contradiction – since Premises P1 through P*n* and Transitivity imply Priority Monster, whereas the proviso denies this.

A distinct but related proposal is that the Priority View should not *apply* to very small benefits.[74] Indeed, a version of the Priority View could stipulate that only benefits that exceed some low size threshold are subject to prioritarian weighting (benefits that fall below the threshold are treated just as the Equal Weight View treats all benefits). Call this version of the Priority View, *the Nontrivial Benefits Priority View*. This View could claim that those premises in P1 through P*n* involving benefits below the relevant size threshold will fail to be true.

I have two replies. First, it is not clear that the way in which the Nontrivial Benefits Priority View denies some of the Premises is plausible. After all, it is implausible to deny the Conditional, and there does not appear to be plausible way to defend a denial of Parity (i.e., a denial that is compatible with or driven by the Nontrivial Benefits Priority View). As I claimed in 3.3.2, I find the claim that, "rather than achieving a very small benefit for an arbitrarily badly off person, it is better to achieve a slightly smaller benefit for a person who is arbitrarily worse off," to be no

---

[74] See Roger Crisp's (2003, 758) suggestion that we might exclude trivial benefits from the "prioritarian half" of the Sufficiency View.

less intuitively plausible than the most plausible claims of form **Ω**. (At least, as long as the smaller benefit is only a small fraction of the size of the larger one). Though, for example, it is less *important* to spare people of mild headaches than it is to spare them from very long and intensely painful ordeals, it seems just as *intuitive that* we should, at least if other things are equal. Similarly, though it is less important that we give priority to the worse off in cases involving small benefits than that we do so in cases involving large benefits, it seems no less intuitive that we should, at least assuming other things are equal. (Again, we could perhaps get a better intuitive grip by imagining a more concrete example involving pain and pain relief; recall the Alfred and Abby example from before. Both have pain-filled lives that are not worth living, but Abby's life is *much* worse. Now suppose that we can relieve one of their many headaches: spare Alfred of a 30 minute headache, or spare Abby of a 29.9 minute headache. Though it does not seem extremely important to spare someone of a 30 minute headache, it does seem plausible that it would be better to spare Abby of the pain, since she's so much worse off).

Some might not be satisfied with this first reply. They might accept my distinction between importance and intuitiveness, but still think, on careful reflection, that the case for giving priority to the worse off evaporates in cases involving very small benefits. Suppose, for the sake of argument, that they are correct. This would mean that Parity is false, and that one could accept the Nontrivial Benefits Priority View without risk of implying Priority Monster. But there are still serious problems.

Relief from one week of intense pain seems to be a fairly big benefit – it is definitely not a very small or trivial benefit. For simplicity, suppose that this is a benefit of size **B**. The Nontrivial Benefits Priority View would apply to benefits which are at least size **B**. But now consider:

Two Priority Monsters with Nontrivial Benefits. We could achieve an arbitrarily large benefit for an arbitrarily badly off person, priority monster #1, or we could achieve a benefit of size **B** for a person, priority monster #2, who is arbitrarily worse off than priority monster #1. It is better to do the latter.

This seems implausible. Even if it were plausible to restrict the Priority View to nontrivial benefits (I am doubtful that it is), this restricted version of the Priority View would face an analog of the problem outlined in 3.2. For consider *Series with Nontrivial Benefits*:

(1): an arbitrarily large benefit for priority monster #1.

(2): a benefit slightly smaller than the one in (1) for a person much worse off than the one in (1).

(3): a benefit slightly smaller than the one in (2) for a person much worse off than the one in (2).

And so on… all the way up to,

(*n*): a size **B** benefit for priority monster #2.

We could also consider *Premises with Nontrivial Benefits*, and *Parity with Nontrivial Benefits* (I will let the reader fill in the details). Instead of facing the challenge of plausibly avoiding Priority Monster, the Nontrivial Benefits Priority View would face the challenge of plausibly avoiding Two Priority Monsters with Nontrivial Benefits. To avoid this conclusion, it would have to deny Tradeoffs, Parity with Nontrivial Benefits, or the Conditional. So our problem remains.

In Appendix D, I discuss some further versions of the Priority View that avoid Priority Monster. I also argue none of them is plausible; each violates Parity. In Appendix E, I explore a crafty attempt to maintain the Priority View, Parity, the Conditional, and the denial of Priority Monster. I show that the only way this feat can be accomplished is by adopting a view which is, in turn, too implausible.

3.5   Taking Stock

Before confronting any more monsters, let us take stock. The foregoing discussion has, I believe, provided considerable evidence for the conclusion that the

Priority View faces a dilemma that the Equal Weight View avoids. The following

diagram illustrates this dilemma:

Priority View   **Equal Weight View**

Accepts both the   Rejects either the Conditional   Accepts both the
Conditional and Parity   or Parity (**Implausible**)   Conditional and Parity

The Premises (given that the Priority   The Equal Weight View **does
View accepts some claims of form   not entail** the Premises, given
Ω)   that it does not accept any
                                                                      claims of form **Ω**)

Priority Monster (**Implausible**)

Figure 3: A Dilemma for the Priority View

The Priority View must either accept or reject the Conditional and Parity. It is

implausible to reject either. If it accepts both, it implies the Premises, which imply

Priority Monster (which is implausible). We can defend the Equal Weight View

without either sort of cost. That is, we can accept the Conditional and Parity without

fear of implying Priority Monster.

What should we conclude from this? I think we should conclude that the

Modest Claim is true. According to the Modest Claim, the Equal Weight View is

more plausible than the Priority View *in one way*. In particular, it is more plausible in

that it can avoid Priority Monster without denying either the Conditional or Parity.

One can accept the Modest Claim without taking a stand on whether the Priority View is more or less plausible, *all things considered*, than the Equal Weight View. After all, though the Equal Weight View does not face many of the difficulties I have presented in this chapter, it genuinely does have other disadvantages that the Priority View does not. (Recall our intuitions about cases like Lollipop, City or Suburbs, Two Children, and Divided World). It may be that no matter what view about the distribution of well-being we adopt, it will have implausible implications. A defender of the Priority View could concede that her view is genuinely embarrassed by the Priority Monster Spectrum Argument, but maintain that she nonetheless espouses the *least implausible* view.

Another possible view to take is that, given the Spectrum Argument here presented, it is very difficult or even impossible to determine which view is all things considered more plausible, or perhaps that these views are roughly equally plausible all things considered.

Yet another possibility is that the *Extreme Claim* is true. According to the Extreme Claim, the Priority Monster type difficulties faced by the Priority View are, on balance, significantly greater than those faced by the Equal Weight View, and so the Equal Weight View is, all things considered, the less implausible view. I have not explicitly argued for the Extreme Claim, though I believe it is defensible. In chapter 4, I will discuss further cases which at least support the Modest Claim.

My conclusion, for now, is that the Priority Monster Spectrum Argument supports the Modest Claim. It reveals the Priority View to be in one way less

plausible than the Equal Weight View, since the former faces a dilemma (see above) that the latter does not.

3.6   Two More Monsters

I will now argue briefly that modified versions of the Priority Monster Spectrum Argument can, respectively, raise an analogous dilemma for the Equality View (3.6.1), and reveal the Priority View to violate a particular constraint in a potentially more worrying way than was already noted in Appendix C (3.6.2).

3.6.1   The Equality Monster Spectrum Argument

Recall, from chapter 2, that according to Parfit's formulation of the Priority View, the moral importance of benefiting people depends only on how well off these people are in *absolute* terms. On this view, the moral importance of benefiting people is unaffected by facts about how well off others besides these very people are. But recall that, according to the Equality View, it does matter how well off people are in *relative* terms; the moral importance of benefiting people *is*, according to this view, affected by facts about how well off people other than those for whom we could achieve benefits are. According to the Equality View, then, whether we should accept

the Premises of the Priority Monster Spectrum Argument might depend on how well off people *other than* the people in the Series are.

The Equality View might seem, at first, to have the resources to avoid Priority Monster. For example, despite the fact that the arbitrarily badly off person is (arbitrarily) worse off than the currently worst off person, it might be that achieving the large benefit for the currently worst off person would do considerably more to reduce overall inequality in world than would achieving a piddling benefit for the arbitrarily badly off person if she existed too (at least, on some measures of "overall inequality"). However, we can retell the case. Consider:

> *Islands*. Maude is the only sentient being alive right now. Moreover, she is the only sentient being ever to have existed. There is only one other sentient individual who will ever exist. It is certain that this other individual will come into existence tomorrow, but his identity and how well off he will be is uncertain.[75] He and Maude will never meet, since Maude is stranded on a large continent and the soon-to-exist person will exist on one of many remote islands. Though it is impossible for them to travel to each other, it is possible for Maude to send him a care package via raft. Maude must choose today

---

[75] It is important to notice that this other person *will* exist in my example, but it is uncertain who it will be (we can suppose there is a fact of the matter who will exist, but it is not known). The defender of the Equality View first compares the state of affairs in which Jeffrey-1 will exist and is benefited with the state of affairs in which Jeffrey-2 will exist and is benefited, and then judges that the latter is better, and so on for each of the below scenarios. Then, she applies expected value theory. She does not apply her view to a person whose existence is contingent on what we do (compare this with how a Person-Affecting Utilitarianism conjoined with expected value theory would handle cases in which it is uncertain who will exist; presumably this is importantly different from how it would handle cases involving persons whose existence is dependent on what we do). There is a similar footnote in Appendix B regarding how the Priority View gets applied to *Only One of Two Will Exist*.

which of the many islands to send the care package, though it will not reach its

destination for 20 years. Consider the following, only one of which will occur:

(1) Jeffery-1 will exist on Island-1 and will be extremely bad off. (Comparable to the

worst off human, in our non-hypothetical world).

(2) Jeffrey-2 will exist on Island-2, and he will be much worse off than Jeffrey-1

would have been if he had existed.

(3) Jeffrey-3 will exist on Island-3, and he will be much worse off than Jeffrey-2

would have been if he had existed.

And so on, all the way up to…

(n) Jeffery-n will exist on Island-n, and he will be arbitrarily badly off.

Next consider the following possible benefits, which would be achieved if a

Jeffrey received the care package from Maude (we can assume that the worse

off a Jeffrey is, the less well-being he is able to extract from the care package):

B1.   Jeffrey-1 receives an arbitrarily large benefit; (1) occurs.

B2.   Jeffrey-2 receives a slightly smaller benefit than the B1 benefit; (2) occurs.

And so on…

Bn.   Jeffrey-n receives an arbitrarily small benefit; (n) occurs.

Just as it is difficult for the Priority View to avoid implying Premises (at least on some formulations of the Series), it is difficult for the Equality View to avoid implying that B2 is better than B1, that B3 is better than B2, and so on…, and that Bn is better than Bn-1. If the Equality View had these implications, then Transitivity would commit it to the conclusion that Bn is better than B1. If, for instance, there were a 50 percent chance that (1) will occur, and a 50 percent chance that (n) will occur, the Equality View would imply that it is expectably better for Maude to send the care package to Island-n than to send it to Island-1. This seems just as implausible as Priority Monster.

The foregoing discussion about Tradeoffs, Parity, and the Conditional applies, *mutatis mutandis*, to the Equality View. If the Equal Weight View is in one way more plausible than the Priority View in that it can avoid Priority Monster without abandoning Tradeoffs, Parity, or the Conditional, then it is similarly in one way more plausible than the Equality View in that it can, without abandoning the Conditional or the relevant analogs to Tradeoffs and Parity, avoid the claim that in Islands it is expectably better for Maude to send the care package to Island-n than to send it to Island-1.

3.6.2   Intrapersonal Priority Monsters

In Appendix C, I discuss the:

Rational Prudence Constraint: we ought not to treat a person in ways in which she could not prudentially rationally choose to be treated, if it were only the well-being or interests of this person our act could affect, and other things were equal.

If it is prudentially rational for one to choose, under risk, to maximize one's expected benefits, then the Rational Prudence Constraint will conflict with the Priority View (more precisely, it will conflict with views which accept Absolute Matters, i.e., views which attach priority weight on the basis of how well off people are in absolute terms). This is because, as is shown in particular cases (see One Child, in Appendix A), the Priority View implies that we should sometimes provide a person with slightly smaller expected benefits rather than larger expected benefits, even when it is only her interests that are at stake. Some people might argue that, insofar as the Priority View is inconsistent with the Rational Prudence Constraint, it is implausible.

I believe that the Priority Monster Spectrum Argument is a significantly more powerful argument against the Priority View than the argument based on the Rational Prudence Constraint. However, it may be worth observing that these two arguments

can be combined, and that some might take the combination of these two arguments to constitute a third and powerful reason for rejecting the Priority View.

In particular, it seems that the Priority View has difficulties avoiding the conclusion that we should sometimes provide a person with arbitrarily small expected benefits rather than arbitrarily large expected benefits, even when it is only her interests that are at stake.

Consider *n* different possible lives that Anna might live:

  (1) A life just as bad as that of the currently worst off person, or

  (2) A much worse life, or

  (3) A much *much* worse life, or

  (4) A much much *much* worse life, or

  …and so on…

  (n) An arbitrarily bad life.

Now consider a spectrum of different possible outcomes:

  A: an arbitrarily large benefit for Anna, who lives the (1) life.

  B: a slightly smaller benefit for Anna, who lives the (2) life.

  C: a slightly smaller benefit for Anna, who lives the (3) life.

  …and so on…

  Z: an arbitrarily small benefit for Anna, who lives the (n) life.

Suppose the Priority View implies that A is worse than B, which is worse than C, …and so on… which is worse than Z. By Transitivity, then, the Priority View implies that Z is better than A. Thus, it would imply:

Intrapersonal Priority Monster. If Anna had a 50 percent chance of living the (1) life and a 50 percent chance of living the (n) life, and we had to decide now whether to do what would give Anna an arbitrarily large benefit if she lives the (1) life or to instead do what would give Anna an arbitrarily small benefit if she lives the (n) life, and other things were equal, we ought to do the latter.

It is implausible that it would be rationally prudent for Anna to prefer this course of action, and it seems hard to believe that we should go against what it would be rationally prudent for Anna to prefer in this case. If it implies Intrapersonal Priority Monster, then the Priority View conflicts with Rational Prudence Constraint in a way which is arguably more worrying and problematic than noted earlier, in chapter 2.

Moreover, it should be noted that if the Priority View implies Intrapersonal Priority Monster, it is hard to see how it would avoid:

Scarier Intrapersonal Priority Monster. If Anna had a 99.999 percent chance of living the (1) life and a 0.001 percent chance of living the (n) life, and we had to decide now whether to do what would give Anna an arbitrarily large

benefit if she lives the (1) life or to instead do what would give Anna an arbitrarily small benefit if she lives the (n) life, and other things were equal, we ought to do the latter.

Intrapersonal Priority Monster is very implausible, and Scarier Intrapersonal Priority Monster is, I believe, absurd. But again, it seems that the Priority View can only avoid such implications if it rejects the Conditional or the relevant analogs to Tradeoffs and Parity.

The main conclusions of this chapter have already been summarized in 3.5. I have here focused on difficulties faced by the Priority View (along with the Equality View and Sufficiency Views) while abstracting from complications about the aggregation of size and number of benefits. In the next chapter, these complications will be brought into the picture.

Chapter 4

Numbers

CHAPTER OUTLINE

4.1   Aggregating Spectrum Arguments

So far I have abstracted from questions about how to *aggregate* the size and number of benefits. Aggregative principles answer questions like: How many benefits of some particular smaller size would be as good as some number of benefits of some particular larger size? (Similar questions arise for bads, or harms; for instance: how many traffic-related inconveniences are as bad as one premature death?) These questions can arise intrapersonally as well as interpersonally, and raise many of the same puzzles in either context. In any case, questions of *weighting* benefits are different from questions of aggregating benefits. In particular, unless we adopt the Equal Weight View, to answer questions about how to weight benefits we

will always have to look at the characteristics or identities of potential beneficiaries. Questions of aggregation are concerned with how to evaluatively combine the size and number dimensions of (weighted or unweighted) benefits.

Because questions of weighting and aggregating are distinct sorts of questions, and because the Equal Weight View, the Equality View, and the Priority View are all views about the former, they do not by themselves commit to further views about aggregation. The Equal Weight View, for example, says that *however* size and number of benefits should be aggregated, *only* size and number of benefits matters. And the Priority View takes the same neutral stance on aggregation, but says that in addition to size and number of benefits, it also matters how well off their recipients are or would be. In comparing the relative merits of these views in previous chapters, I have held constant the number of benefits, and only varied their size (most examples involved comparing one benefit for one person with another benefit of a different size for someone who is worse off).

In this chapter, I will continue to compare the relative merits of these views about weighting (particularly the Equal Weight View and the Priority View), but here I will vary the number of benefits. In some cases, I will *only* vary the number (while holding size constant), and in other cases, I will vary both the size and number of benefits. At the same time, I will compare my Spectrum Arguments, which have to do with the weighting of benefits, with some more familiar Spectrum Arguments involving aggregation.

4.1.1   Three Older Arguments

Recall the general structure of a Spectrum Argument:

P1.	B is better than A
P2.	C is better than B
P3.	D is better than C
…	…
P(n-1).	Y is better than X
Pn.	Z is better than Y
PT.	Transitivity
∴	Z is better than A

This overlapping chain of premises (P1 through Pn) and Transitivity (PT) together imply that Z is better than A. As I mentioned earlier, some philosophers regard certain Spectrum Arguments as arguments against Transitivity. I will assume Transitivity for the time being, and not defend it against these philosophers until chapter 6.

I will now quickly present three Spectrum Arguments. While these arguments raise similar puzzles (at least in part about aggregation), they are importantly distinct

in various ways which I will not fully explore here, and which will anyway already be familiar to some. First, consider the:

> *Utility Monster*. There could be a person who would gain *so* much well-being from being fed people that, if he existed, it would be better (other things equal) if everyone were fed to this person than it would be if everyone, including him, received very large benefits. (We can add that this utility monster is very well off, and that everyone else is very badly off, prior to receiving any further benefit).[76]

This claim seems implausible. We could argue for it using the following *Utility Monster Spectrum Argument*.

Consider the following finite Series of possible outcomes:

(1): very large benefits for everyone, including the utility monster.

(2): slightly smaller benefits than those in (1) for everyone[77] other than the utility monster, and an *arbitrarily larger*[78] benefit than that in (1) for the utility monster.

---

[76] This is an adaptation of Nozick's case (1974, 41).

[77] We could make the Premises (below) even harder to deny if we arranged the Series such that the size of the benefits is gradually decreased for one person at a time (rather than for everyone at once).

[78] In saying that there is an *arbitrarily larger* benefit for the utility monster, what I mean is that, in order to deny the claim that outcome (2) is better than outcome (1) (i.e., in order to deny P1, below), we would have to say that it is false that (2) is better than (1), *no matter how large* the (finitely-sized) benefit for the utility monster in (2) is. Similar remarks apply to the phrase "arbitrarily larger" in other Spectrum Arguments.

(3): slightly smaller benefits than those in (2) for everyone other than the utility monster, and an arbitrarily larger benefit than that in (2) for the utility monster.

And so on… all the way up to:

(*n*-1): piddling benefits for everyone other than the utility monster, and an arbitrarily large benefit for the utility monster.

(*n*): no benefits for anyone other than the utility monster, and an arbitrarily larger benefit than that in (*n*-1) for the utility monster.

Next consider the following Premises:

P1. Outcome (2) is better than outcome (1).

P2. Outcome (3) is better than outcome (2).

And so on… all the way up to:

P*n*. Outcome (*n*) is better than outcome (*n*-1).

The Premises, together with Transitivity, imply that (*n*) is better than (1). That is, they imply Utility Monster. Lastly, note that there are different versions of Utility Monster. For example, in one version, the utility monster is made very well off *at a*

time (e.g., he is given arbitrarily intense pleasure). In another version, the utility monster is made very well off *over a long period of time* (e.g., he is given moderately intense pleasures for billions of years). We can call these two versions, respectively, the *Intensity Utility Monster* and the *Duration Utility Monster*.

Second, consider the:

Repugnant Conclusion. For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.[79]

This claim seems implausible. We could argue for it using the following *Repugnant Conclusion Spectrum Argument*.

Consider the following finite Series of possible populations:

A: ten billion people, all with a very high quality of life.

---

[79] This is quoted from Parfit 1984, 388. There are different ways in which these latter lives, call them the Z-lives, might be just barely worth living. In an unpublished manuscript called "Towards Theory X," Parfit distinguishes between (1) *Short-Lived Z*, where the Z-lives are high quality at each time, but last for a very short time, e.g., for an hour (Parfit is dubious that such lives could plausibly be described as barely worth living), (2) *Roller-Coaster Z*, where the Z-lives are very high quality at some times and very miserable at other times, such that on balance the quality just barely outweighs the misery, and (3) *Drab Z*, where the Z-lives contain no misery, and only very low quality at each time, e.g., they would contain nothing but "muzak and potatoes" or nothing but the experience of "floating very drunk in a warm bath" (Crisp 2006b, 112). I will assume, with Parfit, that the Z-lives are specified by Drab Z.

B: an arbitrarily larger number of people than in A, all with a slightly lower quality of life than in A.

C: an arbitrarily larger number of people than in B, all with a slightly lower quality of life than in B.

And so on… all the way up to:

Z: an arbitrarily large number of people, all with lives that are barely worth living.

Next consider the following Premises:

P1. Population B is better than population A.

P2. Population C is better than population B.

And so on… all the way up to:

Pn. Population Z is better than population Y.

The Premises, together with Transitivity, imply that Z is better than A. That is, they imply Repugnant Conclusion. Lastly, to help appreciate the scope of these puzzles, note that there are different versions of Repugnant Conclusion. Consider, for example, the:

Intrapersonal Repugnant Conclusion. For any possible life of at least one thousand years, with a very high quality of life at each time, there must be some much longer imaginable life whose existence, if other things are equal, would be better, even though it would be just barely worth living at each time.[80]

This claim, which is structurally similar to Repugnant Conclusion, is implausible. And it has nothing to do with population ethics, e.g., about whether it is better if there are more happy people. This suggests that the Repugnant Conclusion Spectrum Argument does not merely raise a puzzle about population ethics, but about aggregation Spectrum Arguments in general. Let us consider just one more similar Spectrum Argument. Consider:

Hangnails for Torture. For any excruciatingly painful torture session lasting for two years, there is a longer session consisting of very mildly annoying hangnail pain which is, other things equal, worse.[81]

This claim seems implausible. We could argue for it using the following *Hangnails Spectrum Argument.*

---

[80] This Intrapersonal Repugnant Conclusion was discussed (and accepted) by J.M.E. McTaggart prior to Parfit's work on the Repugnant Conclusion. See McTaggart 1921, 452-3.

[81] Rachels 1998 and Temkin 1996. My formulation of this argument is closer to Temkin's.

Consider the following finite Series of possible outcomes:

(1): an excruciatingly painful torture session lasting for two years.

(2): an arbitrarily longer session than in (1), containing slightly less intense pain than in (1).

(3): an arbitrarily longer session than in (2), containing slightly less intense pain than in (2).

And so on… all the way up to:

(n): an arbitrarily long session consisting of very mildly annoying hangnail pain.

Next consider the following Premises:

P1. Outcome (2) is worse than outcome (1).

P2. Outcome (3) is worse than outcome (2).

And so on… all the way up to:

Pn. Outcome (n) is worse than outcome (n-1).

The Premises, together with Transitivity (of "is worse than"), imply that (n) is worse than (1). That is, they imply Hangnails for Torture. Lastly, note that there are different versions of Hangnails for Torture. We can distinguish between an intrapersonal version, where the intense torture and the hangnail pain could occur within a single life, and an interpersonal version, where the intense torture and the hangnail pain could occur within separate lives. While there are important differences between these intrapersonal and interpersonal versions of Hangnails for Torture, both seem implausible. This suggests there is a deeper structural issue about aggregation which this sort of case raises.

4.1.2   What these Arguments Have in Common

The above arguments are, obviously enough, similar in that they are all Spectrum Arguments. But they have more in common than this. There is a structural issue about aggregation raised by (each of the above versions of) the Utility Monster Spectrum Argument, the Repugnant Conclusion Spectrum Argument, and the Hangnails Spectrum Argument. Roughly, the issue is whether there are limits on how size and number of well-being increases can be aggregated. Is it true that enough small increases in well-being can outweigh arbitrarily many arbitrarily large increases in well-being? Is it true that a few large enough increases in well-being can outweigh arbitrarily many arbitrarily large increases in well-being? Similar aggregation

questions arise for decreases in well-being, or increases in suffering. Some people

believe that there are certain limits on how size and number of well-being increases

can be aggregated. (I will return to this debate in chapter 7).

For now, I just want the reader to be aware of these *Aggregating Spectrum*

Arguments, for the purposes of comparing and contrasting them with my *Weighting*

Spectrum Arguments. The three arguments in 4.1.1 *scale up* a particular moral factor,

total well-being (or *total ill-being*), in ways which threaten to embarrass certain views

which recognize this factor as relevant and significant.[82] My Spectrum Arguments are

importantly different, and largely novel, in that they do not scale up this factor.

Instead, my arguments scale up *priority weight* (or *equality weight*) in ways which

threaten to embarrass views which recognize this factor as relevant and significant.

Aggregating Spectrum Arguments have received a lot attention since they were first

introduced by Parfit, Rachels, and Temkin. But, as far as I know, this dissertation is

the first attempt to present and explore Weighting Spectrum Arguments.

4.2   Pooled Priority

---

[82] Indeed, Parfit thinks of the Repugnant Conclusion as a kind of Utility Monster where "the greater sum of happiness comes from a vast increase, not in the quality of one person's life, but in the number of lives lived," (Parfit 1984, 389). Also note that Total Utilitarianism (the right act is the one which maximizes total well-being) is not the only view which recognizes *total well-being* as a morally relevant and significant factor.

In chapter 3, I discussed one type of Weighting Spectrum Argument. That type of argument held constant the number of benefits (and number of recipients of benefits), but it varied the size of benefits and how badly off the recipients of those benefits would be; it scaled up priority weight by supposing that there could be priority monsters, or people who are arbitrarily badly off. I will now discuss another type of Weighting Spectrum Argument, which varies the number of benefits (and number of recipients of benefits), and which scales up priority weight without appealing to priority monsters. These arguments raise an importantly distinct sort of challenge to the Priority View, for at least two reasons. First, as I explain below, it is harder to deny (the analog of) Parity for these arguments (it is not that I think that Parity is a vulnerable premise in earlier arguments, but that there is even less room to worry about it here). Second, these arguments cannot be dismissed on the grounds that priority monsters are impossible (see 3.4.1), for they do not appeal to the possibility of priority monsters.

4.2.1   Pooled Priority: Number Drain

Consider the following:

Pooled Priority. There could be a large number *k* of people at some low well-being level *w* such that it is better to achieve *k* benefits of size *b*, one each, for

these people than to achieve $k + m$ benefits of size b, one each, for a large

number $k + m$ of different people each at the slightly greater well-being level

of $w + \varepsilon$ (where ε can be arbitrarily small).

This conclusion, its different specifications, and how the Priority View might

imply them, can be sufficiently explained without a Spectrum Argument. Call a

person with well-being level w a Valjean, and call a person with the slightly greater

well-being level of $w + \varepsilon$ an Improved-Valjean. Perhaps, over the course of their

lives, Valjeans each suffer one more mild headache than Improved-Valjeans do.

Suppose, for example, that there is a size of b such that the Priority View would imply

that it is better to achieve 1,000 benefits of size b for 1,000 Valjeans (one such benefit

per Valjean) than it is to achieve 1,001 benefits of size b for 1,001 Improved-Valjeans

(one such benefit per Improved-Valjean).

(Defenders of the Priority View could reject this particular tradeoff favoring

1,000 benefits for Valjeans over 1,001 benefits for Improved-Valjeans, but accept

other tradeoffs which are similar, e.g., those favoring 1,000,000 benefits for Valjeans

over 1,000,001 benefits for Improved-Valjeans, again assuming that the benefits are

the same size. This would not require any substantial revision in what I say below.

Alternatively, defenders of the Priority View could claim that it is never better to

provide fewer benefits for Improved-Valjeans than to provide more benefits for

Valjeans. But again, this alone might not require any substantial revision in what I say

below. More extremely, defenders of the Priority View could maintain that it is never

better to forgo more benefits for the sake of providing fewer benefits for the worse off. If they did, their view would not imply Pooled Priority).

To keep things relatively simple, let us assume the Priority View accepts the following tradeoff: it is better to achieve 1,000 benefits of size b for 1,000 Valjeans than it is to achieve 1,001 benefits of size b for 1,001 Improved-Valjeans.

If there were one million distinct occasions on which we could either benefit 1,000 Improved-Valjeans or 1,001 Valjeans, and we (following the above tradeoff ratio) favored the Improved-Valjeans each time, then we will have benefited a total of 1,000,000,000 Valjeans when we could have instead benefited a total of 1,001,000,000 Improved-Valjeans. As the reader may have guessed, a series of tradeoffs of the above form, together with the Transitivity of "better than," imply the conclusion that it is better to benefit 1,000,000,000 Improved-Valjeans than it is to benefit 1,001,000,000 Valjeans. But I will omit this Spectrum Argument. But is its conclusion implausible?

Though this is a question we might ask, it is not the best question to ask. Recall the variables in Pooled Priority. We can generate different specifications of this case by plugging in different numbers for these variables. Note that k is the number of Improved-Valjeans that could benefit and that m is how many *more* Valjeans could benefit. When $k = 1,000$, $m = 1$. Obviously, defenders of the Priority View who think it is better to achieve 1,000 benefits of size b for 1,000 Valjeans than it is to achieve 1,001 benefits of size b for 1,001 Improved-Valjeans will not find *this* specification of Pooled Priority implausible at all, for it merely restates their view. On

the other hand, when $k = 1,000,000,000$, $m = 1,000,000$. And when $k = 10^{80}$, $m = 10^{77}$.

As k becomes arbitrarily large, so does m. These scaled-up specifications are *not* mere

restatements of the view which commits to Pooled Priority in the case where $k = 1,000$

and $m = 1$. It is true that, in the versions of Pooled Priority in which, e.g., $k =$

$1,000,000,000$ and $m = 1,000,000$, the ratio of the benefits that could be achieved for

Valjeans to the benefits that could be achieved for Improved-Valjeans is 1: 1,000.

But, while the *ratio* of benefits stays the same as k becomes arbitrarily large, the

absolute difference in the benefits brought about (m) becomes arbitrarily large. We

might, on reflection, believe that when m gets larger and larger, Pooled Priority

becomes less and less plausible.

A better question to ask, vis-à-vis Pooled Priority, is: is there any size of k

such that it would be better to forgo *arbitrarily more* benefits for Improved-Valjeans

for the sake of giving k benefits to Valjeans rather than Improved-Valjeans (keeping in

mind that the difference in well-being level between Valjeans and Improved-Valjeans

is arbitrarily small)?

Reasonable people might have different intuitions about this case. Some might

find Pooled Priority implausible, when m is very (arbitrarily) large. Others might

defensibly claim that Pooled Priority is plausible, however large m is.

Lastly, some might claim that how plausible Pooled Priority is depends, in

part, on how large b is, and what w is. Though they may be right, my interest here is

whether there are *any* versions of Pooled Priority which are not implausible. If there

are *some* sizes of *b* and levels *w* where Pooled Priority is implausible, then we will

have found a potential problem for the Priority View.

Having raised Pooled Priority, I will now set it aside and look at a potentially

more challenging and powerful case.

4.2.2   Scary Pooled Priority: Size and Number Drain

Consider:

Scary Pooled Priority. There could be a large number *k* of people at some low

well-being level *w* such that it is better to achieve *k* benefits of size *B*, one

each, for these people than to achieve *k* benefits of size *B*, one each, for a large

number *k* of different people each at the slightly greater well-being level of *w* +

ε (where ε can be arbitrarily small) and a benefit arbitrarily larger than size *B*

for another person at well-being level *w* + ε.

If the Priority View implied Scary Pooled Priority, it might seem implausible.

Must it have this implication? No, but there is worry for the Priority View here which

is analogous to the one presented in chapter 3, particularly in 3.2. Namely, either the

Priority View has this implication, or else it must deny the Conditional, an analog of

Tradeoffs, or an analog of Parity. (Analogs to claims made in chapter 3 will be denoted with an asterisk, e.g., Parity* is the analog of Parity).

As before, call a person with well-being level w a Valjean, and call a person with the slightly greater well-being level of $w + \varepsilon$ an Improved-Valjean.

The Priority View seems to imply:

Tie-Breaking. There is some size of benefit B such that it is better to achieve a benefit of size B for a Valjean than it is to achieve a benefit of size B for an Improved-Valjean.[83]

If so, then I assume that it also entails that:

Ω*   There is some size of benefit B such that it is better to achieve a benefit of size B for a Valjean than it is to achieve a benefit of size B for an

---

[83] However, it is possible to defend a version of the Priority View which would only imply Tie-Breaking if ε, or the difference between the well-being level of a Valjean and an Improved-Valjean, were sufficiently large. Thus, ε cannot be arbitrarily small if the Priority View clearly implies Tie-Breaking. Nonetheless, versions of the Priority View which require ε to be very large are implausible. For example, it would be implausible for a defender of the Priority View to defend some claims of form **Ω**, but to deny Tie-Breaking if Improved-Valjeans are *non-trivially* better off than Valjeans. Maybe the fact that over the course of their lives Valjeans each suffer one more mild headache than Improved-Valjeans do is a trivial difference. And maybe this trivial difference is not worthy of the attention of the Priority View. But it seems to be a non-trivial difference if, over the course of their lives, Valjeans each suffer one more *day* of *intense* pain than Improved-Valjeans do. Tie-Breaking, when *this* is the difference between Improved-Valjeans and Valjeans, seems no less intuitive than any claim of form **Ω**. But I do not believe Scary Pooled Priority would be any less implausible if we let ε correspond to this extra day of pain, and let w correspond to the well-being level of the currently worst off person.

Improved-Valjean and a benefit of arbitrarily small size *b* for an Improved-Valjean. (Note: *b* could be such an extremely small benefit that it is a smaller difference in well-being than ε, e.g., perhaps *b* is merely the partial relief of one very mild headache).

Again, this is assuming the Priority View goes beyond the trivially modest view that considerations of priority are to function merely as "tie breakers." (Note that **Ω*** is not *exactly* an analog of Tradeoffs, or of any claim of form **Ω**, since the former makes reference to individuals at specific well-being levels – the Valjeans and Improved-Valjeans).

Now, holding *B* and *b* fixed, consider a finite *Series** of possible outcomes:

(1*):   *k* benefits of size *B*, one each, for *k* Valjeans.

(2*):   *k*-1 benefits of size *B*, one each, for *k*-1 Valjeans, 1 benefit of size *B* for 1 Improved-Valjean, and 1 benefit of size *b* for 1 Improved-Valjean.

(3*):   *k*-2 benefits of size *B*, one each, for *k*-2 Valjeans, 2 benefits of size *B*, one each, for 2 Improved-Valjeans, and 1 benefit of size 2*b* for 1 Improved-Valjean.

And so on… all the way up to:

(*n*-1*): 1 benefit of size *B* for 1 Valjean, *k*-1 benefits of size *B*, one each, for *k*-1

Improved-Valjeans, and 1 benefit of size (n-1)*b* for 1 Improved-Valjean.

(*n**):   *k* benefits of size *B*, one each, for *k* Improved-Valjeans, and 1 benefit of size

(n)*b* for 1 Improved-Valjean.

Next consider the following *Premises**:

P1*. Outcome (1*) is better than outcome (2*).

P2*. Outcome (2*) is better than outcome (3*).

And so on… all the way up to:

P*n**. Outcome (*n*-1*) is better than outcome (*n**).

If the Premises* and Transitivity are true, then (1*) is better than (*n**). That is, the

Premises* and Transitivity entail Scary Pooled Priority.

To complete the analogy with the argument in chapter 3, consider *Parity**:

Each of the Premises* is at least as intuitively plausible as **Ω***.[84]

---

[84] Thinking through a pair of somewhat more concrete examples from the *Premises** might help.

Parity* seems very uncontroversial, and even less controversial than Parity

(though again I believe Parity is already sufficiently plausible). The reason is that it

seems that P1*, P2*,…, Pn*, are virtually just further *instances* of **Ω*** – or at any rate

they are further instances of claims of the very same form as **Ω***. Each of the

Premises* is virtually just a reapplication of **Ω*** to a different context. But it does not

seem to matter whether, in trading off a *B* benefit for an Improved-Valjean and *b*

- First, consider Valjean. He is very badly off through no fault of his own. Next, consider Improved-Valjean. He is just like Valjean, only he's slightly less badly off, in virtue of having one (or a few) fewer stubbed toe(s).

- Now consider some concrete benefits we could bring about:
    B1 = relief from Hellish pain for a month.
    B2 = relief from a few seconds of very minor pain. (Or, the "tiniest possible benefit").
    B3 = relief from Hellish pain for *many many years*.

- If we have any prioritarian intuitions at all, the following will most likely seem plausible to us:

    Tradeoff. It's better to bring about,

    B1 for a Valjean than it is to bring about,
    B1 for an Improved-Valjean *and* **B2 for an Improved-Valjean.**

- Now consider a pair of claims. The first claim is an *early* one of the *Premises**, and the second claim a *late* one of the *Premises**.

Claim 1. It's better to bring about,
    1,000,000 B1s for Valjeans than it is to bring about,
    999,999 B1s for Valjeans *and* **1 B1 for an Improved-Valjean** *and* **1 B2 for an Improved-Valjean.**

Now consider,

Claim 2. It's better to bring about,
    2 B1s for Valjeans *and* **999,998 B1s for Improved-Valjeans** *and* **B3 for an Improved-Valjean** than it is to bring about,
    1 B1s for Valjeans *and* **999,999 B1s for Improved-Valjeans** *and* **B3+B2 for an Improved-Valjean.**

Claim 1 and Claim 2 seem to me roughly equally intuitively plausible, when considered independently. They also just seem to be applications of Tradeoff. They are essentially just Tradeoff applied to cases with different "backgrounds."

benefit for an Improved-Valjean in exchange for a *B* benefit for a Valjean, we do so in isolation of all other persons, or in the context of any other pattern of Valjeans and Improved-Valjeans.[85] It seems implausible that the value of engaging in this tradeoff could *depend* on which, or how many, Valjeans or Improved-Valjeans are already around.

Similarly, one might attempt to avoid Repugnant Conclusion by adopting a view which says that, while there is always some value in bringing people with worth living lives into existence, this value depends on how many other people at this same well-being level already exist. The more people at this well-being level exist, the less extra value is added when more such people are brought into existence. This *Diminishing Value View*, though it enables us to avoid Repugnant Conclusion, is implausible. It seems implausible that the value of bringing people into existence could depend on how many other people are already around, or have already existed.

The denial of Parity* seems implausible, and implausible in the same way that the Diminishing Value View is. Both are implausibly sensitive to irrelevant extrinsic factors. It does not seem to matter whether, in trading off a *B* benefit for an Improved-Valjean and *b* benefit for an Improved-Valjean in exchange for a *B* benefit for a Valjean, we do so in isolation of all other persons, or in the context of any other pattern of Valjeans and Improved-Valjeans.   And it does not seem to matter how

---

[85] Defenders of the Equality View would disagree. Nonetheless, I believe I could revise Scary Pooled Priority so that it targets the Equality View. I would have to revise it in something like the way that Priority Monster had to be revised (see Islands in 3.6.1). I will omit the details.

many people are already around, or have already existed, in order to determine how good it is to bring more people into existence.

And again, consider the Conditional:

If we accept claim C1 on the basis of its intuitive plausibility, and if C2 is at least as intuitively plausible as C1, and if other things are equal, then we should also accept claim C2.

If Parity* and the Conditional are true, and if the Priority View accepts Ω^*, then the Priority View implies Premises*, which in turn imply Scary Pooled Priority, i.e., that outcome (1) is better than outcome (n^*). As with Priority Monster, I suspect that many defenders of the Priority View will find Scary Pooled Priority counterintuitive.

4.2.3   A Partner in Crime?

A defender of the Priority View could, in response, point out that the Equal Weight View has counterintuitive implications which are similar to Scary Pooled Priority. For instance, the Equal Weight View might be conjoined with an aggregative principle such as the:

Total View: other things equal, the outcome with more total (weighted) well-being is better.

The word "weighted" appears in parentheses before "well-being" to indicate that the Total View is a principle about aggregating size and number, and that it does not by itself take a stand on how to *weight* benefits (just as the Equal Weight View, the Priority View, and the Equality View do not say how to *aggregate* benefits). One could thus accept both the Total View and the Priority View. The Equal Weight View and the Total View will together imply that there could be a large number of mild headaches *n* (each for a separate person) such that it would be better to prevent them from occurring than to achieve one large benefit for a Valjean (e.g., sparing him from a year of intense torture).[86]

But, even apart from the fact that it seems possible to defend the Equal Weight View without committing to the Total View, this "partners in crime" response outlined in the last paragraph might not be completely satisfactory. Scary Pooled Priority does reveal an implausible implication that the Priority View must face – unless it denies Parity* or the Conditional – that the Equal Weight View does not face. In the below example, assume that *k*, *B*, and *n* can be arbitrarily large. While a defender of the Equal Weight View might be embarrassed by:

---

[86] In addition to the papers by Rachels and Temkin cited above, see Norcross 1997.

Headaches.   *k* benefits of size *B*, one each, for *k* Valjeans **and *n* mild**

headaches prevented (each would befall a separate

person) is better than *k* benefits of size *B*, one each, for

k Improved-Valjeans and 1 benefit arbitrarily larger than

size *B* for 1 Improved-Valjean.

She is not embarrassed by:

Scary Pooled Priority.   *k* benefits of size *B*, one each, for *k* Valjeans **and**

NOTHING ELSE is better than *k* benefits of size *B*,

one each, for *k* Improved-Valjeans and 1 benefit

arbitrarily larger than size *B* for 1 Improved-Valjean.

While both Headaches and Scary Pooled Priority are counterintuitive, Scary

Pooled Priority is *arguably* more counterintuitive (after all, these two claims make the

same "better than" claim, except that in Headaches *n* mild headaches appear on the left

side of the "better than" – that is enough to make the "better than" claim in Headaches

significantly less implausible). To avoid this implication, the defender of the Priority

View will have to deny either Parity* or the Conditional.

Lastly, it is not difficult to show that the Priority View might have more

extreme implications than Scary Pooled Priority. As long as we start out with a big

enough value for *k*, we can "build up" the *number* of benefits, as well as their size, in

($n*$). Accordingly, ($n*$) might as well be: k benefits of size B, one each, for k Improved-Valjeans, and **arbitrarily many arbitrarily large** benefits for **arbitrarily many** Improved-Valjeans. For at least some specifications of ($n*$), it seems implausible that (1*) is better than ($n*$). While the modest $\boldsymbol{\Omega^*}$ seems reasonable, it seems hard to believe that there is any number of miniscule prioritarian gains that could together outweigh arbitrarily many arbitrarily large gains in well-being for arbitrarily many badly off people.

One further point. Pooled Priority is a "number drain" in that arbitrarily *many* benefits are forgone for the sake of many slight prioritarian gains. Scary Pooled Priority is a "size and number drain" in that arbitrarily many *arbitrarily large* benefits are forgone for the sake of many slight prioritarian gains. Priority Monster is only a "size drain" in that one arbitrarily large benefit is forgone for the sake of providing a piddling benefit for a priority monster. We can also imagine a version of Priority Monster which functions as both a size and number drain. We can call this one:

> *Scarier Priority Monster*. There could be a person so badly off that, if she existed, it would be better to achieve an arbitrarily small benefit for her than to achieve arbitrarily many, arbitrarily large benefits for arbitrarily many people each as badly off as the currently worst off person.

I will omit the details of the Spectrum Argument which implies this implausible conclusion, since I believe the reader can, at this point, competently and

charitably fill them in.[87] I simply want to make the reader aware of Scarier Priority Monster.

I quickly considered a "partners in crime" reply to Weighting Spectrum Arguments. This reply simply points out that there are other views (the Equal Weight View conjoined with the Total View) which imply many of the conclusions of Aggregating Spectrum Arguments, which are comparable in implausibility to the conclusions of Weighting Spectrum Arguments. My partial rejoinder to the "partners in crime" reply was that sometimes the conclusions of Weighting Spectrum Arguments (e.g., Scary Pooled Priority) are *more* implausible than what would seem to be the corresponding conclusions of Aggregating Spectrum Arguments (e.g., Headaches). I shall say more about the relative implausibility of the conclusions of Weighting and Aggregating Spectrum Arguments in chapter 7.

But first, I will explore a different possible reply. This reply, which I call *Large Number Skepticism*, is a fairly general reply to Spectrum Arguments which some proponents of the Total View have adopted. Some defenders of the Priority View might also adopt it.

4.3 Intuitions about Large Number Cases

---

[87] Roughly, the idea is to begin the Series with the achieve arbitrarily many, arbitrarily large benefits for arbitrarily many people as badly off as the currently worst off person, and then gradually decrease the size and number of benefits until one reaches an arbitrarily small benefit for the priority monster.

In addition to being Spectrum Arguments, there is something else that both Weighting Spectrum Arguments and Aggregating Spectrum Arguments have in common. In particular, the conclusions of these arguments all involve large numbers or quantities. Consider some of the conclusions of the Weighting Spectrum Arguments: Priority Monster and Scary Pooled Priority. And consider some of the conclusions of the Aggregating Spectrum Arguments: Utility Monster, Repugnant Conclusion, and Hangnails for Torture. In Priority Monster and Utility Monster, we are to consider, respectively, someone *very* badly off and someone *very* well off. In both Scary Pooled Priority and Repugnant Conclusion, we are to consider a *very* large number of people. And in Hangnails for Torture, we are to consider a *very* long session of mildly annoying hangnail pain.

Some philosophers have claimed that since we cannot relevantly imagine the large numbers or quantities involved in the conclusions of these Spectrum Arguments, we should put less epistemic weight on our intuitions about these conclusions. And they sometimes suggest that, if we could relevantly imagine the conclusions of these Spectrum Arguments, we would have different intuitions about them – in particular, we would not find these conclusions implausible. I will call these philosophers *Large Number Skeptics*.[88] To help explain their position, let me point out two things it is not.

First, Large Number Skeptics do not claim that *all* of our beliefs about cases that involve large quantities should be discounted. For instance, consider:

---

[88] They include: Greene 2001, Broome 2004, 55-59, Huemer 2008b, 907-911, and Cureton 2009.

Less Pain. If other things are equal, it is better to spare someone of 100,000 years of pain than it is to only spare her of 1,000 years of pain.

This is a claim about a large number case. But it is importantly different from the conclusions of Spectrum Arguments. This is because neither the existence nor the reliability of our belief that Less Pain is true seems to depend on our imagining thousands of years of pain. We can infer that Less Pain is true from the plausible principle that "other things equal, less pain is better" and the true belief that 1,000 years of pain is less pain than 100,000 years of pain.

But our beliefs or intuitions about the conclusions of Spectrum Arguments at least arguably seem different, in that their existence and reliability do seem to depend on our relevantly imagining large numbers or quantities. By way of illustration, let us reconsider the:

Repugnant Conclusion. For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living.

Recall that population A contains ten billion people, all with a very high quality of life, and that population Z contains an arbitrarily large number of people, all with lives that are barely worth living. Unlike Less Pain, here we cannot apply a

simple principle about *number* "other things equal, it is better if there are more people with worth living lives," or a simple principle about *quality* "other things equal, it is better if people have higher quality lives." We cannot do this because, in Repugnant Conclusion, other things are not equal – in comparing the A population with the Z population we have to determine whether the difference in quality of life between A and Z is compensated by the difference in number between A and Z. And to make this determination, it seems that we need to relevantly imagine both the difference in quality and the difference in number. But, so the above worry goes, we are unable to relevantly imagine the difference in number between A and Z, and this will inappropriately bias our intuitions to favor population A. At worst, we will imagine A and Z to contain roughly the same number of people ("a very large number"), and form the intuition that A is better than Z solely on the basis of the fact that the people in A have a higher quality of life.[89]

Second, Large Number Skeptics do not claim that, to relevantly imagine the conclusions of Spectrum Arguments, we must *precisely* imagine differences between large numbers or quantities. We might need to precisely imagine such differences if, for example, we sought to rely on our intuitions to establish precise tradeoff rates between quality and number of lives – to determine exactly how many people living at quality of life A is as good as 100 billion people living at quality of life Q. But, in considering the conclusions of Spectrum Arguments, establishing tradeoff rates (even imprecise ones) is not our main task. Rather, it is to determine, for example, whether

---

[89] I am here indebted to Greene 2001.

there is *any* size of population Z that would be better than population A. Thus, Large Number Skeptics would be satisfied if we could imagine, even in an imprecise way, the vast difference in number between A and Z. But they doubt that we can do this.

Large Number Skeptics claim that, even if our intuitions tell us that the conclusions of the above Spectrum Arguments are implausible, we should at least be somewhat reluctant to trust them, or put less weight on them than intuitions that do not relevantly involve large numbers. If these Skeptics are right, they would have found a strong objection against my Weighting Spectrum Arguments. Defenders of the Priority View or the Equality View might appeal to this objection.

In 4.3.1 and 4.3.2, I will consider some different forms of Large Number Skepticism. Then in 4.3.3, I will offer a response.

4.3.1 Imagining Monsters versus Imagining Masses

Parfit is only partly a Large Number Skeptic, or subscribes only to a limited sort of Large Number Skepticism. He suggests that our intuition that Utility Monster is false might be unreliable in a way in which our intuition that Repugnant Conclusion is false is not. He writes:[90]

---

[90] Parfit 1984, 389.

…this Monster's quality of life must be *millions* of times as high as that of anyone we know. Can we imagine this? Think of the life of the luckiest person you know, and ask what a life would have to be like in order to be a million times as much worth living. The qualitative gap between such a life and ours, at its best, must resemble the gap between ours, at its best, and the life of those creatures who are barely conscious—such as, if they *are* conscious, Plato's 'contented oysters'. It seems a fair reply that we cannot imagine, even in the dimmest way, the life of this Utility Monster. And this casts doubt on the force of the example. Act Utilitarians might say that, if we really could imagine what such a life would be like, we might not find Nozick's objection persuasive.

He then goes on to write:

Return now to my imagined Z. This imagined population is just another Utility Monster. The difference is that the greater sum of happiness comes from a vast increase, not in the quality of one person's life, but in the number of lives lived. And *my* Utility Monster is neither deeply impossible, nor something that we cannot imagine. We can imagine what it would be for someone's life to be barely worth living. And we can imagine what it would be for there to be many people with such lives. In order to imagine Z, we

merely have to imagine that there would be *very* many. This we can do. So

the example cannot be questioned as one that we can hardly understand.

Parfit evidently believes that there is an important difference in our ability to

imagine extremely large *numbers* of individually imaginable things, on the one hand,

and our ability to imagine things which are *qualitatively* very unlike what we are used

to, such as what it is like to be a utility monster, an oyster, or a bat. (And the

qualitative difference between us and oysters or bats might be much smaller than that

between us and utility monsters). Since our powers of imagination are suited to do the

former, but not the latter, we can relevantly imagine Repugnant Conclusion but not

Utility Monster.

But it seems that if we can imagine extremely large *numbers* of individually

imaginable things, then we can also imagine extremely long *durations* of things which

we can imagine lasting for shorter periods of time. For example, if we can relevantly

imagine what it is like to experience a headache for one minute, we can relevantly

imagine what it is like to experience a headache for one billion years. We could

rephrase this in terms of imagining an extremely large number of individually

imaginable things: if we can relevantly imagine one headache lasting for a minute, we

can relevantly imagine 525,600,000,000,000 headaches each lasting for a minute lined

up back to back – or, equivalently, we can imagine a headache lasting for one billion

years.

Why might Parfit distinguish between imagining large numbers of individually imaginable things and long durations of things which are imaginable for short durations, on the one hand, and imagining things which are qualitatively very unlike what we are used to, on the other hand? Parfit might think that, as long as we can imagine something, we can imagine *replicating* this thing indefinitely. In this way, we can imagine large numbers of individually imaginable things and long durations of things which are imaginable for short durations. But we cannot, in this way, imagine qualitatively very different things. This is because there may be no idea for our imaginations *to* replicate which could approximate these things. For example, we cannot imagine pleasure many times more intense than any pleasure we have experienced by simply imaginatively replicating the latter. More less intense pleasure is not the same thing as more intense pleasure. We can clearly imagine pleasure *somewhat* more or less intense than pleasure we have experienced. But there seem to be limits on how far we can extrapolate from pleasure that that we have experienced, or can imagine.[91]

_____

[91] For reasons which I will not fully explore here, I suspect that John Stuart Mill might draw a similar (or the same) distinction between imagining large numbers of individually imaginable things and imagining things which are qualitatively very unlike what we are used to. Mill believes that to be a competent judge of which of two pleasures is better, one must be acquainted with both (1861, chapter 2, especially paragraphs 4 through 8). However, Mill believes that as long as we are acquainted with the best sorts of pleasure available to humans and the best sorts of pleasure available to pigs, we can competently judge that a life of the former pleasures is better than a life, *however long*, of nothing but the latter pleasures. Mill thinks this, I suspect, because he thinks it does not matter if we do not know anyone who has ever lived a *very* long life with nothing but pig pleasures, for we can relevantly imagine such a life by imaginatively replicating very many individually imaginable pig pleasures. By contrast, someone who was only acquainted with the pig pleasures could not become a competent judge of the human life by imaginatively replicating these pig pleasures (human pleasures are too qualitatively different from pig pleasures).

More generally, there seem to be limits on how far we can extrapolate from *quality of life* that that we have experienced, or imagined. And imagining quality of life which is *millions* of times as high as that of anyone we know seems beyond such limits.

Let us assume, with Parfit, that there is such a relevant and important distinction between imagining large numbers and imagining qualitatively very different things or beings. Before considering another form of Large Number Skepticism, we should perhaps be explicit about what Parfit's distinction would imply about various cases.

Recall that there are different versions of Utility Monster. There is the Intensity Utility Monster, in which the utility monster is made very well off *at a time* (e.g., he is given arbitrarily intense pleasure). And there is the Duration Utility Monster, in which the utility monster is made very well off *over a long period of time* (e.g., he is given moderately intense pleasures for billions of years). Parfit's distinction would rule out Intensity Utility Monster as unimaginable. But, insofar as it rules in Repugnant Conclusion as relevantly imaginable, it would presumably also rule in Duration Utility Monster, Intrapersonal Repugnant Conclusion, Hangnails for Torture, Pooled Priority, and Scary Pooled Priority. And recall two versions of Priority Monster: one in which the priority monster is arbitrarily badly off in virtue of being arbitrarily badly off at a time, and one in which the priority monster is arbitrarily badly off in virtue of having an arbitrarily long life which is significantly but not unimaginably worse than nothing at each time. Parfit's distinction would rule out the

former version of Priority Monster, but rule in the latter version as relevantly

imaginable.

4.3.2   Imagining Large Numbers versus Imagining *Arbitrarily* Large Numbers

The simple form of Large Number Skepticism[92] simply states that since we

cannot relevantly imagine very large numbers or quantities, our intuitions about

certain cases which seem to require us to do so are unreliable, or less reliable. Such

cases would presumably include those ruled out by Parfit's distinction as

unimaginable, i.e., the "intensity" versions of Utility Monster and Priority Monster.

But these cases would also include the "duration" versions of Utility Monster and

Priority Monster, as well as Repugnant Conclusion, Pooled Priority, and so on. I

believe I have already, at the beginning of 4.3, explained this simple form of Large

Number Skepticism. We just saw how it differs from Parfit's limited form of Large

Number Skepticism. I will now discuss a third form of Large Number Skepticism,

introduced by John Broome.[93]

The particular sort of large number claims which concern Broome contain a

quantifier, and are of the form:

---

[92] This version of Large Number Skepticism is espoused by Greene, Huemer, and Cureton (see earlier footnote).

[93] Broome 2004, 55-9. The following discussion in this subsection is largely owed to personal communication with Broome.

(*)   There is no number *n* such that B(*n*) is better than A, where B(*n*) gets better as *n* increases.

To reject this claim intuitively, one might have to imagine a large number. It may be possible to do this. However, to *accept* this claim intuitively, one might have to imagine *every* large number – to intuitively check every possible value for *n*. It is impossible, at least for finite beings like us, to do this.

In addition to the quantifier, "there is no number *n* such that," it is important to note that claims of form **(*)** contain the qualification, "where B(*n*) gets *better* as *n* increases." Suppose, for instance, that A is one billion people living superb lives and that B(*n*) is *n* people living hellish lives. Here B(*n*) would not get better as *n* increases (it would get worse!). Because of this, and because for any *n* we can imagine, it is implausible that B(*n*) is better than A, it is not necessary to imagine every large number to intuitively accept the claim that there is no number *n* such that B(*n*) is better than A.

Suppose that B(*n*) is *n* people living with nothing but muzak and potatoes (Z-lives), and A is ten billion people living superb lives (A-lives). And suppose we deny the Diminishing Value View, and claim that B(*n*) gets better as *n* increases. Then our denial of the Repugnant Conclusion would be of form **(*)**.

Since this denial of the Repugnant Conclusion fits form (*****), to accept this denial of the Repugnant Conclusion intuitively, one might have to imagine every large number, to intuitively check every possible value for n – which, again, is impossible.

Before proceeding, it is important to notice that Broome's form of Large Number Skepticism need not stand opposed to an intuitive rejection of the Total View. (Recall that, according to the latter, if other things are equal the outcome with more total well-being is better). The Total View entails the conclusion:

B(n) is better than A.

Where A and B(n) are as before and n is the next integer bigger than ten billion times the well-being level of a superb life divided by the well-being level of a life with nothing but muzak and potatoes.[94] The denial of this unquantified conclusion is not a claim of form (*****). So we can, consistently with Broome's form of Large Number Skepticism, have reliable intuitions that count against the Total View. However, notice that, insofar as n is a large number, the simple form of Large Number Skepticism would not regard this intuition against the Total View as reliable. Lastly, note that it is not just the denial of the Repugnant Conclusion which might be of form (*****). The denial of Priority Monster, Utility Monster, Hangnails for Torture, and so on, might also be of form (*****). When this is so, Broome's Large Number

---

[94] If n corresponds to the number of lives in B(n), if each person in B(n) has well-being level b (muzak and potatoes life), and if each of the ten billion people in A has well-being level q (superb life), then the Total View would imply that B(n) is better than A if and only if $n > 10$ billion $\times (q / b)$.

Skepticism would imply that these conclusions cannot reliably be denied on the basis of intuitions.

4.3.3   A Response to Large Number Skepticism[95]

All three forms of Large Number Skepticism – the simple form, Parfit's form, and Broome's form – share a common argument structure:

(1)   We cannot imagine X.

(2)   In order to have reliable intuitions about P, we must imagine X.

So,   (3)   We cannot have reliable intuitions about P.

Although I am not sure whether or when (1) is true, I am prepared to concede that it is true it for all three forms of Large Number Skepticism. However, depending on what is meant by "have intuitions about P," I believe we should either deny (2) or else accept these Skeptical arguments but claim that they have remarkably little dialectical force. Below I will focus on Repugnant Conclusion and Intensity Utility Monster, but what I say about these cases will apply, *mutatis mutandis*, to other similar large number cases (e.g., Priority Monster).

---

[95] The ideas contained here have evolved into a short publication titled "Intuitions about large number cases" – see Pummer 2013.

Consider Repugnant Conclusion. It seems true that the more people with Z-lives (barely worth living lives) we add to population Z, the better, and it seems true that we *can* relevantly imagine population Z being *many times bigger* than population A. We can relevantly imagine scaling up the size of Z, such that we imagine it as twice, ten, or perhaps even one hundred times as large as A. We can consider the claim that population A is better than population Z. And crucially, as we imagine scaling up the size of Z (such that we imagine it as twice, ten, or perhaps even one hundred times as large as A), we do not become any *less confident* in our judgment that population A is better than population Z.[96] Now consider the following two hypotheses:

> *Hypothesis 1.* If we could imagine an arbitrarily large population Z, we would judge that there is some size of population Z such that it is better than population A.

---

[96] Joshua Greene (2001, 9-10) considers a different scaled-down version of the Repugnant Conclusion, which it is worth quickly noting. First, he imagines that the quality of life in A is one thousand times greater than the quality of life in Z. Then, if A contains *n* people, for the Total View to imply the Repugnant Conclusion, Z must contain at least $1000n + 1$ people. If A contains ten billion people with high quality lives, then Z contains ten trillion and one people with barely worth living lives (each only one thousandth the quality of an A-life). Then Greene has us consider the following scaled-down case: in A* there is one person with a high quality life and in Z* there are one thousand and one people with barely worth living lives. We are asked which is better. Greene claims that our intuitions about this revised case, involving A* and Z*, are not as firm or strong as our intuitions about the Repugnant Conclusion. He suggests that our insensitivity to the vast number of lives in Z at least partly explains why our intuitions about the revised case are less strong.

This does not, I believe, undermine my point in the main text. A* and Z* are not suitable surrogates for A and Z, respectively. The claim that no number of people with Z-lives can outweigh *one* person with an A-life is significantly less intuitively plausible than the claim that no number of people with Z-lives can outweigh *ten billion* people with A-lives. The crucial difference, which explains why our intuitions about the revised case involving A* and Z* are less strong than our intuitions about the Repugnant Conclusion, thus lies between A and A*. This difference has nothing to do with (insensitivity to) the size of Z.

And:

> *Hypothesis 2*. If we could imagine an arbitrarily large population Z, we would judge that there is no size of population Z such that it is better than population A.[97]

It seems plausible that if *Hypothesis 1* were true, we *would* lose at least some confidence in our judgment that population A is better than population Z, as we imagine population Z containing more and more people with Z-lives (e.g., twice, ten, one hundred times as many people as in A). *But*, since we do *not* lose any confidence in this way when we do this, this casts doubt on the plausibility of *Hypothesis 1*. Indeed, it suggests that *Hypothesis 2* is significantly more plausible.

Thus, we may not need to imagine an arbitrarily large population Z to support plausibly the claim that there is *no* number of people with Z-lives that would be better than ten billion people with A-lives. Instead, we could plausibly arrive at this claim,

---

[97] What I argue here carries over to discussions of Large Number Skepticism in the context of Priority Monster. Consider, for example:

> *Hypothesis 1**. If we could imagine an arbitrarily badly off person (a Priority Monster), we would judge that there could be some sufficiently badly off person such that it would be better to achieve a piddling benefit for her than it would be to achieve a very large benefit for the currently worst off person.

And:

> *Hypothesis 2**. If we could imagine an arbitrarily badly off person (a Priority Monster), we would judge that there could not be a person so badly off that it would be better to achieve a piddling benefit for her than it would be to achieve a very large benefit for the currently worst off person.

via a kind of inductive generalization, from intuitions we have about cases (involving smaller sizes for Z) that we can relevantly imagine.

Return to (2) in the Skeptical argument, "in order to have reliable intuitions about P, we must imagine X." In the context of Repugnant Conclusion, P would be the claim that there is no number of people with Z-lives that would be better than ten billion people with A-lives. If offering a plausible defense of this P via the above kind of inductive generalization from intuitions we have about cases that we can relevantly imagine counts as "having intuitions about P," then I believe that (2) is false. But if this does *not* count as "having intuitions about P," then while (2) might be true and the Skeptical arguments might be sound, these arguments would have remarkably little dialectical force. Those of us who found Repugnant Conclusion implausible would be forced to concede a small point to the Large Number Skeptics: that we do not have reliable intuitions *about* the denial of Repugnant Conclusion. Instead, we would have to say that we have reliable intuitions which, though not about the denial of Repugnant Conclusion, nonetheless plausibly *support* the denial of Repugnant Conclusion.

So either (2) of the Skeptical argument is false, or it – with (1) – only delivers a conclusion which falls very far short of the aim of Large Number Skepticism. The aim of Large Number Skepticism is, I believe, to supply an undercutting defeater of the reasons for denying the conclusions of the various Spectrum Arguments listed above. The reasons for denying these conclusions are given by intuitions we have. At least some of the relevant reasons have not been defeated by any form of Large Number Skepticism. This does not mean that we are justified in believing that the

conclusions of these Spectrum Arguments are false. But it means that Large Number

Skepticism has failed to diminish or undercut our intuition-based reasons for believing

that these conclusions are false.

We might believe, with Parfit, that there is a relevant and important distinction

between imagining large numbers and imagining qualitatively very different things or

beings (recall 4.3.1). If so, we might think that the above reply to Large Number

Skepticism would need to be modified to deal with the "intensity" versions of Utility

Monster and Priority Monster. I will focus on Intensity Utility Monster. And, for

simplicity, I will focus on the intensity of the monster's *pleasure* (but I believe that

similar remarks will apply to versions of Intensity Utility Monster in which the

monster's quality of life at a time is scaled-up along non-hedonic dimensions).

Recall that this utility monster would receive *extremely* intense pleasure from

being fed people. This pleasure would be at least millions of times more intense than

any pleasure any of us has ever had. Perhaps we cannot, as Parfit suggests, relevantly

imagine such extremely intense pleasure. But it seems true that the more intense the

utility monster's pleasure is, the better, and it seems true that we *can* relevantly

imagine some pleasures which are *many times more intense* than others, even those

which are not trivial. We can relevantly imagine scaling up the intensity of the utility

monster's pleasure, such that we imagine it as (roughly) twice, ten, or perhaps even

one hundred times as intense as some moderately intense pleasures we are familiar

with (in my case, at least, petting a cat, stepping into a warm shower on a cold day, or

drinking a can of soda normally induce what I would consider moderately intense

pleasures). We can consider the claim that it would be better if everyone received very large benefits than it would be if everyone were fed to the utility monster. And crucially, as we imagine scaling up the intensity of the pleasure the utility monster would receive from being fed everyone (such that we imagine the monster's pleasure as twice, ten, or perhaps even one hundred times as intense as moderately intense pleasures), we do not become any *less confident* in our judgment that it would be better if everyone received very large benefits than it would be if everyone were fed to the utility monster. Now consider the following two hypotheses:

> *Hypothesis 3*. If we could imagine the utility monster receiving an arbitrarily intense pleasure, we would judge that there is some level of intensity such that it is better if everyone were fed to the monster than if everyone received very large benefits.

And:

> *Hypothesis 4*. If we could imagine the utility monster receiving an arbitrarily intense pleasure, we would judge that there is no level of intensity such that it is better if everyone were fed to the monster than if everyone received very large benefits.

As with *Hypothesis 1*, we might claim that it seems plausible that if *Hypothesis 3* were true, we *would* lose at least some confidence in our judgment that it would be better if everyone received very large benefits than it would be if everyone were fed to the utility monster. And we might again add that since we do *not* lose any confidence in this way, this casts doubt on the plausibility of *Hypothesis 3* – and indeed that this suggests that *Hypothesis 4* is more plausible. Thus, we may not need to imagine the utility monster's arbitrarily intense pleasure to claim plausibly that it would be better if everyone received very large benefits than it would be if everyone were fed to the utility monster, no matter how intense the pleasure the monster would receive from this would be.

However, there might be an important disanalogy between Hypotheses 1 and 2, on the one hand, and Hypotheses 3 and 4, on the other. In particular, while it seems fairly plausible that if Hypothesis 1 were true, we would lose confidence in our judgment that population A is better than population Z as we imagine Z containing more and more people with Z-lives, *it might well be false* that if Hypothesis 3 were true, we would lose confidence in our judgment that it would be better if everyone received very large benefits than if everyone were fed to the utility monster, as we imagine the monster receiving more and more intense pleasure. There might be a disanalogy here insofar as arbitrarily intense pleasure is *qualitatively* different from moderately or even very intense pleasure that we are used to. Imagining more and more intense pleasure is thus a *completely* irrelevant test of our intuitions, insofar as

the pleasure we are imagining remains of a qualitatively different (less intense) kind

from that which the utility monster.

I believe that there probably is such a disanalogy, and that it is important. But

for all that, it would *not* support Hypothesis 3 over Hypothesis 4. It would instead

only remove *one* kind of support we could have had for Hypothesis 4 over Hypothesis

3. But we might think that Hypothesis 4 is more plausible than Hypothesis 3 for other

reasons, which seem strong. We might claim that we simply do not believe that we

need to know how intense the utility monster's pleasure is to be confident that it would

be worse if everyone were fed to the monster. We might plausibly believe that we

would judge this, even if we could imagine every possible level of intensity of

pleasure. And we might plausibly be highly dubious of the Skeptic's claim that the

only reason we find Intensity Utility Monster implausible is that we are not imagining

how intense the monster's pleasure would be. Though the Skeptic's claim remains a

possibility, we might think that all of our available evidence points to the contrary.

While my reply to Large Number Skepticism about the "intensity" versions of

Utility Monster and Priority Monster is thus somewhat different in certain details than

my reply to Large Number Skepticism about the "duration" versions of Utility

Monster and Priority Monster (as well as Repugnant Conclusion, Pooled Priority, and

so on), the basic gist of the reply remains the same. To repeat, either (2) of the

Skeptical argument is false, or it – with (1) – only delivers a conclusion which falls

very far short of the aim of Large Number Skepticism. The aim of Large Number

Skepticism, I believe, is to supply an undercutting defeater of the reasons for denying

the conclusions of the various Spectrum Arguments listed above. The reasons for denying these conclusions are given by intuitions we have. At least some of the relevant reasons have not been defeated by any form of Large Number Skepticism. This does not mean that we are justified in believing that the conclusions of these Spectrum Arguments are false. But it means that Large Number Skepticism has failed to diminish or undercut our intuition-based reasons for believing that these conclusions are false.

In sum: we should continue to rely on our intuitions which are (at least in some sense) about the conclusions of the Spectrum Arguments mentioned above, as an important though defeasible sort of evidence. The intuitions which support the denials of these conclusions are not plausibly discounted or dismissed on the grounds that they involve large numbers or things which are qualitatively very unlike what we are used to.

Chapter 5

Vagueness

CHAPTER OUTLINE

5.1   Sorites Arguments

The most important arguments I offer against rivals to the Equal Weight View are Spectrum Arguments. Recall that these arguments have the following form:

P1.   B is better than A

P2.   C is better than B

P3.   D is better than C

     …and so on…

P(n-1).   Y is better than X

P*n*.   Z is better than Y

P*T*.   Transitivity

∴   Z is better than A

Spectrum Arguments against rivals to the Equal Weight View typically take the form of a dilemma. Either the rival view accepts P1 through P*n* or it does not. If it accepts P1 through P*n*, it will, given Transitivity, imply an implausible conclusion. If the rival view does not accept P1 through P*n*, there are two possibilities: it accepts some but not all of P1 through P*n*, or it rejects all of P1 through P*n*. If the former, the rival view must implausibly deny the Conditional or Parity (or some analog of Parity). If the latter, then the view may differ from the Equal Weight View only in trivial ways, e.g., the Tie-Breaker Priority View.

Some people have claimed that Spectrum Arguments are, or are relevantly like, *Sorites Arguments*.[98] In the classic Sorites Argument, it is "proven" that a single grain of sand is a heap. The reasoning begins with a *Soritical Premise* which says, "if collection of grains of sand *k* is a heap, then collection *k*+1, which contains only one fewer grain of sand, must also be a heap," and then proceeds:

(i)   Take some sizable collection of grains of sand which is a heap, and remove one grain of sand from it,

(ii)   By the Soritical Premise, the resulting collection is a heap,

---

[98] For example, see Griffin 1986, 86-7.

(iii)   Now repeat the steps, i.e., take the resulting collection which (by the

          Soritical Premise) is a heap, and remove one grain of sand from it, and

          so on…

Such reasoning, which entails that a single grain of sand is a heap, is unsound.

Sorites Arguments, of course, do not just pertain to heaps of sand, but a wide variety

of vague predicates, e.g., "is a heap," "is red," "is bald," and so on. In general, if a

predicate is vague, sometimes it is indeterminate whether it applies. Sorites

Arguments have the following form, where predicate P is a vague predicate, and

where each item differs only slightly from its predecessor:

S1.   Item 1 is a P

S2.   If item 1 is a P, then item 2 is a P

S3.   If item 2 is a P, then item 3 is a P

     …and so on…

S(n-1).   If item n-2 is a P, then item n-1 is a P

Sn.   If item n-1 is a P, then item n is a P

MP.   Modus Ponens

∴   Item n is a P

Sorites Arguments might seem to resemble Spectrum Arguments. After all,

both sorts of argument appear to share the following feature:

Slight Differences: These arguments claim that each slight difference between

adjacent items makes no difference as far as the continued

accurate application of the predicate in question at each step is

concerned.

Given this, we might think that Spectrum Arguments are Sorites Arguments, or at least that they are similarly unsound. If Spectrum Arguments are (relevantly like) versions of Sorites Arguments, perhaps we can claim that they too are unsound. In this chapter, I will respond to this and related objections to Spectrum Arguments.

Parfit, Rachels, Temkin and others have written on whether Spectrum Arguments are relevantly disanalogous to Sorites Arguments,[99] but I think what I say below (in 5.3) explains in a more direct and compact way the key disanalogy between the two sorts of argument.

5.2   Slight Differences

[99] See Parfit 1986b, footnote 13, Rachels 1998, 74, and Temkin 2012, chapter 10. My views on this topic were influenced the most by Temkin's work.

First, I will note a different possible disanalogy. Parfit[100] and Rachels[101] once claimed that the difference between Spectrum Arguments and Sorites Arguments is that the slight differences between adjacent items in Sorites Arguments *make no difference* to the application of (for example) "is a heap," whereas the slight differences between adjacent items in Spectrum Arguments *do* make a difference to the application of "is better than." The latter is true, for example, because it is relevant to the application of "is better than" both how intense pain is and how long it lasts; and so the fact that one pain is slightly less intense than another is relevant to the application of "is better than."

While this is true, it is not clear that it amounts to an important disanalogy between Spectrum Arguments and Sorites Arguments. It does not overturn, for example, the fact that both Spectrum Arguments and Sorites Arguments share Slight Differences: each sort of argument claims that each slight difference between adjacent items makes no difference as far as the *continued accurate application* of the predicate in question at each step is concerned. In Spectrum Arguments, the slight differences are *relevant* to whether "is better than" applies (and so "make a difference" to the application of "is better than" in *some* sense), and perhaps that is not true in Sorites Arguments. But for all that, in Spectrum Arguments slight differences do not make a difference to the continued accurate application of "is better than." They do not change the fact that this predicate applies (or seems to apply) at each step, at P1 through P*n*.

---

[100] Parfit 1986b, footnote 13.

[101] Rachels 1998, 74.

Are Sorites Arguments and Spectrum Arguments are truly and importantly disanalogous? What we are looking for is a feature of Sorites Arguments which makes them unsound, which is shared by Spectrum Arguments. But there are many arguments which share Slight Differences, which are obviously sound. Consider the *Old Man Argument*, which begins with the following Series of old men:

(1)   Luke is exactly 90 years old

(2)   Tom is 89 years, 364 days, 23 hours, 59 minutes, and 59 seconds old

(3)   Dick is 89 years, 364 days, 23 hours, 59 minutes, and 58 seconds old

(4)   Harry is 89 years, 364 days, 23 hours, 59 minutes, and 57 seconds old

…and so on…

(*n*-1)   Shelly is 80 years and 1 second old

(*n*)   Wesley is exactly 80 years old

Now consider the following Premises:

P1.   Luke is older than Tom

P2.   Tom is older than Dick

P3.   Dick is older than Harry

…and so on…

P*n*.   Shelly is older than Wesley

P*T*.   "Older than" is transitive

∴   Luke is older than Wesley

This argument is obviously sound. Yet it possesses the feature, Slight

Differences. So it seems that this feature is not an unsoundness-making feature.

But notice that the Old Man Argument does not *essentially rely* on the feature

that each slight difference makes no difference as far as the continued accurate

application of "is older than" is concerned. The Premises of the argument would have

been just as plausible if the age differences between each old man in the Series were

significantly greater (rather than each being a second younger, each might have been a

year younger). So perhaps we should revise Slight Differences thusly:

Slight Differences*:   The argument *essentially relies* on the claim that each slight

difference between adjacent items makes no difference as far as

the continued accurate application of the predicate in question at

each step is concerned.

Though it is shared by Sorites Arguments and Spectrum Arguments, but not the obviously sound Old Man Argument, Slight Differences* still does not seem to be the unsoundness-making feature we are looking for.

As it happens, the predicate "is *just barely* older than" is not transitive. (Each old man is just barely older than his successor, but Luke is more than "just barely" older than Wesley). But suppose, for the sake of argument, that "is just barely older than" were transitive. It is hard to see, then, what would be problematic about giving the Old Man Argument as a Spectrum Argument for the conclusion that Luke is just barely older than Wesley. If there is anything problematic about this Spectrum Argument, I believe, it is simply that "is just barely older than," is not transitive. But then this is an issue about *transitivity*, not about Slight Differences*. And in the case of the predicate "is better than," most people already accept that it is transitive. (I will return to the question of Transitivity in chapter 6).

Lastly, we might think that Slight Differences* is an unsoundness-making feature insofar as we think our intuitions about items which differ only slightly are, in general, less reliable than our intuitions about items which differ significantly.[102] While this hypothesis might be true in some cases, it seems overly broad and clearly false in other cases. For example, the greater the difference between the items in the Old Man Series, the less reliable our intuitions that "is just barely older than" applies will be. Similarly, our intuitions about the Premises of many Spectrum Arguments seem much firmer when the items in the relevant Series differ much less in one way

---

[102] This is suggested by Parfit in "Towards Theory" (unpublished manuscript) and by Voorhoeve 2008.

(e.g., each pain is *slightly* less intense) and much more in another way (e.g., each pain lasts *much* longer).

5.3   Disanalogies

While Slight Differences* is shared by Spectrum Arguments and Sorites Arguments, it does not seem to be an unsoundness-making feature. What makes Sorites Arguments unsound, I believe, is that they essentially rely on the repeated application of a Soritical Premise to a vague predicate. (It is debatable exactly *how* this feature makes these arguments unsound, though, and I will not enter this debate; I just claim that this feature *does* seem to make them unsound).[103] And Spectrum Arguments do not essentially rely on the repeated application of a Soritical Premise to a vague predicate. So they cannot be dismissed as unsound in the way that Sorites Arguments are.

5.3.1   Soritical Premises versus Intuitions

---

[103] See Williamson 1994 and Keefe 2000. For an interesting and plausible solution to Sorites Arguments, see Raffman 1994. For related discussions of Sorites and vagueness in connection with Transitivity, see Quinn 1990, and Alastair Norcross's excellent reply in Norcross 1997, particularly section IV.

Sorites Arguments essentially rely on the repeated application of a Soritical Premise to a vague predicate. It is this feature which renders Sorites Arguments unsound and which makes them even mildly tempting. The latter can be illustrated by considering each collection of grains of sand in a Sorites Argument independently of any Soritical Premise, and asking whether "is a heap" seems to apply to it.

Collection 1:   100,000,000 grains

Collection 2:   99,999,999 grains

Collection 3:   99,999,998 grains

                    …and so on…

Collection 99,999,998:   3 grains

Collection 99,999,999:   2 grains

Collection 100,000,000:   1 grain

You are undoubtedly very confident that "is a heap" does not apply to Collection 100,000,000. Perhaps somewhere between the first and last collection you are not sure whether "is a heap" applies. But without the Soritical Premise, you would not at all be tempted to believe that Collection 100,000,000 is a heap. Our judgments about the collections in this series, considered independently, speak loudly against the Sorites Argument. If and when people are tempted to say that Collection 100,000,000 is a heap, it is probably only because they have been suckered by the Soritical Premise in the following way:

S1.   Collection 1 is a heap

S2.   If Collection 1 is a heap, then Collection 2 is a heap

S3.   If Collection 2 is a heap, then Collection 3 is a heap

…and so on…

S(n-1).   If Collection 99,999,998 is a heap, then Collection 99,999,999 is a heap

Sn.   If Collection 99,999,999 is a heap, then Collection 100,000,000 is a heap

MP.   Modus Ponens

$\therefore$ Collection 100,000,000 is a heap

What partly makes the Soritical Premise (and consequently S2 through Sn) so tempting is that it is a *conditional*. It "merely" says that *if* Collection k is a heap, *then* Collection k+1 is a heap. But it does not say *whether k* is a heap. Nonetheless, the *repeated application* of the Soritical Premise, together with S1 and modus ponens (each of which is incontrovertible), conflicts with our judgments about whether "is a heap" applies to various collections of grains of sand (at least those near Collection 100,000,000).

While our intuitions that "is a heap" applies diminish as we move from Collection 1 toward Collection 100,000,000, many of us find it plausible that "is better than" *always* applies at each step in (certain) Spectrum Arguments, regardless of

whether we are considering P2 or Pn-1. This is an extremely important difference

between Sorites Arguments and Spectrum Arguments.

If the reason we were attracted to the Premises of the Spectrum Argument was

that we accepted and reiteratively applied a Soritical Premise, e.g., "if 'is better than'

applies at Pk, then 'is better than' applies at Pk+1," *then* we may be legitimately

charged with falling for a Sorites. But notice how the Spectrum Argument would have

to be changed to match the structure of the above Sorites Argument:

S1.   B is better than A (the first Premise, or P1)

S2.   If B is better than A, then C is better than B

S3.   If C is better than B, then D is better than C

…   …

S(n-1).   If X is better than W, then Y is better than X

Sn.   If Y is better than X, then Z is better than Y

MP.   Modus Ponens

∴   C is better than B, D is better than C, and so on…, Z is better

                than Y (i.e., the rest of the Premises, or P2 through Pn)

PT.   Transitivity

∴   Z is better than A

But the Premises of Spectrum Arguments are not supported by S1 through Sn.

I never claimed that Pk+1 is true *because* Pk is true. Rather, I simply claimed that for

some formulations of the Series, each of the Premises seems, considered independently, intuitively plausible (or no less intuitively plausible than other Premises). I simply appealed to the intuitions that I and many others have about the Premises. In other words, S1 (the first Premise) and everything in the line following *MP* (the rest of the Premises) just seem plausible; there is no need to appeal to a Soritical Premise to support these claims. Again, by contrast, that "is a heap" applies at each step of a Sorites Argument seems very counterintuitive, and we would not even be mildly tempted to claim that it does if it were not for the repeated application of the Soritical Premise.

Thus, while Sorites Arguments essentially rely on the repeated application of a Soritical Premise to be even minimally tempting, Spectrum Arguments do not. Their steps are independently intuitively plausible.

5.3.2 Conditionals

Recall that, in several weighting Spectrum Arguments (chapters 3 and 4), I made use of the Conditional, which claims that if we accept claim C1 on the basis of its intuitive plausibility, and if claim C2 is at least as intuitively plausible as C1, and if other things are equal, then we should also accept C2.

I just said that is the repeated application of Soritical Premises to vague predicates which seems to make Sorites Arguments unsound. Soritical Premises are

conditionals. We might worry that, also being a conditional, the Conditional is relevantly like a Soritical Premise.

It isn't. A Soritical Premise says that if the predicate of interest applies at step k, then it also applies at $k+1$. It makes no reference whatsoever to how intuitively plausible it is that the predicate of interest applies at step k. For example, a Soritical Premise would imply that "is a heap" applies at step 99,999,998 insofar as the Soritical Premise already "established" that "is a heap" applies at step 99,999,997. And it would imply this regardless of whether it seems intuitively plausible that "is a heap" applies at either step. The Conditional would not have such implications. It would not commit us to the claim that "is a heap" applies at step 99,999,998, since it does not seem intuitively plausible that it would apply at that step (nor does it seem plausible that it would apply at the previous step).

While the Conditional is not itself a Soritical Premise, and while it clearly does not depend on a Soritical Premise, it could commit us to various Soritical Premises. For example, the Conditional might commit us to "if collection k is a heap then $k+1$ is a heap," if it were true that for any collection k we accept that it is a heap on the basis of its being intuitively plausible that it is a heap. But we would never accept the claim that Collection 99,999,998 is a heap on the basis of its own intuitive plausibility. Again, if we ever did accept this claim, it would be because we were suckered by a Soritical Premise (see 5.3.1). On the other hand, if we accept each Premise Pk as intuitively plausible, and each P$k+1$ as at least as intuitively plausible, then the Conditional might commit us to "if 'is better than' applies at Pk, then 'is better than'

applies at Pk+1," and this looks like a Soritical Premise. But it is important to see that this Soritical Premise is argumentatively superfluous; it is not being essentially relied on in the Spectrum Argument. Similarly, S2 through Sn are entirely superfluous in the revised Spectrum Argument in 5.3.1.

In sum: the Conditional is not, and is not relevantly like, a Soritical Premise.

5.3.3   Comparatives

"Is better than" is a comparative. "Is a heap" is not. We might attempt to make Sorites Arguments more analogous to Spectrum Arguments by changing the predicate of interest from "is a heap," to "is no less a heap than." We can call this revised argument *Comparative Sorites*:

S1.   Collection 2 is no less a heap than Collection 1

S2.   Collection 3 is no less a heap than Collection 2

        …and so on…

Sn.   Collection n is no less a heap than Collection n-1

ST.   "Is no less a heap than" (or "at least as much a heap as") is

            transitive

∴   Collection n is no less a heap than Collection 1

However, Comparative Sorites faces an objection that most Spectrum Arguments do not: some of its premises are obviously false. It might seem plausible that *any* Collection *k* is less a heap than Collection *k*-1, precisely because it contains one fewer grain. Though perhaps plausible, we need not advance this claim. It is enough to refute Comparative Sorites if at least *some* Collection *k* is less a heap than Collection *k*-1 (because it contains one fewer grain).

A possibly better attempted analogy might introduce two criteria relevant to the application of "is no less a heap than:" (i) the number of grains, and (ii) the *shape* of the grains.[104] In the *Second Comparative Sorites*, we begin with the following Series of collections of grains of sand:

Collection 1:   a sizable collection of grains of sand which is ideally shaped such that everyone would recognize it to be a heap

Collection 2:   a collection of grains of sand which contains many more grains than Collection 1, but which is shaped slightly less like an ideally shaped heap than (it is slightly more flattened out than) Collection 1

Collection 3:   a collection of grains of sand which contains many more grains than Collection 2, but which is shaped slightly less like

---

[104] In Appendix C of Temkin 2012, Temkin credits Ryan Wasserman with coming up with the following example.

157

an ideally shaped heap than (it is slightly more flattened out

than) Collection 2

…and so on…

Collection *n*:   a collection of grains of sand which contains arbitrarily many

grains of sand, but is a completely flat (as thin as possible)

sheet of grains of sand

Now, predictably, the argument goes:

S1.   Collection 2 is no less a heap than Collection 1

S2.   Collection 3 is no less a heap than Collection 2

…and so on…

S*n*.   Collection *n* is no less a heap than Collection *n*-1

S*T*.   "Is no less a heap than" is transitive

∴   Collection *n* is no less a heap than Collection 1

This Second Comparative Sorites is indeed structurally analogous to a

Spectrum Argument. However, insofar as it is a Spectrum Argument, it is a *bad* one.

It has some false premises. For example, even its first premise seems fairly

controversial. It may well be false that Collection 2 is no less a heap than Collection

1. We might claim that Collection 2 *is* less a heap than Collection 1, because its shape is less heap-like than Collection 1's shape, and the mere fact that it contains more grains of sand does not compensate for this loss off heap-like-ness. Even if this were not true of the first premise of the Second Comparative Sorites, it seems that at least *some* losses of heap-like shape are not compensated by any gain in mere number of grains. (Merely adding more clay to a lump of clay would not make it any more statue-like, and merely adding more grains to a nearly flat non-heap shape would not make it any more heap-like). So, at least some premises of the Second Comparative Sorites are pretty obviously false.

The Second Comparative Sorites is, I think, structurally analogous to a Spectrum Argument. But, unlike the aggregation and weighting Spectrum Arguments which I rely on, this Comparative Sorites has premises which are pretty obviously false.

In sum: unlike standard Sorites Arguments, Spectrum Arguments do not essentially rely on Soritical Premises. And it is this feature, rather than Slight Differences or Slight Differences*, which seems to make Sorites Arguments unsound. Comparative Sorites Arguments, involving "is no less a heap than" rather than "is a heap," do not essentially rely on Soritical Premises. However, unlike Spectrum Arguments, they have premises which are pretty obviously false.

Insofar as there are Comparative Sorites Arguments that have premises which seem very plausible (or at least not obviously false!), they may raise puzzles similar to those raised by Spectrum Arguments. That is, they may either force us to reject

premises which seem plausible, accept conclusions which seem implausible, or seemingly implausibly deny that certain relations are transitive.

It is possible that there is still some further feature which Spectrum Arguments and Sorites Arguments have in common, which putatively renders both sorts of argument unsound. Perhaps this feature is related to the fact that the predicates in Sorites Arguments are vague.

5.4   Vagueness

The predicates in Sorites Arguments are vague. If a predicate is vague, this means that sometimes it is indeterminate whether it applies. In such cases, we cannot say that it is true that the predicate applies to the item in question, but we also cannot say it is false that it applies to this item. Moreover, the reason we cannot say whether the predicate applies is not that we are missing some factual information. Vagueness is different from ignorance. We might know exactly how many grains of sand there are resting on some flat surface. We might also know the precise geometrical relations between these different grains of sand, and so on. Nonetheless, it might be impossible for us to determinately say whether this collection of grains of sand is a *heap* of sand.

There are several different competing accounts of what vagueness consists in.[105] Unfortunately, I cannot, in this chapter, dig deeply into the details of these

---

[105] See the Williamson and Keefe books cited above.

complicated debates. I hope the following discussion of vagueness in connection with Spectrum Arguments is compatible with the most defensible theories of vagueness.

5.4.1 Multidimensionality

We might have thought that the predicate "is better than" is not vague, and thus that this marks another disanalogy between Sorites Arguments and Spectrum Arguments (in addition to the fact that the latter do not essentially rely on Soritical Premises). More generally, perhaps, we might think that though some one-place predicates are vague, the corresponding comparative is not vague.[106] For instance, "is tall" is vague, but "is taller than" seems not to be vague; "is hot" is vague, but "is hotter than" seems not to be vague, "was long ago" is vague, but "was longer ago than" seems not to be vague, and so on. Many believe that "is better than" is the comparative which corresponds to the one-place predicate "is good." They might think that "better than" must be non-vague, even if "is good" is.

However, some comparatives are vague.[107] Some one-place predicates are *multidimensional*. That is, there are multiple factors or dimensions relevant to determining whether they accurately apply to various items. For instance, there are probably many factors relevant to determining whether "is smart" accurately applies to a person. For the sake of simplicity, suppose that the degree to which the person is

---

[106] This is suggested in Cooper 1995, 246.

[107] Keefe 2000, 12-15.

observant, analytical, and imaginative are all that matter. If someone is very observant, analytical, and imaginative, then she is definitely smart. Someone who entirely lacks these three features is definitely not smart. What about someone who is fairly observant, more analytical than average, but not terribly imaginative? It might be that we cannot determinately say whether this person is smart or not. The multidimensionality of "is smart" is at least in part what makes it indeterminate. But notice that this indeterminacy is not lost by shifting to the comparative, "is smarter than." It may be true that a smart person is smarter than a non-smart person, but what can we say about two smart people, A and B, who score dramatically differently in terms of the three relevant factors listed above? There might be no way to precisely combine these factors, and so it might be indeterminate whether person A is smarter than person B. This is one way in which comparatives could be vague.

We might think that "better than" is also vague in virtue of its multidimensionality. The degree to which an outcome "is good" might depend on the degree to which different dimensions of goodness are realized in it. For instance, we might think that both *achievement* and *pleasure* are important dimensions of goodness. Perhaps a lot of achievement is better than a little bit of pleasure, and a lot of pleasure is better than a little bit of achievement. But it might be indeterminate how much achievement is as good as some fixed amount of pleasure. Even if we accept a monistic theory of the good, "is good" and "better than" can be vague due to multidimensionality. For instance, hedonists believe that only pleasure is good, but they also believe that both the intensity and the duration of pleasant experiences

matter. A hedonist might think that there is indeterminacy about how intensity and duration tradeoff against one another. If "is good" is vague due to its multidimensionality, then it is fairly clear that the corresponding comparative, "better than," could also be vague. It might be indeterminate whether a three hour long pleasant experience is better than a somewhat more intensely pleasant experience that lasts two and a half hours.

5.4.2   Indeterminacy versus Imprecision[108]

If we claim that it is indeterminate that A is better than B, we are claiming that it is indeterminate, or that *we cannot say*, whether:

(i)   A is better than B, or

(ii)   B is better than A, or

(iii)   A and B are precisely equally good.

We might instead claim that it is determinate that (i), (ii), and (iii) are false. If we do, we might still claim that:

---

[108] Much of this subsection, and 5.5, is owed to Parfit's "Towards Theory X" (unpublished manuscript). I have been in written and personal correspondence with Parfit about both.

(iv) A and B are *incommensurable*, in the sense that they are *imprecisely* equally good.

Since claiming that (i), (ii), and (iii) are false is different from claiming it is indeterminate whether they are true or false, it seems that there is a difference between claiming that it is indeterminate whether A is better than B, on the one hand, and claiming that A and B are imprecisely equally good, on the other. There is some debate about whether such incommensurability collapses into vagueness or indeterminacy, but I believe (and will here simply assume) that there is conceptual room for both.[109]

We still need to distinguish (iii) and (iv) somewhat more carefully. If A and B are precisely equally good, then each is at least as good as the other. But if A and B are imprecisely equally good, then neither is at least as good as the other. "Imprecisely equally as good as" is not transitive, and moreover if A and B are imprecisely equally good, then there can be some third item, C, which is better than A but is neither better nor worse than B. If A and B are precisely equally good, however, there could be no such third item, C (by stipulation C is neither better nor worse than B, but if C were better than A then C would also be better than B, since A is at least good as B; so we have a contradiction).

Multidimensionality can give rise both to indeterminacy and to imprecision. Consider pleasure again. (I focus on pleasure for the simple reason that, whatever else

---

[109] On incommensurability, see Griffin 1986, Raz 1986, and Chang 1997. On the question of whether incommensurability collapses into vagueness or indeterminacy, see, for example, Broome 2004, and Parfit's reply in "Towards Theory X" (unpublished manuscript).

has intrinsic value, at least pleasure does).[110] There are at least two dimensions

relevant to the goodness of pleasure: intensity and duration. When two pleasures are

the same along one dimension, and differ only along the other dimension, there will be

no indeterminacy or imprecision about their relative goodness. For example, it is

determinately true that, of two equally intense pleasures, the longer one is better. And

it is better *precisely* by however much longer it is. If these two pleasures are exactly

equally intense *and* last exactly equally long, then it is determinately true they are

precisely equally good.

    However, suppose that pleasure A is more intense than pleasure B, but lasts

only half as long as B. Now there are at least five possibilities about the relative

goodness of A and B:

    (1) A is better than B.

    (2) It is indeterminate whether A is better than B, but it is determinately true

        that B is not better than A.

    (3) A and B are imprecisely equally good.

    (4) It is indeterminate whether B is better than A, but it is determinately true

        that A is not better than B.

    (5) B is better than A.

---

[110] As Stuart Rachels says, "In general, if you can make your point with pleasure, then do so" (Rachels 2004). Similar claims are true about pain and intrinsic badness.

Could A and B be *precisely* equally good? It seems dubious. At least, insofar as the difference in intensity and duration between A and B is significant enough, many of us will find it hard to believe that there could be no third pleasure, C, which is slightly better than A in virtue of being the same intensity as A but only slightly longer, and yet neither better nor worse than B. That is to say, we will find it hard to believe that, when the difference in intensity and duration between A and B is significant enough, A and B could be precisely equally good. This seems to be a general truth about the relative goodness of items that differ along multiple relevant dimensions: the more different they are along these relevant dimensions, the less precisely they compare in terms of goodness.[111]

5.5 Parfit's Solution to Spectrum Arguments

Making use of the notions of indeterminacy and imprecision, Parfit has articulated an interesting and resourceful reply to a variety of Spectrum Arguments.[112] There is one other notion which Parfit appeals to, which has already surfaced at a number of points throughout this dissertation but has not yet been explicitly defined.

---

[111] One argument against the view that items can be imprecisely equally good is known as the *Money Pump Argument*. Suppose that B is imprecisely equally as good as both A and C, but that C is better than A. According to the Money Pump Argument, it is a sign that a goodness ranking is false if a series of rational exchanges based on it would result in being worse off. Thus, suppose you have C. It might be rational to exchange C for B, since they imprecisely equally good. It might then be rational to exchange B for A, since they are imprecisely equally good. But now you are worse off than you when you started, since C is better than A. I will discuss Money Pump Arguments in chapter 6.

[112] Again, he does so in his unpublished manuscript "Towards Theory X."

This is the notion of *Lexically Related Values*. There are actually two relations which need to be defined. First:

> P *is lexically better than* Q if and only if: P and Q are both good, the existence of some number of Ps, or some amount of P, would be better than the existence of any number of Qs, however large, or any amount of Q, however large. Moreover, this is true even if the goodness of Q is non-diminishing (even if it is not the case that amount of goodness contributed by more Q diminishes the more Q there is).

And second:

> P *is lexically worse than* Q if and only if: P and Q are both bad, the existence of some number of Ps, or some amount of P, would be worse than the existence of any number of Qs, however large, or any amount of Q, however large. Moreover, this is true even if the badness of Q is non-diminishing (even if it is not the case that amount of badness contributed by more Q diminishes the more Q there is).

To illustrate these notions further, and Parfit's reply to Spectrum Arguments, I will focus on Hangnails for Torture (the Spectrum Argument for this conclusion is one

of the main arguments Parfit seeks to reply to, the other main argument he seeks to

reply to is the Spectrum Argument for the Repugnant Conclusion). Recall:

> *Hangnails for Torture*. For any excruciatingly painful torture session lasting
>
> for two years, there is a longer session consisting of very mildly annoying
>
> hangnail pain which is, other things equal, worse.

This claim seems implausible. For ease of discussion, let us name different

pains according their intensity such that:

> A-Pain is as intense *excruciatingly* painful torture.
>
> B-Pain is slightly less intense than A-Pain.
>
> C-Pain is slightly less intense than B-Pain.
>
> …and so on…
>
> Z-Pain is as intense as a *very mildly* annoying hangnail pain.

Recall the following finite Series of possible outcomes:

> (1): a two year long A-Pain.
>
> (2): a much longer B-Pain.
>
> (3): a much longer C-Pain.

And so on… all the way up to:

(*n*): an arbitrarily long Z-Pain.

And again consider the following Premises:

P1. Outcome (2) is worse than outcome (1).

P2. Outcome (3) is worse than outcome (2).

And so on… all the way up to:

P*n*. Outcome (*n*) is worse than outcome (*n*-1).

The Premises, together with Transitivity (of "is worse than"), imply that (*n*) is worse than (1). That is, they imply Hangnails for Torture. Parfit accepts Transitivity, but he is inclined to believe that there is no amount of Z-Pain that could be worse than two years of A-Pain. That is, he is inclined to believe that A-Pain is lexically worse than Z-Pain, and he is thus inclined to believe that Hangnails for Torture is false. He thus will have to deny at least one of the Premises.

5.5.1   The Vague Way Out

Parfit's solution to Spectrum Arguments (for Hangnails for Torture), in a nutshell, is to divide the Premises up into five zones, making the following claims of the Premises in each zone:[113]

Zone One (from A-Pain to K-Pain): for any pain at an intensity level in this zone, there is some longer period of slightly less intense pain that would be worse.

Zone Two (from L-Pain P-Pain): for any pain at an intensity level in this zone, it is indeterminate whether there is some longer period of slightly less intense pain that would be worse or whether any such period would at most be imprecisely equally as bad.

Zone Three (from Q-Pain to T-Pain): for any pain at an intensity level in this zone, any longer period of slightly less intense pain would at most be imprecisely equally as bad.

Zone Four (U-Pain to W-Pain): for any pain at an intensity level in this zone, it is indeterminate whether any longer period of slightly less intense pain would be at most imprecisely equally as bad or would be less bad.

---

[113] "Towards Theory X," Part One, 23.

Zone Five (X-Pain to Z-Pain): for any pain at an intensity level in this zone, any longer period of slightly less intense pain would be less bad.

There are at least three significant advantages of Parfit's solution.

The first and most obvious advantage is that since it denies at least one of the Premises, it avoids Hangnails for Torture.

Secondly, it seems less implausible to say, in transitioning from Zone One to Zone Two, that (i) we shift from claiming the next item in the Series *is* worse to claiming that *it is indeterminate whether* the next item in the Series is worse, rather than that (ii) we shift from claiming that the next item in the Series *is* worse to claiming that the next item in the Series is *equally bad* or *less bad*. It seems that Parfit's solution makes this transition gentler and so less implausible.

Thirdly, Parfit's solution has the resources to rebut the following argument against the claim that A-Pain is lexically worse than Z-Pain. I will call it *the Slight Differences Argument*:

(1) If A-Pain is lexically worse than pain below certain intensity levels, there must be a *highest* such intensity level. Suppose this is Q-Pain.

(2) Since Q-Pain is the highest such intensity level, A-Pain is *not* lexically worse than P-Pain (which is slightly more intense than Q-Pain). But,

(3) Since P-Pain is only slightly more intense than Q-Pain, it seems implausible that A-Pain is not lexically worse than P-Pain whereas A-Pain is lexically

worse than Q-Pain. Surely such a slight difference in intensity between P-Pain and Q-Pain cannot result in such a large difference in how P-Pain and Q-Pain compare, respectively, to A-Pain.

So, (4) it is implausible that A-Pain is lexically worse than any pain at any lower intensity level.

We might have found this argument tempting if we thought we had to claim either that A-Pain is lexically worse than P-Pain or else that A-Pain is not lexically worse than P-Pain. But Parfit's solution provides us with a third option: we can instead claim that it is *indeterminate* whether A-Pain is lexically worse than P-Pain. Similarly, we might have found Sorites Arguments more tempting if we thought that we had to claim the first (largest) collection of grains of sand that is not determinately a heap is determinately *not* a heap. We have a third option here: we can instead claim that it is indeterminate whether it is a heap.

In other words, (2) in the Slight Differences Argument is false. It does not follow from the claim that Q-Pain is the highest intensity level such that A-Pain is lexically worse than it that A-Pain is *not* lexically worse than P-Pain. Again, instead, it could be that it is indeterminate whether A-Pain is lexically worse than P-Pain. We might run the argument again, with these revised versions of (2) and (3):

(1)   If A-Pain is lexically worse than pain below certain intensity levels, there must be a *highest* such intensity level. Suppose this is Q-Pain.

(2*)   Since Q-Pain is the highest such intensity level, either it is indeterminate

whether A-Pain is lexically worse than P-Pain or A-Pain is not lexically worse

than P-Pain. But:

(3*)   Since P-Pain is only slightly more intense than Q-Pain, it seems implausible

that it is *indeterminate whether* A-Pain is lexically worse than P-Pain (or that

A-Pain is not lexically worse than P-Pain) whereas A-Pain *is* lexically worse

than Q-Pain. Surely such a slight difference in intensity between P-Pain and

Q-Pain cannot result in such a large difference in how P-Pain and Q-Pain

compare, respectively, to A-Pain.

So, (4) it is implausible that A-Pain is lexically worse than any pain at any lower

intensity level.

(3*) is less compelling than (3), but those wielding the Slight Differences

Argument might believe that it is nonetheless sufficiently compelling. I am not sure it

is. (3*) seems analogous to the claim that it is implausible that the slight difference of

one grain of sand can take us from a collection of grains of sand that is determinately a

heap to a collection of grains of sand such that it is indeterminate whether it is a heap.

And it is not clear how plausible this latter claim is.

At the very least, it seems we can here usefully invoke higher-order

indeterminacy. That is, even if we find it implausible that, starting from a heap and

removing one grain of sand at a time, we would go from calling one collection of

grains of sand a heap to saying that it is indeterminate whether its successor is a heap,

it seems harder to deny that we would go from calling one collection of grains of sand a heap to saying that it is *indeterminate whether* it is indeterminate whether its successor is a heap. Or, we might say that it is indeterminate whether it is indeterminate whether it is indeterminate whether, and so on…, it is indeterminate whether its successor is a heap. We can say that it is *super*indeterminate whether its successor is a heap if there is *any* order of indeterminacy about whether its successor is a heap.[114] And it *does* seem plausible that the slight difference of one grain of sand *can* take us from a collection of grains of sand that is determinately a heap to a collection of grains of sand such that it is superindeterminate whether it is a heap. (We might not all agree on *which* collection this is, but we should agree that such a collection exists).

If we are to accommodate such higher-order indeterminacy, we will have to revise the Slight Differences Argument once more:

(1)   If A-Pain is lexically worse than pain below certain intensity levels, there must be a *highest* such intensity level. Suppose this is Q-Pain.

(2**)   Since Q-Pain is the highest such intensity level, either it is superindeterminate whether A-Pain is lexically worse than P-Pain or A-Pain is not lexically worse than P-Pain. But:

(3**)   Since P-Pain is only slightly more intense than Q-Pain, it seems implausible that it is *superindeterminate whether* A-Pain is lexically worse than P-Pain (or

---

[114] Norcross appeals to the notion of superindeterminacy in his reply to Quinn in Norcross 1997, 144. (As Norcross notes, it is self-defeating to postulate indeterminacy about superindeterminacy).

that A-Pain is not lexically worse than P-Pain) whereas A-Pain *is* lexically

worse than Q-Pain. Surely such a slight difference in intensity between P-Pain

and Q-Pain cannot result in such a   large difference in how P-Pain and Q-

Pain compare, respectively, to A-Pain.

So, (4) it is implausible that A-Pain is lexically worse than any pain at any lower

intensity level.

However, not only is (3**) less compelling than (3*) and (3), but it does not

seem compelling at all. We can thus reject the Slight Differences Argument. This, as

I said, is the third main advantage of Parfit's solution.

Lastly, note that we can similarly appeal to higher-order indeterminacy to

describe the transitions between the five Zones mentioned above. In that this would

help make these transitions gentler, it would help strengthen the second advantage of

Parfit's solution.

5.5.2   Why Parfit's Solution Fails

There are different kinds of solutions we can offer to a Spectrum Argument. We

can:

1. Accept the conclusion.

2. Deny some but not all of the Premises (deny Parity or the Conditional).

3. Deny all of the Premises.

4. Deny Transitivity.

5. Claim that the argument is unsound in the way Sorites Arguments are
   unsound (where this is distinct from 1 through 4).

Parfit offers the second kind of solution. I believe that this kind of solution is better than the fourth and fifth kinds of solution, and I believe that Parfit's solution is possibly the best version of the second kind of solution.

Nonetheless, I believe that Parfit's solution is implausible, and I doubt that it could be the least implausible solution (more on the latter in chapter 7). This is because I fail to see how some of the Premises could be plausible, while others implausible. The Premises near Pn seem, at least on the most charitable ways of setting up the Series, no less plausible than those Premises near P1.

Parfit's claim that the earlier Premises are in Zone One, later Premises in Zone Two, and so on…, and Premises near Pn are in Zone Five, *might* be based on an illegitimate sort of *anchoring* to the beginning of the Spectrum, or Series. For example, *if* we assessed the Premises by asking the following corresponding questions:

P1.   Is there an amount of B-Pain that is worse than two years of A-Pain?

P2.   Is there an amount of C-Pain that is worse than two years of A-Pain?

P3.   Is there an amount of D-Pain that is worse than two years of A-Pain?

        …and so on…

P(*n*-1).   Is there an amount of Y-Pain that is worse than two years of A-Pain?

P*n*.   Is there an amount of Z-Pain that is worse than two years of A-Pain?

then I agree that our pattern of intuitive responses would reflect Parfit's five Zones.

That is, it would seem that the answers to each of the earlier questions near P1 is Yes,

then further down the list of questions it would seem that it is indeterminate whether

Yes (but determinately not No), then further down it would seem that it is

determinately not Yes and determinately not No (and so there is imprecise equality),

then further down it would seem that it is indeterminate whether No (but determinately

not Yes), and finally, it would seem that the answers to each of the questions near P*n*

is No.

        But I see no reason why we *should* assess the Premises by asking the

questions, which are anchored around A-Pain. It would seem no less legitimate to

assess the Premises by asking the following questions which are instead anchored

around Z-Pain:

P*n*.   Is there an amount of Z-Pain that is worse than any amount of Y-Pain?

P(*n*-1).   Is there an amount of Z-Pain that is worse than any amount of X-Pain?

        …and so on…

P3.   Is there an amount of Z-Pain that is worse than any amount of C-Pain?

P2.   Is there an amount of Z-Pain that is worse than any amount of B-Pain?

P1.   Is there an amount of Z-Pain that is worse than any amount (two years)

           of A-Pain?

Again, our pattern of intuitive responses would reflect Parfit's five Zones.

That is, it would seem that the answers to each of the earlier questions near P*n* is Yes,

then further down the list of questions it would seem that it is indeterminate whether

Yes (but determinately not No), then further down it would seem that it is

determinately not Yes and determinately not No (and so there is imprecise equality),

then further down it would seem that it is indeterminate whether No (but determinately

not Yes), and finally, it would seem that the answers to each of the questions near P1

is No.

Though our intuitive responses to the questions anchored around A-Pain and to

the questions anchored around Z-Pain are consistent with each other, they provide

inconsistent ways of assessing the Premises. Neither way, it seems, provides a

plausible way of assessing the Premises.

Perhaps this should be obvious. The legitimate and plausible way to assess the

Premises is, well, by looking at *the Premises*, and *not* those other two lists of questions

anchored around A-Pain or Z-Pain, respectively. And the Premises are not anchored

around any particular pain intensity level. The questions we should ask, in assessing

the Premises, are:

P1.   Is there an amount of B-Pain that is worse than two years of A-Pain?

P2.   Is there an amount of C-Pain that is worse than any amount of B-Pain?

P3.   Is there an amount of D-Pain that is worse than any amount of C-Pain?

     …and so on…

P(n-1).   Is there an amount of Y-Pain that is worse than any amount of X-Pain?

Pn.   Is there an amount of Z-Pain that is worse than any amount of Y-Pain?

And *now*, I believe, our pattern of intuitive responses would *not* reflect Parfit's five Zones. Instead, it would seem that answers are Yes, Yes, Yes, and so on…, Yes, and Yes. Our confidence that the answer to each question is Yes would *not* diminish as we move from P1 to Pn, and at no point would it even become superindeterminate whether the answer is Yes.

Of course, we have to consider the relevant and appropriate questions, in assessing P1 through Pn. If we take the difference in intensity between adjacent pain levels to be *significant*, rather than *slight*, then it may not be plausible that the answers to each of the above questions is Yes. And if we compare only *somewhat* longer less intense pain with pain at the previous step, then again it may not be plausible that the answers to each of the above questions is Yes. In either case, it may then be plausible that it is indeterminate what the answer are, or that the answers are No.

But, in assessing P1 through Pn, we should be comparing pains with *slightly* less intense pains which last *arbitrarily* longer. We can, I believe, adjust the relevant intensity and duration differences such that we are taken clearly outside any zone of

indeterminacy; that is, such that it is determinately true that each slightly less intense

and arbitrarily longer pain is worse. Therefore, Parfit's solution to the Spectrum

Argument is implausible.[115]

And while it is true that there are important differences between the Spectrum

Argument for Hangnails for Torture, on the one hand, and Spectrum Arguments for

Repugnant Conclusion, Priority Monster, and so on, on the other hand, I believe that

the reason Parfit's indeterminacy-based solution fails in the case of the Hangnails for

Torture Spectrum Argument is also the reason that this kind of solution would fail in

these other cases.[116]

---

[115] At this juncture, it may be helpful to remember another way in which we might illegitimately anchor to some part of the Spectrum, and thereby fail to appreciate the (roughly equal) plausibility of each of the Premises. (I noted this in 3.3.2, in defense of Parity).

The fact that if we reject at least one of the Premises we can avoid Hangnails for Torture does give us reason to reject at least one of the Premises, but it does not give us any reason to reject some Premises *over* others. Having observed this, we can now see how might have been tricked into thinking that Premises near Pn are less plausible than those near P1. If we went through intuitively judging the Premises *starting* and P1 and moving toward Pn, we might get progressively less and less confident about accepting individual Premises given our belief that if we accept P1 through Pn Transitivity will commit us to Hangnails for Torture. This might bias us against Premises located closer to Pn. We should check whether the bias works both ways. If, when we went through intuitively judging the Premises starting at Pn and moving toward P1, we became progressively less and less confident about accepting individual Premises given our belief that if we accept Pn through P1 Transitivity will commit us to Hangnails for Torture, then it might not be the location of the Premises as such that is driving our intuitions. We should guard against being tricked in this way, and thereby avoid believing that Premises near Pn are less plausible than Premises near P1 for erroneous reasons. (And I would wager a lot of money that if we considered each of the Premises in complete isolation of the others, perhaps in a random order, where each time we considered a new Premise we forgot about any of the others, we would be strongly inclined to accept each).

[116] Qizilbash 2005 offers a solution to Spectrum Arguments which is similar to Parfit's. He considers a spectrum of illnesses, ordered by their seriousness (this illness Spectrum Argument came from Temkin). At one end of the spectrum is AIDS (a serious illness), and at another end is the common cold (a non-serious illness). It is intuitively plausible that, other things equal, it is better to cure one person of AIDS than to cure any number of people from experiencing the common cold. But we can construct a Spectrum Argument against this intuitively plausible claim:

P1. It is better that one person have AIDS than that ten people have a slightly less serious illness, B.

5.5.3   Why this Isn't a Sorites, Yet Again

Instead of offering a Spectrum Argument for Hangnails for Torture, we could have offered the Slight Differences Argument for Hangnails for Torture. Here again, is the latter (revised in order to leave room for indeterminacy):

(1) If A-Pain is lexically worse than pain below certain intensity levels, there must be a *highest* such intensity level. Suppose this is Q-Pain.

P2.   It is better that ten people have B than that one hundred people have a slightly less serious illness, C.
       …and so on…
P*n*.   It is better that one hundred million people have some non-serious illness X than it is for one billion people to have a slightly less serious illness, Y, e.g., the common cold.
P*T*.   Transitivity.
So,   It is better that one person have AIDS than that one billion people have the common cold.

Qizilbash accepts Transitivity, but denies the conclusion of the Spectrum Argument. He appeals to the belief that it is better to cure one person of a serious illness than to cure any number of people from experiencing a non-serious illness. He rightly adds that there is indeterminacy about what constitutes a *serious* illness. Since there is no sharp line dividing serious and non-serious illnesses, there will be a range of steps in the Spectrum Argument wherein it will be indeterminate whether non-serious illnesses are being weighed against serious illnesses. Hence, there will be premises in the Spectrum Argument that we cannot say are true.
       However, I find Qizilbash's solution to be implausible. It may well be indeterminate whether, at some step in the sequence, there is a tradeoff or exchange of a serious illness for a non-serious illness. But this should not prevent us from forming a judgment about whether, at this step, the tradeoff *is* justified. It might be indeterminate whether illness S is a serious or non-serious illness, and whether illness T is a serious or non-serious illness (or it might even be *determinate* that illness S is a serious illness and superindeterminate whether illness T is a serious illness) – but if we *do* know that T is only *slightly less* serious than S, it seems hard to deny that it is better for 100 people to suffer from S than it is for 1,000 people to suffer from T. It is even harder to deny that it is better for 100 people to suffer from S than it is for *one billion* people to suffer from T. It would be implausible to say that *we cannot say* whether this is true, for it clearly *is* true. Each premise thus seems very hard to deny despite the fact that there is indeterminacy about what counts as a serious illness.

(2) Since Q-Pain is the highest such intensity level, it is either superindeterminate whether A-Pain is lexically worse than P-Pain or A-Pain is not lexically worse than P-Pain. But:

(3) Since P-Pain is only slightly more intense than Q-Pain, it seems implausible that it is either superindeterminate whether A-Pain is lexically worse than P-Pain or A-Pain is not lexically worse than P-Pain whereas A-Pain *is* lexically worse than Q-Pain. Surely such a slight difference in intensity between P-Pain and Q-Pain cannot result in such a large difference in how P-Pain and Q-Pain compare, respectively, to A-Pain.

So, (4) it is implausible that A-Pain is lexically worse than any pain at any lower intensity level.

This Slight Differences Argument *is*, I believe, relevantly like a Sorites Argument. That is because (3) is relevantly like a *Soritical Premise*: if A-Pain is lexically worse than pain at some intensity level, then A-Pain is lexically worse than pain at a slightly higher intensity level. But insofar as A-Pain is *not* lexically worse than B-Pain, we can take our Soritical Premise and apply modus tollens, running down the spectrum as follows: since A-Pain is not lexically worse than B-Pain, A-Pain is not lexically worse than C-Pain. And since A-Pain is not lexically worse than C-Pain, A-Pain is not lexically worse than D-Pain, and so on…, to the conclusion that A-Pain is not lexically worse than Z-Pain. *This* is a Sorites. Insofar as we can plausibly

dismiss Sorites Arguments as unsound, so too can we plausibly dismiss this Slight Differences Argument as unsound.

It is true that we *could* offer either the Slight Differences Argument, or a Spectrum Argument, or both, as arguments for Hangnails for Torture. Though they have the same conclusion, these arguments are importantly different. The former, in appealing to a premise relevantly like a Soritical Premise, is relevantly like a Sorites Argument. But the Spectrum Argument for Hangnails for Torture appeals to no such Soritical Premise. It simply appeals to the plausibility of the Premises, and Transitivity. This, I hope, helps highlight yet again how Spectrum Arguments are relevantly disanalogous to Sorites Arguments.

Let me recapitulate some of the most important claims I have argued for in this chapter.

Unlike standard Sorites Arguments (5.1), Spectrum Arguments do not essentially rely on Soritical Premises (5.3.1). And it is this feature, rather than Slight Differences or Slight Differences*, which seems to make Sorites Arguments unsound (5.2). Comparative Sorites Arguments, for example involving "is no less a heap than" rather than "is a heap," do not essentially rely on Soritical Premises. However, they have premises which, if they are not obviously false, might raise the same sort of puzzles that Spectrum Arguments do (5.3.3). Lastly, Parfit's solution to Spectrum

Arguments (5.5.1) based on indeterminacy and imprecision (5.4) is implausible (5.5.2).

Chapter 6

Transitivity

CHAPTER OUTLINE

6.1   The Analytic Truth Argument

One of the most important assumptions behind Spectrum Arguments is Transitivity. According to Transitivity, if A is better than B, and B is better than C, then it follows that A is better than C.

It is important to recall that Transitivity is concerned with *all things considered* betterness. For example, Transitivity does not rule out the following: A is better than B with respect to utility, B is better than C with respect to perfection, and C is better than A with respect to justice. This conjunction of betterness claims does not violate

Transitivity, since the relevant sort of betterness is different in each claim. But "is better than with respect to X" does seem to be transitive (at least, this seems as plausible as Transitivity). For example, if A is better than B with respect to justice, and B is better than C with respect to justice, then it seems to follow that A is better than C with respect to justice. It is also worth noting that while both the "is all things considered better than" and the "is better than with respect to X" relations seem to be transitive, the "regularly beats" relation is not transitive. Consider three chess grandmasters, Bobby, Boris, and Garry. Bobby regularly beats Boris, Boris regularly beats Garry, and Garry regularly beats Bobby. This does not imply that Bobby is a better chess player than Boris, that Boris is a better chess player than Garry, and that Garry is a better chess player than Bobby. What determines how good a chess player one is (or how good one is with respect to chess) is, perhaps among other things, how one compares against all sorts of possible opponents, which is approximated by one's chess rating. Garry's rating is 2850, Bobby's is 2780, and Boris's is 2700. Thus the best player, Garry, regularly loses to the worst of the three, Boris.[117] There is neither a violation of Transitivity nor of the transitivity of "is better than with respect to X" here.

Virtually everyone is strongly inclined to accept Transitivity, once they grasp its content. Many people find it so hard to deny that they will refer to putative counterexamples to it as *contradictions*, or regard such examples as unintelligible.

---

[117] Similarly, in the game of Rock-Paper-Scissors, rock beats scissors, scissors beat paper, and paper beats rock, but each is equally good with respect to the game of Rock-Paper-Scissors (otherwise the game would have little point!). Broome distinguishes "can regularly beat" from "is better than" using the example of football teams (2004, 52).

Nonetheless, there are excellent philosophers who deny Transitivity. Most notably and importantly, Stuart Rachels and Larry Temkin have offered aggregation Spectrum Arguments (particularly, the Hangnails for Torture Spectrum Argument) as support for the claim that Transitivity should be rejected. They believe that the Premises of such arguments are so plausible, and their conclusions so implausible, that the best way to maintain consistency is to deny Transitivity. Rather than regard the weighting Spectrum Arguments I offer in this dissertation as a kind of indirect support for the Equal Weight View, one might likewise claim that my arguments constitute further evidence against Transitivity.

I believe that this would be a mistake. Transitivity seems too hard to deny. I will now consider five arguments for Transitivity, some of which I think might successfully explain why it is so hard to deny.

6.1.1 The Analytic Truth Argument

Some philosophers, including John Broome, believe that Transitivity is an analytic truth. Whether "better than" is transitive is not, according to Broome, an issue in ethics, but "an issue in semantics."[118] That "A is better than C" is just part of what it *means* to say that "A is better than B and B is better than C." Broome elaborates: "A comparative relation is necessarily transitive. This is an analytic

---

[118] Ibid., 51.

feature of the operator 'more … than': the meaning of 'more … than' implies that 'more *F* than' is transitive."[119] Since "better than" means "more *good* than," it is transitive.

This *Analytic Truth Argument* for Transitivity has two premises: (1) that "better than" means "more good than," and (2) that comparative relations are necessarily transitive. It is not clear how compelling either premise is.

Whether "more *F* than" is transitive might depend on what *F* is. Temkin thus considers the following non-normative putative counterexample to (2).[120] Suppose that *F* is "large," and that A is larger than B if and only if *either* A is taller than B *or* A is heavier than B. This relation is not transitive. It could be that A is larger than B (because taller), B is larger than C (because heavier), and that C is larger than A (because taller). It is true that this notion of "larger than" leaves open the possibility that A is larger than B and that B is larger than A, but that is not by itself an objection to its coherence. There are many relations with this feature, e.g., A loves B and B loves A. Or one might object that the above notion of "large" is just another one of those artificial disjunctive predicates contrived by clever philosophers (like Nelson Goodman's color "grue"[121]), and that it can thus be dismissed as not the right kind of *F*. However plausible such a dismissal might be, it *concedes* the point that whether (2) is true depends on what *F* is.

---

[119] Ibid., 50.

[120] Temkin 2012, chapter 7.

[121] Goodman 1983.

There is a class of predicates F such that "more F than" is necessarily transitive. I do not know the exact boundaries of this class. However, I do know at least one subclass of it. If F is measurable along a single linear scale,[122] such that "more F than" is represented by "to the right of" on a straight Euclidean line, then "more F than" is necessarily transitive. This is because "to the right of" on a straight Euclidean line is necessarily transitive. Indeed, many people who accept Transitivity accept what I will call the,

> *Orthodox View*: "is better than" means "is more good than," and "good" is
> measurable along a single linear scale.

However, we might reject this view, and accept either the:

> *First Unorthodox View*: "is better than" does not mean "is more good than,"

Or the:

> *Second Unorthodox View*: "is better than" means "is more good than," but
> "good" is not measurable along a single linear scale. There are different scales

---

[122] It need not be a ratio or even interval scale. An ordinal scale would suffice. Griffin helpfully discusses the differences between these types of scale in the context of the measurement of well-being in his book (1986, 93-105).

of goodness, the appropriateness of which is sometimes dependent on the items being compared,[123] or there are no such scales.

If the Orthodox View is true, then Transitivity is true (for the reason mentioned above). But one could accept Transitivity without accepting the Orthodox View. For example, if we believe that there can be imprecise equality between items A and B, such that A is not better than B, B is not better than A, and there could be a third item C which is better than A but not better than B, then we reject the Orthodox View. Such a relation of imprecise equality cannot be represented on a single linear scale, and this relation is not transitive (see 5.4.2). But we could believe in such imprecise equality, and thereby reject the Orthodox View, and yet still accept Transitivity. Transitivity, after all, claims that a different relation, "is better than," is transitive. While both Parfit and Broome accept Transitivity, I believe that Broome, but not Parfit, accepts the Orthodox View.

Phrased differently, we can accept either of the two Unorthodox Views without denying Transitivity. But the *denial* of Transitivity is consistent with Unorthodox Views, not with the Orthodox View.

As it stands, the Analytic Truth Argument for Transitivity is not convincing. This is because neither (1) nor (2) is convincing. If the Orthodox View were true, then

---

[123] For example, if we deny Hangnails for Torture, we might think that very intense pains belong on a *Higher Scale*, that mildly intense pains belong on a *Lower Scale*, and that no amount of Lower Scale badness could be worse than some amount of Higher Scale badness. We might similarly adopt a multiple scales model if we accept Mill's Higher Pleasures Doctrine (1861, chapter 2), or if we accept other such Lexical Views. This multiple scales model is consistent both with Transitivity and with its denial.

(1) (that "better than" means "more good than") would be true, and "good" would be

the sort of *F* such that "more *F* than" is necessarily transitive.

6.1.2   An Issue in Semantics?

It might still be "an issue in semantics" whether Transitivity is true, in the

sense that it is merely a matter of terminology choice whether we use "better than"

such that it means "more good than," and such that "good" is the sort of *F* such that

"more *F* than" is necessarily transitive. We could just say that *this* (the Orthodox

View) captures what we *shall mean* by "better than," and so "better than" in this sense

is Transitive. But this "semantic defense" of Transitivity is either (1) plausible but

irrelevant or (2) relevant but implausible.

For a sense of "better than" to be relevant, it must be that on this sense of

"better than" the following claim is indubitable:

> *Reason-Giving Thesis*: if A is all things considered better than B, and if either
>
> could be brought about, then there is all things considered more reason to bring
>
> about A than there is to bring about B, other things being equal.[124]

---

[124] The point of this "other things equal" clause is to bracket reason-giving factors other than the goodness of the outcomes the agent could bring about. For example, it might be that one option is slightly worse than another, but the agent has a stronger *desire* to bring about this worse option. It might then be that the agent has more reason to choose the worse option. I shall simply set reason-giving factors other than goodness aside, and take this "other things equal" clause as read from here onward.

If there is a sense of "all things considered better than" for which the Reason-Giving Thesis is false, then it is not a sense of "all things considered better than" with which I am here concerned. So, insofar as the above semantic defense of Transitivity does not make use of a sense of "better than" for which the Reason-Giving Thesis is true, it is irrelevant. In the next two paragraphs I will argue that if, on the other hand, such a defense of Transitivity does make use of a sense of "better than" for which the Reason-Giving Thesis is true, then it is implausible.

Recall the Spectrum Argument for Hangnails for Torture. The Premises of this argument are highly plausible. Let us here assume they are true. If Transitivity is true, then so is Hangnails for Torture – the conclusion that there is some amount of very mildly annoying hangnail pain that is worse than two years of excruciating torture. But if Transitivity is false, then it might be perfectly consistent to deny Hangnails for Torture. Now suppose that we are in the unfortunate situation of having just two available acts: bring about two years of excruciating torture, or bring about an arbitrarily large long period of very mildly annoying hangnail pain (and other things are equal). If Hangnails for Torture is true, then according to the Reason-Giving Thesis we have more reason to bring about the former. But if Hangnails for Torture is false, then according to the Reason-Giving Thesis we have more reason to bring about the latter.

We can now see why, assuming the Reason-Giving Thesis, it is implausible to regard it as a mere matter of terminological stipulation whether we use a sense of

"better than" such that Transitivity is true. Whether we have more reason to bring about the torture pain or the hangnail pain depends on whether Hangnails for Torture is true, according to the Reason-Giving Thesis. But whether Hangnails for Torture is true depends on whether Transitivity is true. And whether Transitivity is true depends on which conception of "better than" we are working with. Thus, what we have more reason to do sometimes depends on which conception of "better than" we are working with. But it clearly *cannot* be that what we have more reason to do could be determined by mere terminological stipulation, or which conception of "better than" we are working with. Therefore, it is implausible to regard it as merely a matter of terminological stipulation whether we use a sense of "better than" such that Transitivity is true, assuming the Reason-Giving Thesis. That is, assuming that our sense of "better than" is not irrelevant.

Transitivity, then, is not merely an issue in semantics, and is primarily a substantive issue in normative theory. To determine which conception of "better than" is most plausible, and whether it is transitive, we have to carefully consider the theoretical costs and benefits of these rival conceptions. And it is, to my mind, a significant theoretical benefit of Unorthodox Views like Rachels's and Temkin's that they enable us to consistently accept the Premises and deny Hangnails for Torture. (They also enable us to consistently accept the Premises of a variety of other Spectrum Arguments whilst denying their implausible conclusions). The positive case for Transitivity will have to be fairly strong if it is to outweigh this benefit. Let us now turn to further arguments for Transitivity.

6.2   The Money Pump Argument

In what follows, I will use ">" as the symbol for "better than" (and "=" as the symbol for "as good as"). That is: "A > B" means "A is better than B."

The *Money Pump Argument* is a famous argument for Transitivity.[125] On its standard formulation, it specifically works as an argument against intransitive betterness *cycles* of the form A > B > C > D > A.[126] It could easily be revised so as to target acyclical intransitive rankings of the form A > B > C > D = A, too.[127] I will distinguish between two broad versions of the Money Pump Argument: a *theoretical* version, which argues that the denier of Transitivity cannot plausibly avoid the implausible claim that it is rational to get money-pumped, and a *practical* version, which argues that because the denier of Transitivity sometimes cannot plausibly avoid being money-pumped, the denial of Transitivity is implausible.

6.2.1   The Theoretical Money Pump Argument

---

[125] It comes from Davidson, McKinsey, and Suppes 1955, 145-6.

[126] More accurately, the standard formulation is an argument against cyclical *preferences*.

[127] For example, we could run the argument exactly as below, but revise the step between D and A such that one *accepts* (rather than pays) a nickel to move from A to D.

The theoretical version of the Money Pump Argument has been taken up by Michael Huemer.[128] Here is his formulation of the argument, copied nearly word-for-word from his article:

(1)   If "better than" is intransitive, then there could be a situation such as that described in the "money pump" scenario, where A is worse than B, B is worse than C, and C is worse than A (in each case by a nontrivial margin). (Premise)

(2)   If *x* is better than *y* (by a nontrivial margin), then it is rational to choose *x* over (*y* plus a small amount of money) when given the choice between those two alternatives, and this is true regardless of what previous choices one has made. (*Optimizing*)[129]

(3)   Therefore, if "better than" is intransitive, then it would be rational to make *each* of the trades in the money pump scenario, regardless of what previous choices one has made. (1 & 2)

(4)   If it is rational to do *x*, and it is rational to do *y* whether or not one has done *x*, then it is rational to do *x* and *y*. (*Agglomeration*)

(5)   Therefore, if "better than" is intransitive, it would be rational to make the series of trades described in the money pump scenario. (3 & 4)

---

[128] In Huemer 2013, 332-5.

[129] Optimizing follows from the Reason-Giving Thesis and the plausible claim that it is rational to choose the option, of one's two available options, that one has more reason to choose.

(6)   It would not be rational to make the series of trades described in the money pump scenario. (Premise)

(7)   "Better than" is not intransitive. (5 & 6)

Huemer guesses that, in response to this argument, Temkin would reject (2), Optimizing. His guess is based on Temkin's thought that "it is often rational to refrain from taking some action, to prevent yourself from taking some other action later."[130] But I believe that Temkin, in offering that thought, is offering a kind of practical strategy for avoiding what I call the *practical* version of the Money Pump Argument (discussed in the next section). Temkin's thought is that rational agents, when they can foresee that they are in danger of being money-pumped, can *commit* or *bind* themselves to choosing in ways that do not result in their being money-pumped. That is, perhaps at time T_1 an agent can rationally force herself to choose *y* over *x* at time T_2, even if at T_2 it *would* be rational for her to choose *x* over *y*. But this is clearly compatible with (2), Optimizing.

Moreover, I do not believe that Temkin *could* deny Optimizing, since Temkin endorses the Reason-Giving Thesis.[131] And the Reason-Giving Thesis implies that if

---

[130] These are Huemer's words (2013, 333); they are based on Temkin's (2012, 188-93) words.

[131] Temkin refers to what I call the Reason-Giving Thesis as "[his] reason-implying sense" of "all-things-considered better than" (2012, 10-18). It is worth noting that this sense of betterness does not presuppose the existence of any bearers of reasons. For example, the view could be amended easily to say that O_1 is better than O_2, in an agent-less world, if and only if there *would* be more reason for agents to choose to bring about O_1 than to bring about O_2, *if* agents existed.

O_1 is better than O_2, then there is more reason to choose O_1 than O_2. It would thus be *rational* to choose O_1 over O_2, if they were the only options available.[132]

Instead, I believe that Temkin, insofar as he denies Transitivity, would and should deny (4), Agglomeration (which states that, "If it is rational to do *x*, and it is rational to do *y* whether or not one has done *x*, then it is rational to do *x* and *y*"). Indeed, Agglomeration does not seem to be *independently* intuitively plausible. Its intuitiveness appears to be restricted to cases in which Transitivity is preserved, and thus it cannot plausibly be invoked to rule out cases in which Transitivity is violated. To see this, simply consider how plausible Agglomeration appears on the assumption that there are intransitive rankings: temporarily assume that $A > B$, $B > C$, and $C > A$. Given this, it is entirely plausible that, given the choice between keeping A and exchanging it for C at the cost of a nickel, it is rational to make this exchange (holding all else constant). And it is entirely plausible that, given the choice between keeping C and exchanging it for B at the cost of a nickel, it is rational to make this exchange (holding all else constant). And so on. However, it is *not* at all plausible that it is rational to make a *series* of such exchanges which would result in getting money-pumped. If, for example, one could press a button which would force one to make all of the exchanges, it would be very irrational to press it. This is perfectly consistent with claiming that it would be rational to make *each* of the exchanges *in* the series,

---

[132] Stuart Rachels (1998) appears to reject (2). He writes, "A rational person with intransitive preferences and adequate information cannot be money-pumped because she will reject the principle, *it is always wise to give up something to get something better*. By denying this principle she may avoid both contradiction and poverty. For example, having Z, she may consistently (and wisely) refuse to embrace Y, even though she knows that Y is better than Z" (82). Insofar as the principle Rachels is denying is that, *at time T_1, it is* rational *to choose Y over Z, given than Y is better than Z*, he is denying the Reason-Giving Thesis (Temkin's reason-implying sense of "better than").

taken individually. Given the temporary assumption of intransitivity, the fact that

these claims are inconsistent with Agglomeration is not troubling, nor is it

counterintuitive. An implication of the intransitive ranking that we assumed is that it

very much matters what one's options at various particular times are. It matters a lot

whether one's options are only **(i)** stick with one item (C) or **(ii)** go the next (A), or

whether they are instead **(i)** stick with one item (C) or **(ii)** go to the next (A) *or* **(iii)** go

to the next (B), etc. When we assume intransitive rankings, Agglomeration loses its

intuitive bite. This suggests that it is not independently intuitively plausible, that its

intuitive plausibility is restricted to cases in which Transitivity is maintained, and that

it cannot plausibly be invoked in an argument for Transitivity.

6.2.2   The Practical Money Pump Argument

In the theoretical Money Pump Argument, it is argued that those who deny

Transitivity, and accept cyclical betterness rankings, cannot avoid the implication that

it would be rational to be money-pumped without denying plausible assumptions. I

just argued that this argument fails. There is, however, a different kind of Money

Pump Argument. The argument is that those who deny Transitivity, and accept

cyclical betterness rankings, sometimes cannot plausibly avoid *being* money-pumped,

and that this reveals the implausibility of such rankings.

My reply to this version of the Money Pump Argument is considerably shorter than my reply to the theoretical version. That is because I believe the practical version of the Money Pump Argument is simply a *non-sequitur*. It confuses, I believe, *practical reasons* for holding beliefs with *epistemic reasons* for holding beliefs. Roughly, whereas we have practical reason to hold beliefs which have good effects, we have epistemic reason to hold beliefs which are true. Showing that we do not have practical reasons for holding some belief, or that we have practical reasons *not* to hold some belief, would not by itself imply that we lack epistemic reasons for holding it. Suppose that if I believe that determinism is true, I will become severely depressed. This might give me practical reason not to believe that determinism is true, but it would not give me epistemic reason not to believe that determinism is true. That believing P would make me depressed is not evidence that P is false. This is a familiar observation. Similarly, believing that A > B > C > A might have bad effects, in the form of susceptibility to money pumping.[133] But if so, this would not be evidence that this belief is false.

This is not to deny that it is worthwhile project to figure out whether, and how, people who believe that cyclical betterness rankings are true (and who have cyclical preferences based on these beliefs) can avoid being money-pumped. But this project is like the project of figuring out whether, and how, people who believe that consequentialism is true can avoid doing worse on their own terms. These projects are

---

[133] Parfit notes that cyclical preferences might have good effects: "Suppose that, whenever our situation changed in some way that we preferred, that change would give us some pleasure. If we had three such cyclical preferences about three easily changeable situations X, Y, and Z, this would be, in a minor way, good for us. We could go round and round this circle, getting pleasure from each move. This merry-go-round would be, hedonically, a perpetual motion machine" (2011, 128).

not about whether such beliefs are *true*. Even if believing consequentialism made things worse according to consequentialism, this would at most yield a practical reason not to believe it. It would not imply that consequentialism is false.

Perhaps I have not accurately portrayed the practical version of the Money Pump Argument. Perhaps the point of this argument isn't, or needn't be, that cyclical betterness rankings are false because susceptibility to money pumping is merely a *bad effect* of believing such rankings to be true. Perhaps the point is that an agent's belief that such rankings are true is responsible for *very irrational behavior* that the agent will engage in. But this too seems irrelevant from the standpoint of what we have epistemic reason to believe. Even if believing P would cause one to behave very irrationally, in addition to having other bad effects, this would not be evidence that P is false. In sum, we have not heard how or why susceptibility to money pumping is a sign of the epistemic irrationality of holding certain beliefs which make one so susceptible. Susceptibility to money pumping seems, at most, to give us practical reason not to accept cyclical betterness rankings. But this is not our question here. Our question is whether any such cyclical ranking is true.

It seems that neither the theoretical nor the practical version of the Money Pump Argument delivers a plausible case against cyclical betterness rankings. I will now turn to a different argument.

6.3   The Dominance Argument[134]

In what follows, I will use "&" to refer to *combinations* of outcomes. That is: "A & B" refers to the outcome in which both outcomes A and B obtain.

Michael Huemer appeals to the following principle in an argument for Transitivity:[135]

> *Dominance Principle*: for all outcomes x_1, y_1, x_2, y_2 … x_n, y_n, if (i) $x_1 > y_1$, (ii), $x_2 > y_2$ … (n) $x_n > y_n$, and there are no "evaluatively significant relationships" among any of these outcomes, then $(x_1 \& x_2 \& … \& x_n) > (y_1 \& y_2 \& … \& y_n)$.

Although Huemer does not himself elaborate, here is roughly what I take his "no evaluatively significant relationships" condition to amount to: (a) the value of a combination of outcomes is solely dependent on the values of its constituent outcomes, and (b) the values of these outcomes are solely dependent on their internal features, and so are in no way dependent on the combination of these outcomes with other outcomes.[136]

---

[134] I am indebted to Tim Campbell for very helpful discussions of the Dominance Argument.

[135] Huemer 2008b, 905-6. Also see Huemer 2013, 335-6.

[136] In other words, this condition brackets evaluative holism, or organic unities. See Moore 1903/1988 for an early defense of holism about the good.

At an intuitive level, the Dominance Principle says: as long as (a) and (b) hold, the *combination* of the outcomes ranked as better is better than the combination of the outcomes ranked as worse. I will now look at the argument in which this principle figures.

6.3.1   Asymmetry, Dominance, and Transitivity

Huemer offers the Dominance Principle as part of an argument for Transitivity, which I will call the *Dominance Argument*. Here is how it goes. According to Transitivity, if A > B, and B > C, then A > C. Again, this makes cyclical rankings, e.g., A > B > C > A, impossible. But we can offer independent support for the view that cyclical rankings are impossible, by appealing to the Dominance Principle and another principle called *Asymmetry*. According to Asymmetry: if A > B, then *not* B > A. The Dominance Argument, then, is a purported *reductio ad absurdum* which accepts the Dominance Principle and Asymmetry as premises, and assumes (for a reductio) that a given cyclical ranking is true. For example:

1. A > B > C > D > A   Assumption

2. The ranking in (1) meets conditions (a) and (b)   Assumption

3. (A & C) > (B & D)   1, 2, Dominance

4. (D & B) > (A & C)   1, 2, Dominance

5. (B & D) > (A & C)   4, Rearranging

6. *Not* (B & D) > (A & C)   3, Asymmetry

7. Contradiction   5 and 6

Notice that if there were a sufficiently strong positive evaluatively significant relationship between, say, A and C, and no other evaluatively significant relationships, Dominance would no longer apply, and it would fail to be true that (D & B) > (C & A). We would thus be unable to derive a contradiction.

But the Dominance Argument is only being offered against cyclical rankings for which (a) and (b) hold. Once the contradiction is derived, the Argument continues: since the Dominance Principle and Asymmetry are true, cyclical rankings which meet (a) and (b) are impossible.[137]

Nonetheless, it is a mistake to think that the Dominance Argument supports Transitivity. This is because, at best, the Dominance Argument will support Transitivity when conditions (a) and (b) hold. It cannot rule out at least some cyclical rankings for which (a) or (b) are false. (As I just noted, it cannot rule out A > B > C > D > A, where there is a sufficiently strong positive evaluatively significant relationship between, A and C, and no other evaluatively significant relationships).

---

[137] Remember that Transitivity does not just rule out cyclical rankings. It also rules out the following acyclical ranking: A > B > C > D = A. As it is formulated above, the Dominance Principle, when combined with Asymmetry, would not rule out intransitive acyclical rankings. But Huemer might suitably expand the Dominance Principle such that it would. For example, he might argue that if the ranking A > B > C > D = A meets conditions (a) and (b), then we can appeal to an expanded Dominance Principle to derive not just that (A & C) > (B & D), but that since D = A, it follows that (A & C) > (B & A), and thus that C > B. But B > C, and so given Asymmetry, we have a contradiction.

But, since Transitivity is inconsistent with any and all cyclical rankings, the Dominance Argument will be unable provide a complete argument for it.

One might still hope that the Dominance Argument can show, for many Spectrum Arguments, why it is implausible to accept their Premises P1 through Pn, and yet deny their conclusions. In this way, the Dominance Argument would serve the same purpose in a Spectrum Argument as Transitivity. Presumably the sorts of Spectrum Arguments the Dominance Argument would be relevant to, if any, would be those in which the outcomes concerned meet conditions (a) and (b). The Hangnails for Torture Spectrum Argument seems to be one of best candidates for such a Spectrum Argument. This is because it seems plausible that the badness of a painful episode is solely dependent on its internal features, and thus that (b) holds, and the badness of a combination of painful episodes is solely dependent on the badness of each such episode, and thus that (a) holds. (At least, these claims seem plausible if the painful episodes are realized in separate persons).

Now recall that the Dominance Principle says, for outcomes $x_1, y_1, x_2, y_2 \ldots x_n$, y_n, as long as (a) and (b) hold, if (i) $x_1 > y_1$, (ii), $x_2 > y_2 \ldots$ (n) $x_n > y_n$, then (x_1 & x_2 & $\ldots$ & x_n) > (y_1 & y_2 & $\ldots$ & y_n).

Consider a concrete example from Stuart Rachels, *Twelve Bad Headaches*[138] (in this example, n = 6):

---

[138] Taken from Rachels 1998. In personal communication, Parfit expressed some doubt as to whether it is even at all tempting to rank these twelve bad headaches intransitively. We might, however, simply regard this example as an abbreviated version of the Hangnails for Torture spectrum (where each pain is

x_1: 5 minutes: a wrenching migraine headache. Your head is ready to explode.

y_1: 10 minutes: a pounding migraine headache somewhat less bad than the headache in x_1.

x_2: 20 minutes: a hideous headache somewhat less bad than the headache in y_1.

y_2: 40 minutes: a terrible headache somewhat less bad than the headache in x_2.

x_3: 90 minutes: a dreadful headache somewhat less bad than the headache in y_2.

y_3: 3 hours: a headache somewhat less bad than the headache in x_3.

x_4: 6 hours: a headache somewhat less bad than the headache in y_3.

y_4: 12 hours: a headache somewhat less bad than the headache in x_4.

x_5: 1 day: a headache somewhat less bad than the headache in y_4.

y_5: 2 days: a headache somewhat less bad than the headache in x_5.

x_6: 4 days: a headache somewhat less bad than the headache in y_5.

y_6: 1 week: a headache somewhat less bad than the headache in x_6. Its pains are only slightly worse than temporary unconsciousness.

Suppose, as Rachels believes, that $x_1 > y_1 > x_2 > y_2 > x_3 > y_3 > x_4 > y_4 > x_5 > y_5 > x_6 > y_6 > x_1$. This cyclical ranking can be represented as the following "betterness mapping" (the arrows point from the better outcomes toward the worse outcomes):

---

only *slightly* less intense but *much* longer). The point is that we can devise some series of pains A, B, C, …, Z, such that it *is* intuitive to claim that A > B, that B > C, …, that Y > Z, and that Z > A.
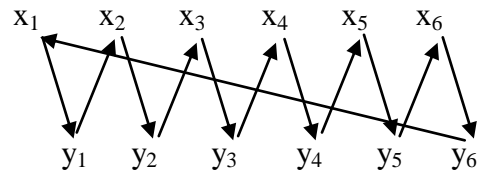
Figure 4: Betterness Mapping #1[139]

The Dominance Principle implies that, since each x is better than each y, the combination of the x's is better than the combination of the y's, or $(x_1 \,\&\, x_2 \,\&\, \ldots \,\&\, x_6)$ > $(y_1 \,\&\, y_2 \,\&\, \ldots \,\&\, y_6)$. Here the Dominance Principle "focuses on" the following half of the above betterness mapping:

Figure 5: Betterness Mapping #2

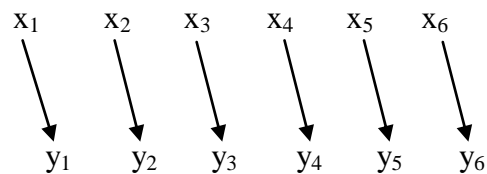But the Dominance Principle also implies that, since each y is better than each x, the combination of the y's is better than the combination of the x's, or $(y_6 \,\&\, y_1 \,\&\, \ldots \,\&\, y_5)$ > $(x_1 \,\&\, x_2 \,\&\, \ldots \,\&\, x_6)$. Here the Dominance Principle "focuses on" the complementary half of the betterness mapping:

---

[139] These Betterness Mappings are owed to Tim Campbell.
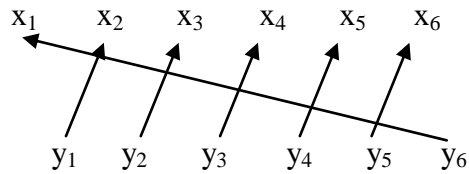
Figure 6: Betterness Mapping #3

Given the cyclical ranking, $x_1 > y_1 > x_2 > y_2 > x_3 > y_3 > x_4 > y_4 > x_5 > y_5 > x_6 > y_6 > x_1$, the Dominance Principle implies both that $(x_1 \ \& \ x_2 \ \& \ \ldots \ \& \ x_6) > (y_1 \ \& \ y_2 \ \& \ \ldots \ \& \ y_6)$ *and* that $(y_1 \ \& \ y_2 \ \& \ \ldots \ \& \ y_6) > (x_1 \ \& \ x_2 \ \& \ \ldots \ \& \ x_6)$. This amounts to: $A > B$ and $B > A$, which, given Asymmetry, is a contradiction.

6.3.2   A Reply to the Dominance Argument

I am doubtful that the Dominance Argument has much, if any, force against people who are tempted to accept cyclical betterness rankings. This is because the Dominance Principle seems to assume the Orthodox View, that "is better than" means "is more good than," and that "good" is measurable along a single linear scale. If A is better than C and B is better than D, and the Orthodox View is true, then A is more good than C and B is more good than D. If we then *add* the goodness of A and B, and *add* the goodness of C and D, then it is a mathematical fact that the two things with more goodness (A and B) will together contain more goodness than the two things with less goodness (C and D). If we find the Dominance Principle intuitively

plausible, it is, I believe, because we are (implicitly) assuming that the Orthodox View is true.

However, arguments like the Hangnails for Torture Spectrum Argument are the very arguments which call the Orthodox View into question. Unless it is supplemented with some independent reason for accepting the Orthodox View, it seems that the Dominance Argument fails in roughly the same way that the Analytic Truth Argument fails (see 6.1).

If the Orthodox View is false, it does not seem clear that if A is better than C and B is better than D, then A and B is better than C and D, even assuming conditions (a) and (b) are met. In the Hangnails for Torture Spectrum Argument, it seems plausible that conditions (a) and (b) are met. Nonetheless, if the Orthodox View is false, it is false that the badness of mild pains and extremely intense torture are measurable along a single linear scale. And so it is false that we can, in any straightforward way, *add* together the badness of mild pains and intense tortures, to arrive at their combined badness. If we thought, for example, that some amount of intense torture is worse than any amount of mild pain, adding the badness of the mild pains to the badness of the torture would be like adding finite numbers to infinity. Addition does not make sense here.

Consider the cyclical ranking, $x_1 > y_1 > x_2 > y_2 > x_3 > y_3 > x_4 > y_4 > x_5 > y_5 > x_6 > y_6 > x_1$. If the Orthodox View were true, then we could, using the Dominance Principle, generate the contradiction that $(x_1 \& x_2 \& \ldots \& x_6) > (y_1 \& y_2 \& \ldots \& y_6)$ while $(y_1 \& y_2 \& \ldots \& y_6) > (x_1 \& x_2 \& \ldots \& x_6)$. But if the Orthodox View were

false, the Dominance Principle would fail to be true as a general principle, and only

apply in a restricted set of cases. If the Orthodox View were false, it might be

plausible that, although (viewed one way) each x is better than each y, $(x_1$ & x_2 & …

& $x_6)$ and $(y_1$ & y_2 & … & $y_6)$ are equally bad or imprecisely equally bad. Or it might

be indeterminate which is worse. Or we might simply be uncertain which is worse.

To summarize, the Dominance Argument appeals to the Dominance Principle,

but the Dominance Principle seems plausible only if the Orthodox View is true. But

Spectrum Arguments (including the Hangnails for Torture Spectrum Argument) tempt

people to doubt the Orthodox View. For the Dominance Argument to succeed, we

require some independent reason to accept the Orthodox View. Therefore, at least

considered by itself, the Dominance Argument does not deliver a very compelling case

against cyclical betterness rankings. I will now turn to a different and more promising

argument.

6.4   The No Dilemmas Argument

It is difficult to deny the Premises (P1 through P*n*) of many Spectrum

Arguments; at the same time, it is difficult to accept their conclusions. The Premises

entail a betterness ranking: A > B > C > … > Y > Z. Such arguments, then, have the

conclusion that A > Z. It is typically difficult to accept the conclusions of such

arguments, because it is difficult to deny that Z > A. If we accept Transitivity, we are

forced to decide: either deny one of the Premises, or accept the conclusion. But it may be that, whether or not Transitivity is true, it is implausible to accept P1 through P*n*, or A > B > C > … > Y > Z, and to claim that Z > A. If we did, we would be accepting a cyclical betterness ranking: A > B > C > … > Y > Z > A. There might be some plausible principle which, though distinct from Transitivity, is inconsistent with the cyclical betterness ranking: A > B > C > … > Y > Z > A. Such a principle would, I believe, be just *as good as* Transitivity for the purposes of this dissertation, since the main question here is whether we can plausibly accept P1 through P*n*, or A > B > C > … > Y > Z, and concurrently claim that Z > A. In this section, I will present and explore the plausibility of a candidate principle of this sort. I call it the:

No Worse Option Principle: there is no set of options such that, within this set, each option is worse than another.

In other words, for each set of options an agent could have a time, there is at least one option such that there is *no worse option* than it. The No Worse Option Principle is consistent with sets of options, as in Buridan's Ass, where each option is equally good.

In 6.4.1, I will discuss which cyclical betterness rankings this principle rules out, and which it fails to rule out. In 6.4.2, I will present an argument for the No Worse Option Principle. And in 6.4.3, I will discuss a seemingly powerful challenge to it.

6.4.1   Option Sets: Within and Across

Unlike Transitivity, the No Worse Option Principle is consistent with *some*

types of cyclical betterness rankings. I will now quickly illustrate this.

Let A > B (A, B) read:

A is better than B, where the set of options is: bring about A or bring about B.

An *option set* is the set of outcomes an agent can bring about at a given time. I

may sometimes refer to an outcome as an option, as shorthand for "it is an option to

bring about this outcome."

According to some, it is possible that A > B (A, B), but B > A (A, B, C). That

is, it is possible that how we rank outcomes A and B ought to vary depending on

which option set they are in. There are several purported instances of this

phenomenon. Here is one discussed by Frances Kamm.[140] Consider three outcomes:

A: A paraplegic person becomes unparalyzed.

---

[140] Kamm 2009. Strictly speaking, Kamm is not here concerned with *betterness*, but with which options
we *ought* to choose. But this example still serves its illustrative purpose.

B: A paraplegic person remains paralyzed.

C: An unparalyzed person remains unparalyzed.

Kamm defends a moral theory which implies that: A > B (A, B), B = C (B, C), and *not* A > C (A, C). This, at least according to some people, is an intransitive ranking.[141] However, her theory also implies that: A = C > B (A, B, C), which is not intransitive. I am not here commenting on the plausibility of Kamm's theory; I mention it merely as an example of a view which implies that how we rank A, B, and C ought to vary depending on which outcomes are in the option set. In particular, because it has this feature, Kamm's theory can consistently espouse intransitivity *across* option sets whilst retaining transitivity *within* option sets.

It should now be clear that we can distinguish between two types of cyclical rankings:

(1) *cyclical rankings* across *option sets*, e.g., A > B (A, B), B > C (B, C), and C > A (A, C)

And:

(2) *cyclical rankings* within *option sets*, e.g., A > B > C > A (A, B, C)

---

[141] Some people claim that Transitivity only *applies* within option sets, such that there cannot be such a thing as intransitivity across option sets. This seems to be a terminological issue. But to the extent that Transitivity does only apply within option sets, the No Worse Option Principle is very close to it.

Transitivity is inconsistent with both (1) and (2). However, the No Worse Option Principle is only inconsistent with (2). If an agent only had two options at a time, e.g., A and B, or B and C, or A and C, she could always avoid choosing an option which is worse than another option. But if an agent had to choose between all three options at the same time, e.g., A, B, and C, then she could not.

It might sometimes seem intuitively plausible to claim that A > B (A, B), that B > C (B, C), and that C > A (A, C), but intuitively *implausible* to claim that A > B > C > A (A, B, C). This can happen when the relative goodness of A, B, and C seems to depend on which option set they are in. But then, other times, it seems to us that the relative goodness of A, B, and C does *not* or can*not* depend on which option set they are in. Recall, once more, a specific formulation of the Series in the Hangnails for Torture Spectrum Argument (from Rachels):

A: 1 year of excruciating agony.

B: 100 years of pain slightly (or somewhat) less intense than the pain in A.

C: 10,000 years of pain slightly less intense than the pain in B.

D: 1 million years of pain slightly less intense than the pain in C.

           …and so on…

Y: 10^{48} years of pain slightly less intense than the pain in X.

Z: 10^{50} years of pain slightly less intense than the mild pain in Y. Each moment of the very mild pain in Z is only slightly worse than temporary unconsciousness.

Intuitively, it is very hard to deny that A > B (A, B), that B > C (B, C), ... and that Y > Z (Y, Z). It also seems intuitively plausible that Z > A (A, Z). Taken together, these intuitions imply a cyclical betterness ranking across option sets.

However, perhaps unlike Kamm's paraplegia example, we are not here tempted to *change* this ranking when the option set is different, and even when the option set includes all of the outcomes in the Series. In particular, it still seems intuitively plausible that A > B, (A, B, C, …, Y, Z), that B > C (A, B, C, …, Y, Z), ... and that Y > Z (A, B, C, …, Y, Z). Our intuitions, taken together, favor the cyclical ranking A > B > C > … > Y > Z > A, whatever the option set is. [142] And even if they did not favor this *cyclical* ranking, it seems independently implausible that our intuitions about how to rank the outcomes in the pain spectrum should vary depending on what the option set is. For example, one year of excruciating agony seems better than one hundred years of slightly less intense pain, regardless of whether these are the only two outcomes in the option set, or whether every outcome in the above Series is.

---

[142] One possibility is that, when our option set contains A through Z, options near the middle (e.g., N) will seem more choiceworthy than options at either end (e.g., A or Z). Though a possibility, this seems implausible. Insofar as this possibility seems plausible, I suspect that this is merely an artifact how the options are presented. We rule out A and Z as not choiceworthy because these options stand out when the options are presented as (A, B, C, D, E, F, G, H, I, J, K, L, M, N, O, P, Q, R, S, T, U, V, W, X, Y, Z); our attention is drawn to the implausibility of A and to the implausibility of Z. If the options were instead presented as (N, O, P, Q, R, S, T, U, V, W, X, Y, Z, A, B, C, D, E, F, G, H, I, J, K, L, M, N, O), we might be more tempted to claim that options around the new middle (e.g., between W and C) are more choiceworthy. Perhaps these presentational or ordering effects could be avoided if the options were represented by equally-spaced points along the circumference of a circle, rather than as points along a straight line segment. I make a similar point about a similar ordering effect in 3.3.2, in the discussion of Parity.

In the Hangnails for Torture Spectrum Argument, the distinction between cyclical betterness rankings across option sets and cyclical betterness rankings within option sets seems unimportant. Indeed, this is generally true of the Spectrum Arguments most relevant to my purposes in this dissertation, including those discussed in chapters 3 and 4. In these Spectrum Arguments, how the outcomes are ranked cannot plausibly depend on what the option set is. Moreover, in these Spectrum Arguments, if our intuitions, taken together, support the cyclical betterness ranking A > B > C > … > Y > Z > A, they also, taken together, support A > B > C > … > Y > Z > A (A, B, C, …, Y, Z).

This is all to say that the Premises of the Spectrum Arguments which here concern us, combined with the denials of their conclusions,[143] imply cyclical betterness rankings of the form A > B > C > … > Y > Z > A (A, B, C, …, Y, Z). Thus, a principle which is inconsistent with such cyclical betterness rankings within option sets is just as good as Transitivity for our purposes here.

The No Worse Option Principle is such a principle. It states that there is no set of options such that, within this set, each option is worse than another. But if there are cyclical betterness rankings within option sets, there are sets of options such that, within these sets, each option is worse than another. It is, therefore, just as important for our purposes whether the No Worse Option Principle is true as it is whether

---

[143] Strictly speaking, the denial of the conclusion of a Spectrum Argument is of the form *not* A > Z. But we do not imply a cyclical betterness ranking unless, in addition to the Premises, we accept Z > A. But this detail is not important, since for all of the Spectrum Arguments I discuss, we do not merely find it counterintuitive that A > Z, but also that *not* Z > A.

Transitivity is true. I will now consider an argument for the No Worse Option

Principle.

6.4.2 From No Dilemmas to No Worse Option

Recall:

Optimizing: if an agent chooses the worse of her only two options, and other

things are equal, she chooses irrationally. (Note: this is a slight reformulation

of Optimizing as it appeared in the theoretical Money Pump Argument).

We can now consider the more general claim:

Optimizing*: if an agent chooses an option that is worse than another option,

and other things are equal, she chooses irrationally.

If this claim and

No Dilemmas: there is no set of options such that, whichever option is chosen,

it is irrational to choose it.

is also true, then so is the No Worse Option Principle. If there is no set options such that choosing any option is irrational, and if choosing a worse option is irrational, then there is no set of options such that each is worse than another.

In assessing the plausibility of Optimizing* and No Dilemmas, it is important that we distinguish two possible readings of "irrational:"

Failure to Conform to Reason: an agent chooses irrationally when she knowingly, though perhaps faultlessly, chooses an option she has all things considered less reason to choose than another.

And:

Fault: an agent chooses irrationally when she fails to conform to reason (see above), and this failure is *her* fault.

On the Fault reading of "irrational," it is hard to escape the conclusion that No Dilemmas is true and that Optimizing* is false. If, for example, an agent were presented with three options, A, B, and C, and A > B > C > A (A, B, C), she would be forced to choose an option it is all things considered less reasonable to choose than another. Assuming that the existence of this option set is not the agent's fault, her not choosing an option that conforms to reason is not *her* fault, since such an option is *unavailable*. And thus, given the Fault reading, she would not choose irrationally in

choosing contrary to reason, or in choosing an option that is worse than another. Thus, on the Fault reading, Optimizing* is false.

However, I am here assuming the Failure to Conform to Reason reading. I do not claim that this reading uniquely captures the true meaning of "irrational." I am merely stipulating that that is what *I* will mean by "irrational."

Given this stipulation, Optimizing* follows from the Reason-Giving Thesis (that if A is all things considered better than B, and if either could be brought about, then there is all things considered more reason to bring about A than there is to bring about B, other things being equal). So Optimizing* is true. No Dilemmas is perhaps less obviously true on the Failure to Conform to Reason reading than it is on the Fault reading. Nonetheless, many people still find No Dilemmas intuitively plausible, even on the Failure to Conform to Reason reading. In assessing the plausibility of No Dilemmas, it may first help to note two kinds of option sets that it does not rule out.

First, there are tragic situations in which, no matter which option an agent chooses, she must knowingly cause or allow something *terrible* to happen. For instance, the leader of an earthquake rescue team may be forced to choose which group of people to save. Whichever group she chooses to save, the other group will die buried beneath the rubble. Though forced to make a painful choice, the rescue team leader was not forced to choose an option she has all things considered less reason to choose than another. It would not have been irrational of her to choose to save the larger group of people, since other things equal more deaths are worse than fewer. No Dilemmas does not rule out this kind of decision-making tragedy.

Second, there are cases in which reason favors no single option over all others. An ass might have just as much reason to choose one bale of hay as another, and might have no other options. In such a case, it would be rationally permissible to choose *either* option. And there are cases where it is imprecise, indeterminate, or uncertain which options we have more reason to choose. No Dilemmas is consistent with these kinds of cases.

On the other hand, suppose an agent is placed in a situation in which no matter what she chooses she must knowingly choose contrary to reason. That is, she must choose an option she knows to be all things considered less reasonable to choose than another option. (We can suppose that failure to *actively* choose an option still constitutes the selection of an option). In such a case, it is rationally impermissible to choose any option. If, for example, an agent were presented with three options, A, B, and C, and A > B > C > A (A, B, C), reason would recommend the agent *not* to choose *any* option: don't pick A, for you've got all things considered less reason to pick it than C! Also, don't pick C, for you've got all things considered less reason to pick it than B! Oh, and also, don't pick B, for you've got all things considered less reason to pick it than A! An agent receiving such recommendations might regard them as ridiculous or incoherent, for there is nothing else she *can* pick besides A, B, or C. She might, then, find it hard to deny that she always has *some* option it is not irrational to choose.

No Dilemmas thus might seem intuitively plausible. And Optimizing*, as I have explained, is true. And from these two claims follows the No Worse Option

Principle, which rules out cyclical betterness rankings within option sets.

Unfortunately, there are cases in which No Dilemmas and the No Worse Option

Principle appear to be false.

6.4.3   The Challenge of Infinite Options

The No Worse Option Principle is inconsistent with cyclical betterness

rankings within option sets. However, there are transitive betterness rankings which

also apparently conflict with the No Worse Option Principle, which seem harder to

abandon. Consider the following:

Devil's Deal.[144] It is 9:00am now. Satan approaches you and hands you a

form which reads:

> PLEASE INDICATE BELOW THE FINITE NUMBER OF YEARS
>
> YOU WOULD LIKE TO BE IN HEAVEN. IF YOUR ANSWER IS
>
> OR IMPLIES AN INFINITE NUMBER OF YEARS, YOU WILL BE
>
> SENT TO HELL FOREVER. RETURN THIS FORM TO SATAN
>
> BY 10:00AM OR YOU WILL BE SENT TO HELL FOREVER.
>
> THANK YOU.

---

[144] This sort of case is discussed in Arntzenius, Elga, and Hawthorne 2004.

This appears to be a case in which each option is worse than at least one other option: no matter what finite number of years in heaven you pick, you could have picked something better – a *larger* finite number of years in heaven. This means that whatever you pick, you pick something *worse* than another option. Thus, given Optimizing*, in this case, the choosing of any option is irrational (or rationally impermissible). But this is inconsistent with No Dilemmas, which says that there is *no* set of options such that, whichever option is chosen, it is irrational to choose it. So *Devil's Deal* appears to be a counterexample to No Dilemmas. There are several attempts to resist this counterexample, which I will now consider briefly.

First, we might claim that there is some number of years in Heaven such that more years would not be better. Although some have defended it,[145] I believe that this claim is implausible. For any number of years in Heaven n, it seems that $n + 1$ years would be better.

Second, we might deny Optimizing* in this case. We might claim that, as long as we pick a *large* number of years in Heaven, we do not choose irrationally. But this too seems implausible, since it seems that there is all things considered more reason to pick better options, other things equal, and since the "irrational" in Optimizing* and No Dilemmas simply refers to the choosing of an option the agent has all things

---

[145] See Williams's 1978. For replies to Williams, see Fischer 2012 and Rosati 2012. McMahan (2002, 98-103) also offers a short but plausible rejoinder to Williams. Lastly, note that all I need for *Devil's Deal* to be an effective case is the claim that there is no number of years in Heaven n such that $n + 1$ years would fail to be better. Although some people who deny that immortality could be desirable might deny this claim, some might accept it. That is, it is consistent to claim both that a larger finite number of years in Heaven is always better, and that an infinite number of years is worse, or even bad.

considered less reason to choose than another (the Failure to Conform to Reason reading of "irrational").

Third, we might deny that we truly have *infinite* options. We might claim that, between 9:00am and 10:00am, there is only a finite number of finite numbers that you can consider and specify. You should therefore simply write down the largest number that you *can*, in the hour allotted. But we could stipulate that, between 9:00am and 10:00am, Satan gave you the power to consider and specify numbers at an infinite speed (and an infinitely large sheet of paper on which to write your number!). Then there would be no *largest* number that you could specify.

Fourth, we might claim that it is impossible for infinite duration or infinite speeds to exist. These sorts of infinities do seem to me to be metaphysically possible, but even if they were not, surely they are *logically* possible. And if No Dilemmas were false in a logically possible world, it would seem to be false *period*, unless we were given some good reason for restricting its application to a subset of logically possible worlds that excludes those which contain infinite durations and speeds.

Fifth, we might argue that cases like *Devil's Deal* raise puzzles not about No Dilemmas, but about *infinity*, and that, since we are doing ethics, and not the mathematics or metaphysics of infinity, we can ignore such cases involving infinite options, durations, speeds, and so on.[146] But unless one explains *why* such problems really *are* puzzles about infinity, and not various normative principles like No

---

[146] For example, Derek Parfit has expressed doubts about whether anything about moral philosophy can be learned by looking at infinite cases.

Dilemmas, this is not a satisfactory response to *Devil's Deal*. (E.g., *how* might a better mathematical or metaphysical account of infinity solve our problem?) And it is not a sufficient defense of No Dilemmas to point out that cases involving infinity raise puzzles for a wide range of familiar and intuitively plausible normative principles, in addition to No Dilemmas. For a clear possibility is that many such principles, *in addition to* No Dilemmas, are shown to be *implausible* by cases involving infinity. It seems to me an unjustified prejudice to assume that the problem must lie with the unfamiliar or odd cases, rather than with the more familiar principles they challenge.

Unless one of the above five replies to *Devil's Deal* succeeds, or we can find some other successful reply, this case appears to be an effective counterexample to No Dilemmas and to the No Worse Option Principle.[147] Thus, we cannot appeal to the No Worse Option Principle to argue against cyclical betterness rankings within option sets. I will now consider another argument against such cyclical betterness rankings. If this new argument succeeds, it will provide a defense of a version of the No Worse Option Principle that is limited to option sets with a finite number of options.

---

[147] A diachronic version of *Devil's Deal* provides a money pump for transitive rankings (this case is also from the 2004 Frank Arntzenius, Adam Elga, and John Hawthorne paper).

> *Diachronic Devil Deals*. Suppose that just a few minutes before your death in the year 2100, Satan approaches you, and offers to give two years in Heaven if you first suffer for a year in Hell. Call this *Deal #1*. It seems rational to take this deal. Suppose you do. As you are finishing up your year in Hell, Satan offers to give you two *more* years in Heaven if you stay in Hell for another year, from 2101 to 2102. Call this *Deal #2*. It again seems rational to take this deal. Suppose you do. At the end of each year, Satan offers you another such deal. It seems rational to take each deal. Of course, if you do take each, you will remain in Hell forever.

Moreover, this case also seems to provide a counterexample to Agglomeration. This is because it seems rational to take *each* deal, but irrational to take *all* of them.

6.5   The Implosion Argument

I will offer what we might call the *Implosion Argument* against cyclical betterness rankings within option sets. The general strategy of this argument is to first assume the existence of any such ranking, and to then show that this ranking, when combined with plausible claims, *self-refutes*. Different versions of the Implosion Argument are required to target separate types of intransitive ranking. Here are some of these types:

(a)   A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and each outcome is better than *n* others to degree *d* and worse than *n* others to degree *d*.[148]

(b)   A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and it is false that [each outcome is better than *n* others to degree *d* and worse than *n* others to degree *d*].

---

[148] One might worry that in order to sensibly refer to the *degree* to which A, B, and C are better than one another, they must be locatable along a single linear scale of goodness. If so, this intransitive ranking would be incoherent, because if A, B, and C are locatable along a single linear scale of goodness, they cannot possibly violate Transitivity. While this is a very natural worry to have, I do not believe that we must appeal to a single linear scale of goodness in order to sensibly refer to the *degree* to which A, B, and C are better than one another. For example, we could imagine that A, B, and C are represented by equidistant points along the circumference of a circle. That they are points on a circle captures the fact that they are ranked intransitively, and that they are equidistant captures the fact that each item is better than *n* others to degree *d* and worse than *n* others to degree *d*. (We could ignore distances between non-adjacent points on the circle.) There are other geometrical models which illustrate the same point, but one illustration seems enough.

(c)   A > B (A, B, C), B > C (A, B, C), and C is (roughly) as good as A (A, B,

C).

However, to conserve space, and because it appears to be fairly straightforward how to extend the Implosion Argument to target (b)-type and (c)-type rankings, I will focus on (a)-type rankings. (Though below I will briefly spell out an extension of the argument targeting (b)-type rankings).

It is again important to notice that the Spectrum Arguments discussed earlier threaten to imply intransitive rankings of the form A > B (A, B, C), B > C (A, B, C) and C > A (A, B, C), but the rankings that Spectrum Arguments might imply contain more than three outcomes, and instead look like: A > B (A, B, …, Z), B > C (A, B, …, Z)…and so on, Y > Z (A, B, …, Z), and Z > A (A, B, …, Z). The Implosion Argument, if it successfully rules out cyclical rankings with three outcomes or options, would clearly also rule out cyclical rankings with as many outcomes or options as there are in typical Spectrum Arguments. Without further ado, here is the version of the Implosion Argument designed to rule out (a)-type rankings.

(1)   A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and each outcome

is better than n others to degree d and worse than n others to degree d.

(Assumption, reductio)

(2) There is more reason to choose A than there is to choose B, there is more reason to choose B than there is to choose C, and there is more reason to choose C than there is to choose A, and A, B, and C are the only options, and each option is more choiceworthy than n others to degree d and less choiceworthy than n others to degree d. (1 and the Reason-Giving Thesis)

(3) *If* there is more reason to choose A than there is to choose B, there is more reason to choose B than there is to choose C, and there is more reason to choose C than there is to choose A, and A, B, and C are the only options, and each option is more choiceworthy than n others to degree d and less choiceworthy than n others to degree d, *then* A, B, and C are equally choiceworthy. (*Implosion Premise*)

(4) A, B, and C are equally choiceworthy. (2 & 3)

(5) *Contradiction*. (2 & 4)

Assuming the Reason-Giving Thesis, this version of the Implosion Argument against type-(a) cyclical rankings really has only one premise. (2) follows from (1) and the Reason-Giving Thesis (recall that the latter implies that if outcome O_1 is better than outcome O_2 then there is more reason to choose to bring about O_1 than O_2).

Given this, the only *real* premise, and so the only *contestable* premise, is (3) – what I am calling the *Implosion Premise*. It is important to be clear that the Implosion Premise is a *conditional* claim. It is thus not to be confused with (4), the claim that A, B, and C *are* equally choiceworthy. In 6.5.1, I will consider and respond to five

objections to the Implosion Premise. But first, I need to address a concern about the relevance of this Implosion Argument.

The concern is that this Implosion Argument is irrelevant, since while it only targets type-(a) cyclical rankings, the Temkin pain spectrum does not threaten to imply a type-(a) cyclical ranking. Instead, this sort of spectrum threatens to imply a type-(b) cyclical ranking. Recall that type-(b) rankings take the form: A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and it is *false* that [each outcome is better than *n* others to degree *d* and worse than *n* others to degree *d*].

In the Temkin pain spectrum there appears to be something that *does* plausibly make some options more choiceworthy than others – even assuming they are ranked cyclically. Though it is true that no matter which outcome we choose there are some other outcomes that are worse than it, *and* some other outcomes that are better than it, it seems that the *degree* to which some outcomes are worse than others is greater than the degree to which others are worse than others. Moreover, some outcomes appear to be better or worse than a greater *number* of other outcomes than others are; for example, maybe when A through Z are the outcomes to choose from, outcome P is better than 144 others, whereas outcome T is better than 300 others. (I am assuming that Temkin's pain spectrum contains more than 26 different outcomes; letters are merely shorthand for rough locations in the spectrum.)

On the far side of the spectrum are the most intense pains, starting with A, excruciating torture for two years. In the middle we have moderately intense pains that last for *many* years. At the other end, we have very low intensity pains, like Z, the

very mildly annoying hangnail pain, that last for *zillions* of years. While it might be

that A > B (A, B, …, Z), B > C (A, B, …, Z)…and so on, Y > Z (A, B, …, Z), and Z >

A (A, B, …, Z), it seems intuitively plausible that the *degree* to which B is worse than

A is greater than the degree to which Z is worse than Y – it accordingly seems

plausible that the preference for A over B should be *stronger* than the preference for Y

over Z. Thus, it is not true that, for anything that could be said in favor of choosing

one option, an exactly similar thing could be said in favor of choosing any other.

  In response to such type-(b) intransitive rankings, we can offer the following

revised version of the Implosion Argument:

(1) A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and it is false that

  [each outcome is better than *n* others to degree *d* and worse than *n* others to

  degree *d*]. (Assumption, reductio)

(2) There is more reason to choose A than there is to choose B, there is more

  reason to choose B than there is to choose C, and there is more reason to

  choose C than there is to choose A, and A, B, and C are the only options,

  and it is false that [each option is more choiceworthy than *n* others to

  degree *d* and less choiceworthy than *n* others to degree *d*]. (1 and the

  Reason-Giving Thesis)

(3) *If* there is more reason to choose A than there is to choose B, there is more

  reason to choose B than there is to choose C, and there is more reason to

  choose C than there is to choose A, and A, B, and C are the only options,

and it is false that [each option is more choiceworthy than *n* others to degree *d* and less choiceworthy than *n* others to degree *d*], *then* among A, B, and C, some option is the most choiceworthy, or multiple options are (roughly) tied for most choiceworthy. (Implosion Premise)

(4)   Among A, B, and C, some option is the most choiceworthy, or multiple options are (roughly) tied for most choiceworthy. (2 & 3)

(5)   *Contradiction*. (2 & 4)

As with the first version of the Implosion Argument, this revised version only has one contestable premise, which is (3) – again, the Implosion Premise. Therefore, I will now defend the Implosion Premise against some possible objections. However, for the sake of simplicity and illustration, I will do so while focusing on the version of the Implosion Argument targeting type-(a) cyclical rankings (again, rankings that have the form A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), where each outcome is better than *n* others to degree *d* and worse than *n* others to degree *d*.) What I say in defense of the Implosion Premise in this context applies, *mutatis mutandis*, to the Implosion Argument targeting type-(b) cyclical rankings.

6.5.1   Objections and Responses

I will now consider and respond to five objections to the Implosion Premise, as it is invoked in the first version of the Implosion Argument (targeting type-(a) cyclical rankings).

Objection One:

In his earlier work on Transitivity, Temkin considers the situation in which one is presented with all of the pains in his spectrum – A through Z – as options, and he writes:

> Perhaps all the alternatives are equally reasonable in cases of intransitivity. But we don't believe this. We don't believe a long life including two years of intense torture is an equally reasonable alternative to a long life with a hangnail. Moreover, reflection suggests that most alternatives might be placed on a large, intransitive, continuum, analogous to the one involving extreme torture at one end and mild discomfort at the other. This opens the possibility that there would be no rational basis for choosing between virtually any alternatives.[149]

But Temkin's response is more effective against step (4) of the Implosion Argument than it is against *my* claim, which is that *given* (1) – that A > B (A, B, C), B > C (A, B, C), and C > A (A, B, C), and each outcome is better than *n* others to degree *d* and worse than *n* others to degree *d* – *then* (4) is plausible. Temkin is surely right that, to many of us, *each* of the claims about his spectrum seems, considered *individually*, very intuitively plausible: A > B (A, B, …, Z), B > C (A, B, …, Z), C > D (A, B, …, Z),…and so on, Y > Z (A, B, …, Z), and Z > A (A, B, …, Z). And many of us no doubt find *each* of these intuitively plausible, considered *individually*, when

---

[149] 1996, 209.

read as claims about choiceworthiness, or what we have reason to choose (indeed, this is to be expected insofar as we find the claims about betterness plausible and we find the Reason-Giving Thesis plausible). But that is a much different matter from what we would believe about the choiceworthiness of A through Z, if we were simply *given* that, or instructed to *assume* that A through Z form a cyclical ranking, that is, that A > B (A, B, …, Z), B > C (A, B, …, Z)…and so on, Y > Z (A, B, …, Z), and Z > A (A, B, …, Z). Given *that*, I claim, each option seems equally choiceworthy. When we attend narrowly to adjacent options, or when we doubt that A through Z *truly* form an intransitive ranking, we find it implausible that each option is equally choiceworthy. But when we take seriously the assumption that A through Z form an intransitive ranking, and attend widely to our choice situation as a whole, it actually seems very difficult to deny that they are equally choiceworthy.

Why do I claim that, given (1), each option seems equally choiceworthy, or reasonable? Because, *given (1)*, for anything that could be said in favor of choosing any option, an *exactly similar* thing could be said in favor of choosing any other. In particular, no matter which option one chooses, it is true that there is one other option that is less choiceworthy than it, and one other that is more choiceworthy than it. But given that these facts about what could be said in favor of each option are thus relevantly tied, and since there is nothing else that could make any option more choiceworthy than any other, they are *equally* choiceworthy or reasonable.[150] This is

---

[150] I should clarify that I don't think that the fact that one option is more choiceworthy than another is itself, or itself constitutes, a *further* reason for choosing it, but rather that it is reflective of the reasons for choosing this option over others; when it is true that no matter which option one chooses, there is one other option that is less choiceworthy than it, and one other that is more choiceworthy than it, and

what the Implosion Premise claims. Given an intransitive set of reasons, we get a transitive set of reasons.

Objection Two:

We could deny the Implosion Premise if we thought that, when we face type-(a) rankings, it is *merely* rationally permissible to choose any option. In these cases, we might claim, we do not fail to comply with reason or choose irrationally if we choose any option, but it is not the case that we have *equal* reason to choose any.

While we *could* deny the Implosion Premise in this way, it would not be *plausible* to do so. It is true that, given a type-(a) ranking, it is permissible to choose any option. But, I believe, that is *because* each option is equally choiceworthy. And, again, each option is equally choiceworthy because, given the type-(a) ranking assumed in (1), for anything that could be said in favor of any option, an exactly similar thing could be said in favor of any other. Because the "countings in favor" are relevantly tied, the reasons for choosing are tied.

Objection Three:

Suppose that we faced all the options in Temkin's pain spectrum, and that they constituted a type-(a) ranking. One might claim that, even if it were true that, assuming this type-(a) ranking, each option is equally choiceworthy, it would remain

---

other things are equal, *this* is likewise reflective of the fact that the reasons for choosing among these options are relevantly tied. Similarly, as Parfit (2011, 39) notes, the fact that some book is the *best* would not constitute a *further*, or independent, reason for choosing it (on top of the considerations that already make the book better or more choiceworthy than others).

*im*plausible that each option is equally *good*. While such facts about one's choice context can affect one's reasons for choosing, and make (for instance) pain A and pain B equally choiceworthy, they cannot affect how *good* these pains are. And it remains very implausible that these pains are equally good, or equally bad.

In response to this objection we should clarify what *sense* of goodness, or betterness, we are talking about. If we are working with a sense of betterness according to which the Reason-Giving Thesis is true (which implies that if outcome O_1 is better than outcome O_2 then there is more reason to choose to bring about O_1 than O_2), then this objection is nonsense. For according to this Reason-Giving Thesis, there is a necessary link between goodness and reasons. But, there are other senses of goodness and betterness. We might understand goodness and betterness completely independently of reasons. Or we might accept a reason-implying sense of betterness, but one according to which facts about one's choice context of the sort noted above *cannot* affect betterness.[151] If we are working with any of these importantly different senses of betterness, then I simply concede the point. I intend for the Implosion Argument to support Transitivity (or, more accurately, I intend for it to support the No Worse Option Principle restricted to finite option cases) for a reason-implying sense of betterness according to which the Reason-Giving Thesis is true; I do not claim that the Implosion Argument supports Transitivity for other importantly different senses of betterness. This concession may disappoint some Transitivity-defenders, but I believe it is nonetheless the right move to make.

---

[151] For example, perhaps instead of understanding betterness in terms of reasons for *choosing*, we could understand it in terms of reasons for *wanting* or *hoping*. See Parfit 2011, 41.

Objection Four:

One possible reaction to the Implosion Premise is that it is itself deeply confused. After all, it is a conditional of the following form: *if* P, *then* not-P. "But surely," one might react, "*if* we *accept* P, then we cannot also accept not-P!" "If we *really* accept that A is more choiceworthy than B, in the antecedent of the Implosion Premise, then we cannot also claim that A and B are equally choiceworthy, in the consequent of the Implosion Premise."

While this reaction to the Implosion Premise might be natural, I do not believe that it is correct. We in fact already believe in a variety of conditional claims structurally similar to the Implosion Premise, and are often quite right for doing so. Indeed, we correctly accept such conditionals about claims that are *self-refuting*. We can say that P is self-refuting when it is true that P implies not-P. Maybe sometimes P is self-refuting when it is true that P, together with some other plausible claim, implies not-P. In general, I take it that P is *self*-refuting when, if we assume P, we find not-P plausible (at least partly) *on the basis of* P. It is not merely the case that, if we assume P, we nonetheless independently find not-P to be plausible. Many examples of self-refuting claims involve self-reference. Consider: "This statement is false" or "We are justified in believing that no one is justified in believing anything" or "Only empirically verifiable statements can be true." These cases yield a host of conditional claims of the form "if P, then not-P," for example, "if this statement is true, then it is false."

Cyclical betterness rankings, I suggest, are also self-refuting. The Implosion

Premise is importantly like these other conditional claims of the form "if P, then not-

P." Where P is "we have more reason to choose A than B, more reason to choose B

than C, and more reason to choose C than A," we can plausibly conclude that "we

have equal reason to choose A, B, and C," or "not-[we have more reason to choose A

than B, more reason to choose B than C, and more reason to choose C than A]."

Cyclical rankings yield another kind of case in which P seems to imply not-P.

Whether they are self-refuting in the way that several self-referential claims are, they

are self-refuting nonetheless. We might then reserve the notion of *implosion* to refer

to this distinct class of self-refuting claims, to which cyclical rankings belong. These

are claims which are self-refuting, but not in virtue of self-reference.

I believe that the Implosion Premise captures the main intuition had by those

very many people who accept Transitivity. It captures the phenomenology of being

utterly flabbergasted that we have, when presented with a putative cyclical ranking of

reasons. We tend to simply describe such rankings *themselves* as "contradictions" or

"inconsistent" or "choice-defeating."[152] These rankings are not literally inconsistent,

but because of the deep plausibility of the Implosion Premise, we move from these

rankings to their denials rather quickly. If these rankings were true, then this would

plausibly imply that they would actually *not* be true. In the case of type-(a) rankings,

these rankings would not be true because *if* they *were*, then for anything that could be

said in favor of any option an exactly similar thing could be said in favor of any other.

---

[152] For example, note the language that Michael Otsuka (2004, 413-26) and Frances Kamm (2007, 71-4) use to describe intransitive deontic rankings in the context of aggregation.

This antecedent of the conditional *makes* the "countings in favor" relevantly tied across these options – this is why the consequent of the conditional is the denial of its antecedent.

Objection Five:

Finally, one might object that assuming the cyclical ranking in (1) simply *implies* that the Implosion Premise is false. For convenience, let "P" stand for the cyclical choiceworthiness ranking in (2) of the Implosion Argument. Then (3), the Implosion Premise, would be "if P, then not-P." One might reason as follows: If I accept (1), then given the Reason-Giving Thesis I must also accept (2), or P. But then I will see that it cannot be the case that "if P, then not-P." After all, if P is true, then it is false that "if P, then not-P," since if both were true we could derive "P and not-P," a contradiction.

In response to this objection, I contend that even when we are temporarily assuming P, it *still* seems plausible that "if P, then not-P." This is because it *still* seems plausible, when we assume P, the cyclical reasons ranking, that we can plausibly conclude *from* this cyclical ranking that our options are equally choiceworthy. It is true that P and "if P, then not-P" are inconsistent. What this shows, I contend, is that insofar as we find P intuitive (or insofar as we have intuitions which jointly imply P), our intuitions are inconsistent. And what we must do when our intuitions are inconsistent is try to determine which are stronger, more likely to be reliable, etc. But, importantly, we *would* still find "if P, then not-P," or the Implosion

Premise, to be intuitively plausible, even when assuming P. This is what sets the intuition behind the Implosion Premise apart from other intuitions (in favor of premises in other arguments) for Transitivity. I believe that the intuition in support of the Implosion Premise is strong enough to rule out cyclical betterness rankings and cyclical reasons rankings.

This concludes the objections and responses.[153]

6.5.2   Concluding Remarks

The Implosion Argument is, I tentatively believe, successful. I will end this chapter by summarizing this argument's implications and advantages (over other arguments discussed in this chapter).

---

[153] Another note about how the Implosion Argument can be extended to type-(b) rankings: In the context of type-(b) rankings, the Implosion Premise is plausible. When given type-(b) rankings, it is not true that, for anything that could be said in favor of choosing one option, an exactly similar thing could be said in favor of choosing any other. There is *more* to be said in favor of some options than others. This is true, for example, when these options are more choiceworthy than a greater number of options than other options are. Perhaps option T is more choiceworthy than 300 options, whereas every other option is at most more choiceworthy than only 298 options. There is also more to be said in favor of T when, for example, the degree to which T is more choiceworthy than other options is, on average, *d*, whereas the degree to which every other option is more choiceworthy than other options is at most, on average, significantly less than *d*. If these claims were true of T, perhaps it would be the most choiceworthy option. Finding the (roughly tied for) most choiceworthy option(s) in the case of type-(b) rankings may not be easy, but all I need to claim is that there *is* some option that is the most choiceworthy, or that there *are* multiple options are (roughly) tied for most choiceworthy. I need not spell out what, in particular, such options are. The consequent of the Implosion Premise does not require that. And, in this revised version of the Implosion Argument targeting (b)-type rankings, the Implosion Premise again seems plausible. This yields another sort of case where, given P, not-P seems to follow.

- The Implosion Argument avoids the problems with Broome's Analytic Truth Argument (6.1), and with Huemer's Dominance Argument (6.3), in that it doesn't presuppose that the Orthodox View is true.

- It avoids the shortcomings of the Money Pump Argument (6.2).

- It avoids the objection to the No Dilemmas Argument in 6.4.3, as it is consistent with infinite option sets where the options are ranked transitively (as in *Devil's Deal*). It provides a plausible rationale for accepting the No Worse Option Principle when the option sets are a finite size, while rejecting it when the option sets are an infinite size.

- The argument is fairly simple and straightforward, and it seems to capture our intuitive sense that cyclical betterness rankings (and cyclical reasons rankings) within option sets are inherently implausible.

- The key premise of the Implosion Argument – the Implosion Premise – does not seem plausible *merely* in cases in which cyclical betterness rankings within option sets seem, for independent reasons, false or implausible. *Even* in the very cases where we might be tempted to accept such cyclical rankings – e.g., the Hangnails for Torture Spectrum Argument – the Implosion Premise *also* seems plausible. That is, the Implosion Premise seems *independently* intuitively plausible, and indeed it seems plausible in precisely the sort of relevant "test cases" that it should. Note, by contrast, that the Orthodox View, Agglomeration, the Dominance Premise, and No Dilemmas, do *not* seem to be similarly independently intuitively plausible; they seem to lose their intuitive appeal in the relevant "test cases," and so

cannot plausibly be invoked in arguments for Transitivity or for the No Worse
Option Principle.

Finally, recall the reason we were concerned to defend Transitivity in the first
place. It was to show that it is implausible to accept the Premises of Spectrum
Arguments and to simultaneously deny their conclusions. In 6.4.1, I explained that we
would essentially show this if we showed cyclical betterness rankings of the form $A >
B > C > … > Y > Z > A$ (A, B, C, …, Y, Z) to be implausible. And the Implosion
Argument does show that such rankings are implausible.

While I have not shown that Transitivity is true, I believe I have given us a
decent argument for accepting the No Worse Option Principle (restricted to finite
option sets). And that principle is as good as Transitivity, for the purposes of this
dissertation.

Chapter 7

Concluding Chapter

CHAPTER OUTLINE

7.1   Undiagnosed Paradoxes?

7.2   Least Implausible Solutions

7.3   Revisionary Intuitionism

7.1   Undiagnosed Paradoxes?

This dissertation has relied heavily on Spectrum Arguments to show that rivals to the Equal Weight View face implausible implications. Earlier I mentioned five different kinds of solution we can offer to Spectrum Arguments:

1. Accept the conclusion (e.g., claim the intuitive evidence against the conclusion is outweighed by other intuitions, or argue with the Large Number Skeptics, from 4.3, that the intuitions that the conclusions of Spectrum Arguments are false are unreliable).

2. Deny some but not all of the Premises (deny Parity or the Conditional).

3. Deny all of the Premises (e.g., this is how defenders of the Equal Weight View would respond to the Priority Monster Spectrum Argument).

4. Deny Transitivity (see chapter 6).

5. Claim that the argument is unsound in the way Sorites Arguments are unsound (see chapter 5).

Below I will discuss briefly whether this list is exhaustive. But first let me make some general observations about what Spectrum Arguments purportedly do.

There are many intuitively plausible principles that imply claims of the form: *a little less X and a lot more Y is better*. (E.g., "a slightly smaller benefit for a much worse off person is better," or "a population with a slightly lower per capita quality of life that is much larger is better"). We can call these principles *Tradeoff Principles*. Many of these Principles, which endorse very modest tradeoffs, seem very hard to deny. Spectrum Arguments purportedly show that these Tradeoff Principles imply conclusions of the form: *for any amount of X, there is some amount of Y that is better*. (E.g., "for any arbitrarily large benefit for any arbitrarily badly off person, there is some arbitrarily worse off person for whom it would be better to achieve the very tiniest benefit," or "for any arbitrarily large population with an arbitrarily high per capita quality of life, there is some arbitrarily larger population with a barely positive per capita quality of life that would be better"). We can call these conclusions *Disturbing Conclusions*.

These Disturbing Conclusions involve large numbers, and recall that the Large Number Skeptics (4.3) argue that since we cannot relevantly imagine very large numbers, our intuitions about such conclusions should not be trusted, or should be trusted less. I offered a defense against the Large Number Skeptics. I argued that there is strong intuitive support for denying the Disturbing Conclusions, and I argued that this intuitive support does not depend on whether or not we can imagine very large numbers.

One might attempt to support the Skeptics, arguing that since Spectrum Arguments reveal a wide range of intuitively plausible Tradeoff Principles to have Disturbing Conclusions (which involve large numbers) and since either our intuitions which support these Principles are wrong or our intuitions which support the denials of these Disturbing Conclusions are wrong, it is our intuitions about the Disturbing Conclusions which must be jettisoned.

This, I believe, is a mistake. I have already argued that these intuitions about the Disturbing Conclusions cannot plausibly be discarded on the grounds that they are about large number cases. And the fact that the Disturbing Conclusions are about odd or less familiar cases does not, by itself, provide any reason to doubt our intuitions about them. Moreover, the fact that these Conclusions embarrass *many* Tradeoff Principles is, by itself, no defense against these Tradeoff Principles. For it is a clear possibility that *all* such Tradeoff Principles *are* indeed revealed to be implausible by Spectrum Arguments.

Instead, we might doubt the reliability of the *arguments* from these Tradeoff Principles to these Disturbing Conclusions (that is, we might doubt whether Spectrum Arguments are sound). We might claim that the Tradeoff Principles are so plausible, and the Disturbing Conclusions are so implausible, that there *must* be *some* problem with arguments from the former to the latter, even if we do not know what it is. We might count this as a *sixth* possible response to Spectrum Arguments: that they are unsound in some way which we have not yet been able to successfully detect.

Take an analogy. We often make claims about whether objects are heaps, about whether scalps are bald, and about whether people have beards. We are often very confident about these claims, e.g., Santa Claus *definitely* has a beard. But then we encounter *Sorites Arguments*, which themselves seem to make plausible assumptions (e.g., one hair cannot make the difference between a beard and a non-beard). These arguments have conclusions which are very hard to believe – for example, that one grain of sand is a heap, that very hairy scalps are bald, and that a completely hairless face has a beard. The claims about heaps, scalps (bald or not bald), and beards that we often make seem so plausible, and the conclusions of Sorites Argument seem so implausible, that we might sensibly claim that there *must* be *some* problem with these arguments, even if we cannot detect what it is. Indeed, this is how many philosophers who do not work on Sorites Arguments respond to them. They admit that they do not know *what* is wrong with these arguments, but they insist that there is nonetheless *something* wrong with them.

Still, many of us who do not specialize in vagueness *are* aware of a feature of

Sorites Arguments, which appears to render them unsound: their reiterative

application of a Soritical Premise to a vague predicate. It seems to us that this feature

exploits vagueness in an illegitimate way, and we thus have a *sense* of what is wrong

with Sorites Arguments. This, I believe, is part of the reason that we can plausibly

dismiss such arguments as unsound without knowing exactly how they are unsound

(i.e., without being experts on vagueness).

Things are different when we have *no* such particular reason for doubting the

soundness of an argument, other than that its conclusion seems implausible (despite

the plausibility of its premises). The general possibility that an argument is unsound is

not by itself a reason for doubting its soundness. Even if the five possible responses to

Spectrum Arguments listed above are not exhaustive, and a sixth *possibility* is that

Spectrum Arguments are unsound for some distinct and heretofore undetected reason,

it seems that we have no reason to believe this sixth possibility is a *plausible* one.

Indeed, we have excellent reason to *deny* that this sixth possibility is plausible: it

seems that, since Spectrum Arguments are *valid*, the *only* way to deny their

conclusions is to deny at least one of the Premises, or else deny Transitivity. (Indeed,

for this reason, the fifth solution is not truly independent of others – though it remains

useful to distinguish it and discuss it separately).

And insofar as we not only find the Disturbing Conclusions implausible, but

we also find it implausible to deny the Premises, and implausible to deny Transitivity,

it seems that *however* we respond to Spectrum Arguments, our response will be

implausible. What, then, can we do? The best we can do is look for the *least*

implausible solutions.

7.2   Least Implausible Solutions

Spectrum Arguments are complex, and not all such Arguments are created

equally. The least implausible solution varies from Spectrum Argument to Spectrum

Argument, and depends not only on the form or structure of the argument, but on its

specific content. The Devil is in the details. Below is a table, summarizing some

tentative claims about least implausible solutions.

7.2.1   Possible Solutions

Proposed solution	When it is the least implausible solution
1. Accept the conclusion	Sometimes; e.g., in response to Aggregating Spectrum Arguments (see 4.1)
2. Deny some but not all of the Premises	Sometimes; e.g., in response to certain Comparative Sorites Arguments (see 5.3.3)
3. Deny all of the Premises	Sometimes; e.g., in response to Weighting Spectrum Arguments (see chapters 3 and 4)
4. Deny Transitivity	Never; at least insofar as this allows cyclical betterness rankings within option sets (see 6.4 and 6.5)
5. Unsound for the reason Sorites Arguments are unsound	Never (see chapter 5)

Figure 7: Possible Solutions

For reasons offered in chapters 5 and 6, I do not believe that the fourth and fifth solutions are *ever* the least implausible solution to a Spectrum Argument. Spectrum Arguments *are* relevantly disanalogous to Sorites Arguments, at least if we are referring to standard Sorites Arguments, and not the Comparative ones (mentioned in 5.3.3). And the intuitive and theoretical costs of giving up Transitivity are too great (as I argued in 6.5, cyclical betterness claims within option sets are implausible).

Regarding the second solution: we could deny some but not all of the Premises of a Spectrum Argument in two ways. We could either deny the

Conditional, or we could deny Parity. It never seems to be the least implausible

solution to deny the Conditional. But it sometimes seems to be the least implausible

solution to deny Parity. Recall, from 5.3.3, the Comparative Sorites. It started with

the following Series:

        Collection 1:   100,000,000 grains

        Collection 2:   99,999,999 grains

        Collection 3:   99,999,998 grains

                  …and so on…

        Collection 99,999,998:   3 grains

        Collection 99,999,999:   2 grains

        Collection 100,000,000:   1 grain

And then argued:

        S1.   Collection 2 is no less a heap than Collection 1

        S2.   Collection 3 is no less a heap than Collection 2

          …and so on…

        Sn.   Collection n is no less a heap than Collection n-1

        ST.   "Is no less a heap than" is transitive

        ∴   Collection n is no less a heap than Collection 1

According to Parity, S1 through S*n* are (roughly) equally intuitively plausible. It might seem plausible to some of us that, when comparing two very large heaps of sand, the heap containing one fewer grain is *no less a heap* than the heap containing one more grain (we could defensibly say, "yes, this heap contains one fewer grain, but *that* doesn't make it *less a heap*"). At the same time, we might think that, when comparing two "medium-sized" heaps of sand, the heap containing one fewer grain *is* less a heap than the heap containing one more grain. That is, we might find the premises near S1 more plausible than the "middle" premises, or those near S*n*. (Similar remarks apply to the Second Comparative Sorites).

Some, however, might deny *each and every* premise of this Comparative Sorites (and the Second Comparative Sorites). If they did, they would be invoking the third solution. It might also be plausible to deny the transitivity of "is no less a heap than" or "is at least as much a heap as" where it would not be plausible to deny the transitivity of "is all things considered better than." Denying transitivity in the former case might not, for example, face an analog of the Implosion Argument (that argument relied on a connection between reasons and goodness, and there may be no analogous connection between reasons and *heapness*!). Let us now turn to Aggregating and Weighting Spectrum Arguments.

7.2.2   Aggregating versus Weighting

Recall that, in chapter 1 and more fully in chapter 4, we distinguished between two kinds of Spectrum Argument: *Aggregating* Spectrum Arguments and *Weighting* Spectrum Arguments. Both threaten to imply that certain Tradeoff Principles are implausible.

In the case of aggregation, these Principles make claims about how to tradeoff different normatively significant dimensions of benefits and burdens, e.g., their size and number, or their quality or intensity and duration. These normatively significant dimensions of benefits and burdens are, in a certain sense, intrinsic to benefits and burdens. That is, they are relevant to how good various collections of benefits are, and how bad various collections of burdens are, independently of any relations they might have to various possible beneficiaries. Aggregating Spectrum Arguments scale up a particular moral factor, *total well-being* (or *total ill-being*), in ways which threaten to embarrass views which recognize this factor as relevant and significant. Since any plausible view must recognize this factor as relevant and significant, Aggregating Spectrum Arguments threaten to embarrass all views worthy of our consideration.

In the case of weighting, on the other hand, the relevant Tradeoff Principles make claims about how to tradeoff total well-being (or some dimension of total well-being) against the priority that ought to be given to those worse off. The latter is, in a certain sense, extrinsic to benefits themselves. That is, combining the latter with the intrinsic facts about benefits, we are attempting to determine the *overall goodness* or *moral importance* of achieving these benefits relative to particular beneficiaries. Weighting Spectrum Arguments, rather than scaling up total well-being, scale up

priority weight or *equality weight* in ways which threaten to embarrass views which recognize these factors as relevant and significant. And whereas Aggregating Spectrum Arguments have received a lot philosophical attention, this dissertation (as far as I know) marks the first discussion and exploration of Weighting Spectrum Arguments.

What, then, are the least implausible solutions to the puzzles these arguments raise? In the case of both Aggregating and Weighting Spectrum Arguments, I believe that the second solution is too implausible to be the least implausible solution. The second solution is to deny some of the Premises but not others. This would require us to deny either Parity or the Conditional.

Denying the Conditional is, I believe, *too* implausible. It seems almost absurd to claim that, while we are justified in holding belief-1 on the grounds of its intuitive plausibility, we are *not* justified in holding belief-2 on the grounds of its intuitive plausibility, assuming that belief-2 is at least as intuitively plausible as belief-1, and other things are equal.

And denying Parity is also, I believe, *too* implausible. At least, it seems too implausible in the case of both Aggregating and Weighting Spectrum Arguments. Recall that Parity is the claim that the Series behind Aggregating and Weighting Spectrum Arguments can be constructed such that each of the Premises is (roughly) as intuitively plausible as any other. Each such Premise takes the form of a Tradeoff Principle: *a little less X and a lot more Y is better*. (E.g., "a slightly smaller benefit for a much worse off person is better," or "a slightly less intense pain that is much

longer is worse"). As long as the difference in the amount of X is arbitrarily slight, and the difference in the amount of Y is arbitrarily large, it seems that each of these Premises would be as intuitively plausible as any, and that the *location* of a Premise in these Spectrum Arguments (i.e., whether it is closer to the first Premise or the last one) does not, by itself, seem to make a difference to its intuitive plausibility. While we have not, and cannot, directly check every possible such Premise, the plausibility of Parity seems strongly supported by taking random, representative samples of Premises from different locations (see 3.3.2), and I am not aware of any convincing counterexamples to Parity.

This leaves us with just two possible solutions to Aggregating and Weighting Spectrum Arguments: accept their conclusions (the first solution), or deny all of their Premises (the third solution).

I believe, albeit somewhat tentatively, that the least implausible solution to Aggregating Spectrum Arguments is to accept their conclusions, but that the least implausible solution to Weighting Spectrum Arguments is to deny all of their Premises. This asymmetric pair of solutions might initially seem odd. We might have thought, for example, that if the first solution were the correct solution to Aggregating Spectrum Arguments, then it must also be the correct solution to Weighting Spectrum Arguments. This, I think, is mistaken. While these two kinds of Spectrum Argument are structurally analogous, there are important disanalogies in their content, or substance. There are at least the following two major differences.

(1) The denial of the Premises of Aggregating Spectrum Arguments seems more counterintuitive than does the denial of the Premises of Weighting Spectrum Arguments. Suppose, for example, that the following two claims were inconsistent, and that you *had* to deny one of them: **(WP)** a slightly smaller benefit for a much worse off person is better than a slightly larger benefit for a much better off person, other things being equal, and **(AP)** many slightly smaller benefits is better than one slightly larger benefit, other things being equal. It might seem crazy to deny either claim, but which does it seem *crazier* to deny? It seems to me that the answer is AP. ("WP" stands for "Weighting Premises," and "AP" stands for "Aggregating Premises").

If WP were false, that might be because the Equal Weight View is true. But for AP to be false, it would have to be because the following view is true:

> *Lexical Size View*: the *size* of benefits is always lexically better than the *number* of benefits. That is, *no* loss in benefit size, however small, could be compensated by *any* gain number of benefits, however great.

But it seems that the Equal Weight View is less implausible than this Lexical Size View. The Lexical Size View is *absurd*. The Equal Weight View, though it has implausible implications, is by no means absurd. Indeed, it seems to be a defensible option.

Here is the second major difference between Aggregating and Weighting Spectrum Arguments:

(2) The acceptance of the conclusions of Weighting Spectrum Arguments seems more counterintuitive than does the acceptance of the conclusions of Aggregating Spectrum Arguments. Suppose, for example, that you could not consistently deny both of the following two claims, and that you *had* to accept one of them: **(WC)** for any arbitrarily large benefit for any arbitrarily badly off person, there is some arbitrarily worse off person for whom it would be better to achieve the very tiniest benefit, other things being equal, and **(AC)** for any arbitrarily large benefit we could achieve (e.g., prevention of one excruciatingly painful two-year-long torture session), there is some arbitrarily large number of very tiny benefits (e.g., prevention of many mildly annoying hangnail pains) that it would be better to achieve, other things being equal. It might seem crazy to accept either claim, but which does it seem *crazier* to accept? It seems to me that the answer is WC. ("WC" stands for "Weighting Conclusion," and "AC" stands for "Aggregating Conclusion").

If WC were true, that would have to be because some fact extrinsic to the very tiniest benefit – namely, the fact that its recipient would be *so* badly off – could so dramatically magnify its moral importance, that this fact would make this tiny benefit more morally important than any benefit, however large, for a badly off person. There would have to be *no limit* on how morally

important such an extrinsic fact could make this very tiniest benefit. I do not see how this could be true.[154] On the other hand, if AC were true, that would have to be because *enough* of the tiniest benefits could always add up to be as good as any number of arbitrarily large benefits. Though this seems implausible, I at least think I do see how this could be true. At least, it does not seem quite as unbelievable as WC.

Let us assume, for the sake of argument, I am right about major differences (1) and (2). A *weak* version of (1) is merely a claim about the comparative plausibility of Not-WP and Not-AP, and a *weak* version of (2) is merely a claim about the comparative plausibility of WC and AC. These claims are, respectively, that Not-WP is more plausible than Not-AP, and that AC is more plausible than WC. These weak versions of (1) and (2), if true, would establish two important disanalogies between Weighting and Aggregating Spectrum Arguments, but they would not establish that the first solution is the least implausible solution to Aggregating Spectrum Arguments, or that the third solution is the least implausible solution to Weighting Spectrum

---

[154] Indeed, WC is even more implausible than I suggested above. Above I spelled out WC as Priority Monster. But recall *Scarier Priority Monster* from Chapter 4. Scarier Priority Monster is the claim that there could be a person so badly off that, if she existed, it would be better to achieve an arbitrarily small benefit for her than to achieve arbitrarily many, arbitrarily large benefits for arbitrarily many people each as badly off as the currently worst off person. But surely that is absurd, and far more implausible than any AC.

For a more concrete version of this example, suppose there were millions of very badly off people slowly dying in London, and that I have some pills I could distribute to them if I hopped on the next flight to Heathrow. These pills would cure their illnesses, and give each of them an additional 100 years of extremely high quality life. Or, I could hop in a space shuttle, fly at the speed of light to some distant galaxy in order to find someone who has been alive at a low quality of life for a very long time (and is therefore a priority monster), and give this person an aspirin, or a lollipop lick – some piddling benefit. Suppose these are the only two acts I could perform. It strikes me as absurd that I should here trivially help a priority monster, and this seems to me considerably more implausible than Hangnails for Torture, Utility Monster, Repugnant Conclusion, and the like.

Arguments. Indeed, these weak versions of (1) and (2) are compatible with each of

the following possibilities:

(i)   AC is more plausible than WC is more plausible than Not-WP is more

        plausible than Not-AP. Here we would accept the conclusions of both

        Aggregating and Weighting Spectrum Arguments as the least

        implausible solutions. (We would accept the first solution for both).

Or:

(ii)   AC is more plausible than Not-WP (or Not-WP is more plausible than

        AC) is more plausible than WC is more plausible than Not-AP (or WC

        is more plausible than Not-AP). Here we would accept the conclusions

        of Aggregating Spectrum Arguments, but deny all the Premises of

        Weighting Spectrum Arguments, as the least implausible solutions.

        (We would accept the first solution for the former, and accept the third

        solution for the latter).

Or:

(iii)   Not-WP is more plausible than Not-AP is more plausible than AC is

        more plausible than WC. Here we would deny all the Premises of both

Aggregating and Weighting Spectrum Arguments as the least

implausible solutions. (We would accept the third solution for both).

But there are *stronger* versions of (1) and (2) which do not *merely* make the

comparative claims that AC is more plausible than WC, and that Not-WP is more

plausible than Not-AP. In addition, these stronger versions make the following non-

comparative claims: WC is *unbelievable*, period, and Not-AP is *unbelievable*, period.

Thus, insofar as AC is not unbelievable and Not-WP is not unbelievable, we can

conclude that Not-WP is more plausible than WC, and that AC is more plausible than

Not-AP. That is, we can conclude that the least implausible solution to Weighting

Spectrum Arguments is the third solution, and that the least implausible solution to

Aggregating Spectrum Arguments is the first solution. We could also arrive at this

conclusion, if we accepted the *moderate* versions of (1) and (2), which do not make

the above non-comparative claims, but which simply directly claim that Not-WP is

more plausible than WC, and that AC is more plausible than Not-AP (in addition to

claiming that AC is more plausible than WC, and that Not-WP is more plausible than

Not-AP). I believe that *at least* the moderate versions of (1) and (2) are true.

I admit that, in making the above claims in favor of the first solution for

Aggregating Spectrum Arguments and the third solution for Weighting Spectrum

Arguments, I am appealing to my own considered intuitions. I am reporting which

claims seem to me less implausible. While many people do and will share these

intuitions, some others do and will have different intuitions. To make progress in such

cases of disagreement, we may have to go significantly deeper into intuitionist moral epistemology, and into the epistemology of disagreement.[155] Regrettably, I cannot do that here.

7.2.3 Reconsidering Utilitarianism

Suppose that, as I tentatively suspect, we should accept the conclusions of Aggregating Spectrum Arguments, but deny all the Premises of Weighting Spectrum Arguments. If this is right, then we should reconsider Utilitarianism, a view which many have rejected for all of its implausible implications. We should reconsider whether Utilitarianism, though a view with many implausible implications, is nonetheless the *least implausible* option. In asking this, I mean to ask about a suitably scope-restricted Utilitarianism. Thus, we are not reconsidering whether:

> *Utilitarianism*: an act is permissible if and only if it promotes at least as much total well-being as any alternative act.

is the least implausible option. But rather whether:

---

[155] For instance, see Christensen 2007. Also see Crisp's discussion of intuitionism and disagreement in 2006b, 88-97.

Outcome Utilitarianism: an outcome is all things considered better only insofar as it contains more total well-being.[156]

is the least implausible option. Utilitarianism is a general normative theory. Outcome Utilitarianism, by contrast, is not. It only ranks outcomes. It avoids many of the implausible implications that Utilitarianism faces, e.g., that it is impermissible *not* to harvest an innocent, non-consenting person's vital organs to save five others. Outcome Utilitarianism follows from the following three claims:[157]

- *Equal Weight View*: when it comes to distributing benefits, only their size and number matters, facts about the possible recipients of these benefits do not ultimately matter,

- *Total View*: other things equal, the outcome with more total (weighted) well-being is better, and

- *Welfarism*: outcomes are better only insofar as they contain a better distribution of well-being.

---

[156] This definition of Outcome Utilitarianism is equivalent to Sen's (1979) definition. However, whereas Sen factors Outcome Utilitarianism into the *Unweighted Total View* and Welfarism, I factor it into the Equal Weight View, the Total View (which is agnostic on weighting), and Welfarism.

[157] Sen factors Outcome Utilitarianism into two components, Welfarism and *Sum-Ranking*, where the latter is equivalent to an *Unweighted Total View*. I factor this second component further, into the Equal Weight View (which is agnostic on aggregating) and the Total View (which is agnostic on weighting).

Each of these claims has implausible implications. The Equal Weight View has implausible implications in the following cases, as noted in chapters 2 and 3: Lollipop, City or Suburbs, Two Children, and Divided World. The Total View has implausible implications in the following cases, as noted in chapter 4: Utility Monster, Repugnant Conclusion, and Hangnails for Torture. While I believe that Welfarism is plausible, some have claimed that it too faces implausible implications.[158] But suppose Welfarism is true. The defender of Outcome Utilitarianism might now argue:

> While it is true that the Equal Weight View has implausible implications, it is the least implausible option available (the third solution to Weighting Spectrum Arguments), and while it is true that the Total View has implausible implications, it is the least implausible option available (the first solution to Aggregating Spectrum Arguments). It is a regrettable fact that Outcome Utilitarianism is implausible, but it turns out that *all* available views are also implausible. And Outcome Utilitarianism is, of those views available, the least implausible. So, given that we are justified in accepting the least implausible view available, we are justified in accepting Outcome Utilitarianism.

---

[158] For example, we could compare two outcomes which contain equally good distributions of well-being, but imagine that one contains much more beauty than the other. We might, with G. E. Moore, think that the one with more beauty is better (1903, 83-85). Other putative Non-Welfarist goods might include humanity's survival, and its scientific and moral progress (cf. Parfit 1984, 454).

This argument deserves serious consideration. If my claims in 7.2.2 are correct, and if Welfarism is true, then it seems that we are indeed justified in accepting Outcome Utilitarianism.[159]

There are, however, some important rivals to the Equal Weight View which I have not been able to address in this dissertation. For instance, there are views which place greater weight on achieving benefits for persons over sentient non-persons, and there are views which place greater weight on achieving benefits for the more virtuous over the less virtuous (or the vicious).[160] I believe that these distributive principles face at least some analogs of the Weighting Spectrum Arguments I have discussed here. For example, both seem to threaten to imply, via Spectrum Arguments, conclusions analogous to Scary Pooled Priority (4.2.2). But these distributive principles raise many complications, which regrettably I cannot address here.[161]

7.3 Revisionary Intuitionism

---

[159] This is slightly inaccurate. Even if we should accept the conclusions to Aggregating Spectrum Arguments as the least implausible solutions, it would not follow that accepting the Total View is the least implausible solution. If we accept the Repugnant Conclusion, for example, we only claim that there is *some* size of population Z such that it is better than population A. We need not thereby claim that population Z is better than population A if and only if it contains more total well-being.

[160] See the footnotes in chapter 1 for some of the proponents and critics of these views.

[161] Also recall that I have largely confined the discussion here to Same-People Choices, or choices which affect neither the identities nor the number of future people.

As noted in chapter 1, I accept *Normative Intuitionism*. This is the view that, if we have the intuition that some normative proposition is true, we thereby have reason to believe this proposition, absent sufficiently strong defeaters.

When we have more reason overall to believe a normative proposition than not, we are justified in believing that proposition. Intuitions can fail to justify beliefs, or be defeated, in two ways. First, they can be rebutted. When the intuition that P faces a sufficiently strong *rebutting defeater*, though the intuition that P provides some reason to believe P, it is outweighed by a countervailing reason not to believe P. Second, intuitions can be undercut. When the intuition that P faces a sufficiently strong *undercutting defeater*, this intuition fails to provide any reason to believe P.

We might have thought that, if we accept Normative Intuitionism, this leads fairly quickly to *Common Sense Morality*, or the set of normative beliefs that most people accept.[162] Here are some intuitions that are fairly widely shared. It seems that lying is prima facie wrong. It seems that promise-breaking is prima facie wrong. It seems that the guilty should be punished, and that the innocent should not be. It seems good and just to promote equality. It seems prima facie wrong to restrict the liberty of consenting persons. It seems that we have prima facie duties to help others. It seems that we are not required to sacrifice our personal projects to help others. It seems prima facie wrong to steal. It seems that the good and hard-working deserve to be rewarded more, and that the evil and lazy deserve to be rewarded less (if at all). It seems that it is prima facie wrong to harm others, even if doing so will prevent

---

[162] For example, Normative Intuitionism led both Prichard (1957, 16) and Ross (1988, 21-2) to Common Sense Morality.

considerably greater harms to others. It seems that there are many things which make life worth living, in addition to pleasure and desire-satisfaction. It seems that we have special obligations to those near and dear to us that we do not have to strangers.

If each of the above intuitions is enshrined in a corresponding stand-alone principle, or belief, the resulting set of beliefs is, roughly, Common Sense Morality. We can call a normative theory which significantly differs from Common Sense Morality a *revisionary* normative theory. Utilitarianism, for example, is a highly revisionary normative theory in that it denies many or most of the beliefs of Common Sense Morality.

It is possible, however, to accept both Normative Intuitionism and a revisionary normative theory. I will call such a pairing of views *Revisionary Intuitionism*.[163] A standard version of Revisionary Intuitionism *weights* some intuitions more heavily than others, where the weights intuitions receive need not correspond to their *strength* (how strongly the propositions in question seem to be true). And an intuition might receive more or less weight depending on various factors. To a first approximation, this is uncontroversial. Most agree with John Rawls that we should distinguish between intuitions rendered "under conditions favorable for deliberation and judgment in general"[164] and those rendered under unfavorable conditions. For instance, Rawls notes that our capacity for rational moral judgment tends to be distorted or unreliable in circumstances in which we are hesitant,

---

[163] Following Huemer 2008a.

[164] Rawls 1971, 48.

frightened, or excessively self-interested. Giving less weight to intuitions formed under these particularly unfavorable conditions will probably not, however, result in a revisionary normative theory.

But there are more extreme proposals about weighting intuitions. Some have argued that intuitions about more general or abstract propositions (e.g., "the right act has the best consequences") should be given greater weight than intuitions about concrete cases (or the "mid-level principles" I listed above as the principles of Common Sense Morality).[165] Many have argued just the reverse.[166] Others have argued that there are several particular classes of intuitions which are subject to powerful undercutting defeaters, such as intuitions about deontological constraints on harming others,[167] or intuitions about special obligations to the near and dear.[168] If these proposals were right, we could be both Normative Intuitionists and defenders of revisionary normative theories. We could say, for example, that our intuitions *would* support Common Sense Morality, *if* they all received roughly equal weight. But, some of our intuitions should receive more weight than many others, and many of our intuitions face strong undercutting defeaters, and so should be given little or no weight at all. If enough intuitions are weeded out of our initial stock of intuitions, the

---

[165] See the footnotes in chapter 1 for some of the proponents and of these views.

[166] See, for example, Dancy 1993, and Kamm 2007, 5.

[167] For example, much of Joshua Greene's work in cognitive neuroscience is an attempt to reveal that deontological intuitions are misguided. For a critical response to Greene, see Berker 2009. And for a response to Berker, see Greene's "Notes on 'The Normative Insignificance of Neuroscience' by Selim Berker" (in progress).

[168] Again see Huemer 2008a and Crisp 2006b, 142-4.

surviving intuitions will, on balance, support a normative theory other than Common Sense Morality.

   Some of the above specific proposals are more plausible than others. For example, I am skeptical that, in general, intuitions about concrete cases should receive less weight than intuitions about abstract principles. But the broad project of searching for various possible undercutting defeaters of our intuitions seems extremely worthwhile, and it seems to be a way to make genuine progress in ethics.

   There is another version of Revisionary Intuitionism which could be pursued in conjunction with, or separately from, the above version which assigns significantly different weights to different kinds of intuitions. Rather than weighting intuitions differently, we might simply search for *more* intuitions than those included in the average person's stock of intuitions which, if weighted equally, would result in Common Sense Morality. The average person's stock of intuitions is, in part, the result of considering a particular range of familiar cases and principles. However, there are many unfamiliar cases which we do not ordinarily consider. Our intuitions about these unfamiliar cases, then, add to our stock of intuitions. If, as I believe is often though not always true, our intuitions about these unfamiliar cases should be given significant weight, then they may yield strong rebutting defeaters of many of our commonsense beliefs. If enough of our intuitions about familiar cases and principles are relevantly rebutted by intuitions about unfamiliar cases and principles, our expanded stock of intuitions will, on balance, support a normative theory other than Common Sense Morality. This is another important way to make genuine progress in

ethics. I hope, in this dissertation, to have provided but one example of how such

progress might occur.

Appendix A: The Separateness of Persons Argument for RM

Consider a variant of City or Suburbs:

Two Children. Two children will develop a mobility-affecting condition. One will suffer from *Slight Impairment*, "a condition that renders it difficult to walk more than 2km." The other child will suffer from *Very Severe Impairment*, "a condition that leaves one bedridden, save for the fact that one will be able to sit in a chair and be moved around in a wheelchair for part of the day if assisted by others."[169] You are a morally motivated stranger with the ability to provide a preventative treatment for one but not both of these children. The treatment for the child who will develop the Slight Impairment would prevent any mobility impairment whatsoever. The treatment for the child who will develop the Very Severe Impairment would prevent *Very* Severe Impairment, but leave the child with *Severe Impairment*, "a condition in which one is no longer bedridden; rather, one is able to sit up on one's own for the entire day but requires the assistance of others to move about." Assume that the treatment for the Slight Impairment would induce a *slightly larger benefit* than the treatment

---

[169] Ibid., 171-4.

for the Very Severe Impairment would.[170] Other things are equal (e.g., there are no considerations of desert in play).

Many judge that you ought to provide the treatment for the Very Severe Impairment, even though this treatment would induce a slightly smaller benefit than the benefit that would be induced by the treatment for the Slight Impairment. This is because the recipient of the former, slightly smaller benefit would have been much worse off than the recipient of the latter, slightly larger benefit would have been. Both RM and AM can capture the judgment that this is what you ought to do.

Michael Otsuka and Alex Voorhoeve[171] maintain RM. Their argument for this claim begins with an examination of the following sort of case:

> *One Child*.[172] A child will develop a mobility-affecting condition. There is a 50 percent chance that she will suffer from Slight Impairment, and a 50 percent chance that she will suffer from Very Severe Impairment. You are a morally motivated stranger with the ability to provide a preventative treatment for one

---

[170] In the original version of the case, these treatments would bring about same-sized benefits. It is perhaps natural to expect that the treatment for the Very Severe Impairment would induce a larger benefit than the treatment for the Slight Impairment would. But we can imagine that this is not the case in this example; we can tell many scenarios in which the treatment for the Slight Impairment would result in a larger well-being gain.

[171] Ibid.

[172] This case differs from the one Otsuka and Voorhoeve initially offered in that it involves a child, rather than a young adult. A reason for favoring an example involving a child rather than a young adult is that it helps abstract from considerations of respect for autonomy.

but not both of these conditions. You have to decide now whether to administer the treatment for the Slight Impairment or instead for the Very Severe Impairment. As in Two Children, the treatment for the Slight Impairment would bring about a slightly larger benefit than the treatment for the Very Severe Impairment would. Other things are equal (e.g., there are no considerations of respect for autonomy in play).

AM deems it more important to achieve a benefit for someone with the Very Severe Impairment than to achieve a slightly larger benefit for someone with the Slight Impairment. In a case like One Child, AM would plausibly follow the orthodox theory of decision making under risk, and thereby recommend maximizing *expected moral value*.[173] Since a 50 percent chance of a more important benefit (treatment for Very Severe Impairment) is better than a 50 percent chance of a less important benefit (treatment for Slight Impairment), AM would imply that you ought to administer the treatment for the Very Severe Impairment.[174]

---

[173] Otsuka and Voorhoeve (2009, 195-6) plausibly assume that the Priority View (the AM version) first ranks the moral value of outcomes, and then recommends maximizing expected moral value under risk; they note that Parfit would agree. I shall follow them in making these assumptions. However, it is worth noting that one could defend the Priority View as a theory about how to rank the moral value of outcomes, but leave it open how to deal with risk. One might believe, for example, that maximizing expected moral value is justified only as a strategy for dealing with large numbers of gambles. Following this strategy will, in the long run, maximize *actual* moral value. But one might have doubts about the wisdom of maximizing expected moral value in "one shot" gambles, where there is no chance of recouping one's losses if one loses. I am grateful to Dick Arneson for discussing this possibility with me. For an explanation and defense of the axioms of expected value theory, see chapter 5 of Broome 1991.

[174] If AM is true, then it is true in cases like One Child. It would be implausible to claim that whether it morally matters that some are worse off in Absolute terms is dependent on how many people exist, or how many people we can benefit. If it matters generally, then it matters in One Child. This point was first made by Rabinowicz 2002, 17. Also see Otsuka and Voorhoeve 2009, 188-190.

In One Child no one will be worse off relative to anyone else, since there is only one person involved (at least, we could bracket others). Those who accept RM but reject AM could consistently favor the treatment for the Slight Impairment in One Child, but favor the treatment for the Very Severe Impairment in Two Children. Those who believe *both* RM and AM would favor the treatment for the Very Severe Impairment in both cases, but would recognize an additional reason for favoring this treatment in Two Children (since in this case but not in One Child someone is worse off relative to someone else). But those who deny that it matters how well off people are in relative terms recognize no shift or asymmetry between the one-person case and the analogous multi-person case; according to them, the importance of providing the treatment for the Very Severe Impairment depends only on the size of the benefit this treatment induces, and how well off the recipient of the benefit would have been without it.

Otsuka and Voorhoeve seek to illustrate the truth of RM by showing that there is, as RM would imply, a shift or asymmetry between the one-person case and the analogous multi-person case. Presumably, their defense of the claim that there is such a shift should not simply presuppose RM.

Their defense appeals to the moral significance of the *separateness of persons*. They claim that a single person "has a unity" that a group of different people do not have, and that this unity renders it easier to justify balancing or trading-off (expected) benefits and (expected) burdens intrapersonally than interpersonally.[175] If in One

---

[175] Ibid., 179. Also see: Rawls 1971, 27-33, Parfit 1984, 329-347, and Brink 1993.

Child you decided to administer the treatment for the Slight Impairment, and the child developed the Very Severe Impairment, Otsuka and Voorhoeve point out that you "can justify [your] decision to [administer the treatment for the Slight Impairment] on the grounds that [you] were looking after that *very same* child's interest in flourishing in the event that she had turned out [to have the Slight Impairment]."[176] This justification – call it the *Intrapersonal Compensation Justification* – is obviously unavailable in the Two Children case, which involves an *inter*personal trade-off. Otsuka and Voorhoeve believe that the Intrapersonal Compensation Justification renders it reasonable to favor the treatment for the Slight Impairment in One Child.[177]

The argument appealing to the Intrapersonal Compensation Justification, so far, would seem to imply not only that there is a shift or asymmetry between One Child and Two Children (as RM would imply), but also that AM is *false* in One Child. Defenders of AM might find this implausible. Indeed, Roger Crisp has pointed out that the child might well acknowledge that the treatment for the Slight Impairment offered her a larger expected benefit, but complain that you should have paid attention to the fact that she had two possible futures, one in which she developed the Very Severe Impairment and one in which she only developed the Slight Impairment. She could add, "You should then have asked yourself which possible future person should

---

[176] Otsuka and Voorhoeve 2009, 188 (italics mine).

[177] Again, their original case involved same-sized benefits; accordingly, their claim about that case was that the Intrapersonal Compensation Justification rendered it reasonable to be *indifferent* between the two treatments.

be given priority to the other – and the answer would have been clear: the one who is worse off."[178]

But Otsuka and Voorhoeve, to defend RM with the above argument, need not deny AM. They might argue as follows:

(a) Even if *all things considered* you should provide the treatment for the Very Severe Impairment in both cases (as Crisp claims), there is at least *a* reason for favoring the treatment for the Slight Impairment in One Child that is absent in Two Children, since you can offer the Intrapersonal Compensation Justification in the former but not the latter.

(b) The presence of a reason for favoring the treatment for the Slight Impairment in One Child that is absent in Two Children supports:

> *The Shift*: however much weight it is appropriate to give to the interests of someone who turns out badly off in one-person cases, it is appropriate to give *more* weight to the interests of badly off persons in analogous multi-person cases.[179]

---

[178] Crisp 2011, 107.

[179] See Otsuka and Voorhoeve 2011, 111.

(c) The Shift explains why RM is true.

But this, I argue, is a questionable argument for RM.

Consider another case:

Three Children. One child, Anna, has a 50 percent chance of developing the Slight Impairment and a 50 percent chance of developing the Very Severe Impairment. Let Anna-1 refer to her life under Slight Impairment, and let Anna-2 refer to her life under Very Severe Impairment. Another child, Becca, will develop the Slight Impairment. There is a 50 percent chance that her condition is untreatable. And another child, Calla, will develop the Very Severe Impairment. There is a 50 percent chance that her condition is untreatable. You are a morally motivated stranger with the ability to provide only one preventative treatment. You have to decide now whether to administer the treatment for Anna-1, Anna-2, Becca, or Calla (not now knowing whether Anna-1 or Anna-2 will exist, and not now knowing whether Becca and Calla have untreatable conditions). These treatments, if successful, would bring about the same benefits as in previous cases. Other things are equal.

A diagram representing the four options in *Three Children*:
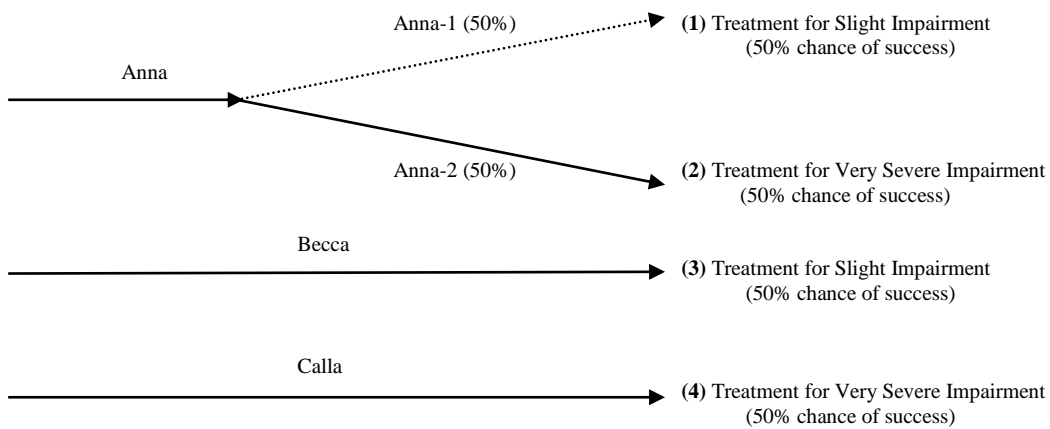
Figure 8: Three Children

I assume that Anna would acquire her mobility-affecting condition at the same time as Becca and Calla. In terms of their overall well-being scores (prior to receiving any of the treatments in question), the Anna-1 life and the Becca life are equivalent, and the Anna-2 life and the Calla life are equivalent.

Now observe that just as in One Child you could justify your choice to administer the treatment for the Slight Impairment to the child if she ended up with the Very Severe Impairment on the grounds that she is identical to the person who would have had the Slight Impairment, you could offer the same justification in defense of choosing (1) over (2) in Three Children.

Insofar as Otsuka and Voorhoeve are right that the availability of the Intrapersonal Compensation Justification in One Child provides a reason to favor the treatment for the Slight Impairment, it seems plausible that since one could offer the

Intrapersonal Compensation Justification to Anna in Three Children there is a reason in favor of (1) over (2).

And insofar as Otsuka and Voorhoeve are right that the presence of a reason to provide the treatment for the Slight Impairment (in One Child) that is absent between two other treatments (in Two Children) supports The Shift, it seems plausible that since such a reason is present between (1) and (2) but absent between (3) and (4), it is appropriate to give *more* weight to the interests of Calla than to the interests of Anna if she were to end up with the Anna-2 life. This implies that (4) is more morally important than (2).

Note that the Anna-2 life and the Calla life suffer from the very same impairment, and these treatments would, if successful, bring about the very same improvements. Consider two diagrams:

Treatment (2):   Utility

50 Percent Chance, Success   50 Percent Chance, Unnecessary

Anna-2   Becca   Calla   Anna-1   Becca   Calla

Treatment (4):   Utility

50 Percent Chance, Success   50 Percent Chance, Untreatable

| Anna1 | | | | Anna1 | |
| Anna2 | | | | Anna2 | |

Anna   Becca   Calla   Anna   Becca   Calla
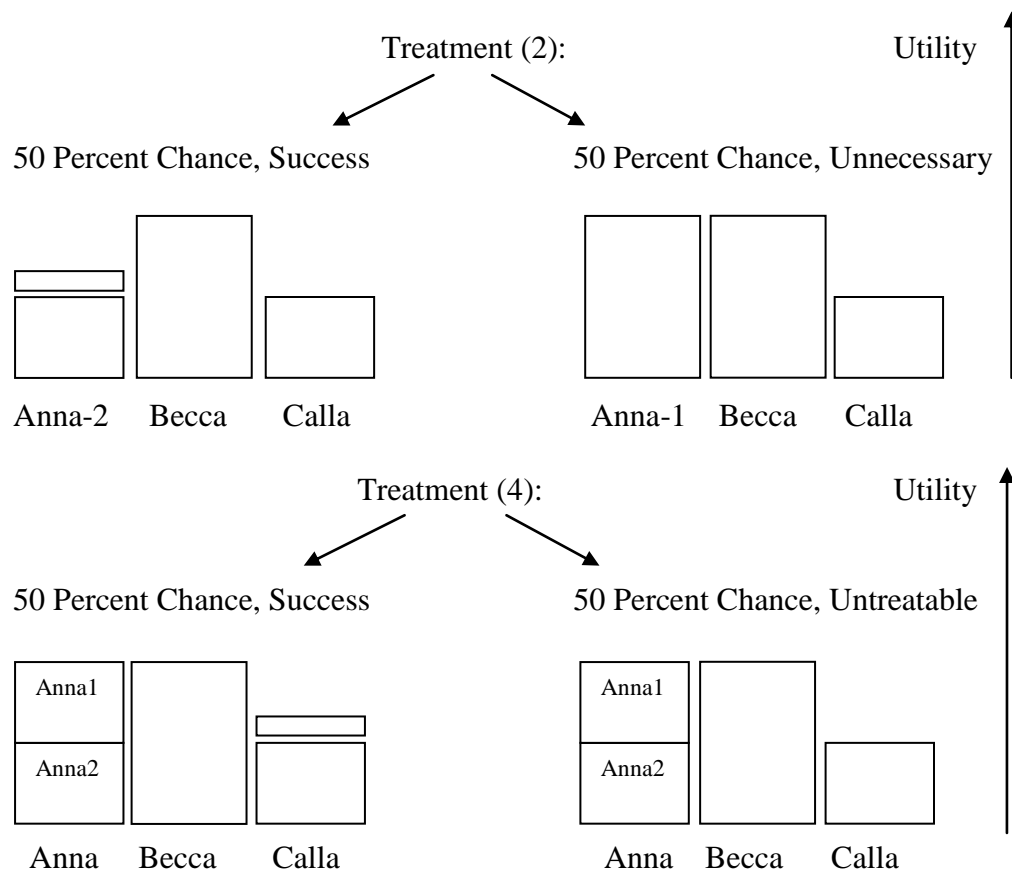
Figure 9: Two Treatments

In these diagrams, the boxes represent each person's utility level, and the thin

horizontal boxes represent the benefits induced by a successful treatment. For

treatment (4), there are four possibilities, depending both on whether Calla's condition

is treatable and on whether Anna ends up living the Anna-1 life or the Anna-2 life.

As the diagrams illustrate, treatment (2), if successful, would benefit those

worse off relative to others at least as effectively as treatment (4), if successful, would.

And so, *according to RM*, (2) is at least as important as (4). But the argument Otsuka

and Voorhoeve offer for The Shift, which was intended to illustrate or reflect the truth

of RM, apparently implies otherwise – since their argument implies that there is more reason to provide (4) than (2).

Otsuka and Voorhoeve might deny that their argument for The Shift implies that (4) is more morally important than (2). But if the Intrapersonal Compensation Justification provides a reason in favor of the treatment for the Slight Impairment in One Child, then *why* does it *not* provide a reason in favor of (1) over (2)? Any explanation of the insignificance of the Intrapersonal Compensation Justification in Three Children would seem just as good an explanation of its insignificance in One Child. And claiming that the Intrapersonal Compensation Justification is insignificant in Three Children *because* it is a multi-person case would be circular; it is the *significance of the Intrapersonal Compensation Justification* that is supposed to *explain* the moral difference between one-person cases and analogous multi-person cases, not the other way around.

Alternatively, Otsuka and Voorhoeve might simply accept the conclusion I think their argument implies in Three Children – that treatment (4) is more morally important than treatment (2). They might believe that this is both plausible and consistent with RM.

They might believe, for example, that RM should not just be concerned with how well off people *end up* relative to others, but also with how good people's *prospects* are relative to others. If so, then Calla, who has a 100 percent chance of having the Very Severe Impairment, is clearly worse off in prospect than Anna, who

only has a 50 percent chance of having the Very Severe Impairment. For these reasons, one might think that treatment (4) *is* more important than treatment (2).

But then, to fairly compare AM with RM, we should consider a version of the former that is not just concerned with how well off people end up, but also how good their prospects are. When so amended, AM could likewise imply that (4) is more morally important than (2). Thus even if this conclusion about Three Children *were* plausible, it would not favor RM over its denial.

There is further evidence that the argument for The Shift does not reflect the truth of RM. Consider a pair of cases:[180]

500 Lottos. 500 children are in exactly the same predicament as the child in One Child. There are 500 *separate* lotteries, one for each child, that determine which impairment each child ends up with. A likely result is 250 children with the Very Severe Impairment and 250 with the Slight Impairment.

Single Lotto for 500. Same case as 500 Lottos, except there is a *single* lottery that determines which impairment *all* 500 will end up with.

Suppose again that, in these cases, you can provide only one of the two treatments for each child. Each child faces the same prospect, whichever case she is in. But it is nearly certain that in 500 Lottos some will *end up* worse off relative to

[180] Crisp presents *500 Lottos* in Crisp 2011, 107-8, and Otsuka and Voorhoeve present *Single Lotto for 500* in Otsuka and Voorhoeve 2011, 111-4.

others, whereas there is no chance of this in Single Lotto for 500. Therefore RM would imply that there is an *extra* reason to provide the treatment for the Very Severe Impairment in 500 Lottos that is absent in Single Lotto for 500. But since the Intrapersonal Compensation Justification is equally available in both cases, the argument Otsuka and Voorhoeve offer in defense of The Shift cannot explain the difference between the cases which is implied by RM. Moreover, this fact suggests that their argument gets the right result for the wrong reason in cases like One Child. In that case, two distinct types of reason happen to pull in the same direction:

(1) An extra reason to benefit those worse off relative to others (whether in outcome or in prospect).

(2) A reason to maximize (expected) benefits intrapersonally.

As we have seen, these two types of reason are distinct, and need not always pull in the same direction. RM entails (1). But I hope to have shown that (2) is not the same as, and does not entail, (1). The argument Otsuka and Voorhoeve offer for RM, based on The Shift and justification the Intrapersonal Compensation Justification, therefore appears dubious.[181]

We might nonetheless ask whether (2) is plausible, and what bearing that could have on the debate between RM and AM. Moreover, we might consider an even

---

[181] For a reply to Otsuka and Voorhoeve similar to mine, which I learned about after developing my own, see Parfit 2012.

stronger claim than (2) – stronger in that it is refers to what we *ought* to do, and not merely what we have some reason to do. I will call it the:

> *Rational Prudence Constraint*: we *ought not* to treat a person in ways in which she could not prudentially rationally choose to be treated, if it were only the well-being or interests of this person our act could affect, and other things were equal.[182]

If it is prudentially rational for one to choose, under risk, to maximize one's expected benefits, then the Rational Prudence Constraint will conflict with AM. But I will return to this in Appendix C. There is a pressing question that I would like to address first, in Appendix B. There I will consider a question which is in some respects the mirror image of the question I have just been considering. I have just been considering, independently of whether or not we accept AM, whether we must *also* accept RM. I will now ask: *insofar as* we believe we should give priority to the worse off – and independently of whether we accept Egalitarianism or RM – must we *also* accept AM? I will present a series of cases which suggest the answer is "yes."

---

[182] The "other things equal" clause should not be read as abstracting from considerations of Equality or Priority, otherwise the Rational Prudence Constraint could never conflict with the Equality View or the Priority View. But I am supposing that it could, and moreover that this fact might provide some reason for thinking these views are implausible.

Appendix B: Should Defenders of RM Also Accept AM?

Suppose we have the belief, with Egalitarians and Prioritarians (and with defenders of AM and RM alike), that it is better to achieve a slightly smaller benefit for a considerably worse off person than it is to achieve a slightly larger benefit for a considerably better off person. I will now present some cases which elicit beliefs that seem *no less plausible* than this belief. It turns out that, to capture these beliefs about these cases, we must appeal to AM. Consider:

Only One of Two Will Exist. No one exists, nor has existed, nor will exist,[183] other than either Adele or Belda. It is a fact that either Adele or Belda will exist, but only one of them will. There is a 50 percent chance that Adele will exist, and a 50 percent chance that Belda will exist. If Adele existed, she would live a blissful life, filled with a great abundance of that which makes life worth living. If Belda existed, her life would be barely worth living. We can bring it about that, if Adele exists, she will receive a large benefit, or that, if Belda exists, she will receive an equally large benefit. We have to decide what to do now. Other things are equal.

---

[183] These qualifications about time are important because some Egalitarians (e.g., Larry Temkin) – as well as non-Egalitarian defenders of RM – might believe that it matters how well off people are relative not just to persons who exist now, but also to persons who did or will exist.

The belief that a benefit to Belda has greater moral importance than an equal benefit to Adele strikes me as no less plausible than the most plausible Egalitarian and Prioritarian beliefs. If this is right, it seems uncontroversial that we should bring about the 50 percent chance that Belda will benefit rather than the 50 percent chance that Adele will. We cannot capture this belief by appealing to RM, since Belda would be worse off than no one (a person can only be worse off in relative terms if she is worse off *than someone else*). It seems we must therefore appeal to AM.[184]

First, one might object to *Only One of Two Will Exist* as follows: "if *our* act could have an effect on either Adele or Belda, we must exist. But if we exist, then some people besides Adele and Belda exist." In response, we could suppose that "we" are suitably hypothetical spectators. Alternatively, we could suppose that we are real, but bracket facts about how well off we are. *Alternatively*, we might simply ask about the moral value of benefits *befalling* Adele or Belda, without the intervention of any agent.

---

[184] It is worth mentioning that, in applying AM to this case, we are *not* thereby including *persons whose existence is contingent on what we do*, or *persons who certainly will not exist unless we choose make them exist*, among those persons whose interests should receive greater priority the worse off they are. One person *will* exist in my example, whatever we do, but it is uncertain which one (we can suppose there is a fact of the matter about which person will exist, but we do not know it). The Prioritarian first compares the state of affairs in which Adele will exist and is benefited with the state of affairs in which Belda will exist and is benefited, and then judges that the latter is better. Then, she applies expected value theory. She does not apply her view to a person whose existence is contingent on what we do (compare this with how a Person-Affecting Utilitarianism conjoined with expected value theory would handle cases in which it is uncertain who will exist; presumably this is importantly different from how it would handle cases involving persons whose existence is dependent on what we do). The reason that this is worth mentioning is that, if AM *did* apply to persons whose existence is contingent on what we do, it might have even more troubling implications, regarding the Repugnant Conclusion, than straight Utilitarian views do. For discussions of the Priority View in connection with population ethics, see: Nils Holtug 1999 and 2010 and Brown 2007.

Next, one could deny that we are forced to appeal to AM to capture the judgment that it is more morally important to provide the benefit for Belda. One could claim that it is better to benefit Belda because she has a 50 percent chance of ending up worse off *relative to* how well off Adele has a 50 percent chance of ending up. That is, Belda is worse off *in prospect* relative to Adele. So RM can, one might argue, accommodate our judgment about *Only One of Two Will Exist* after all.

Note that if it does not matter how well off in prospect *people who might not exist* are relative to each other,[185] then it is clearly necessary to appeal to AM to capture our judgment about *Only One of Two Will Exist*. But suppose *arguendo* that it does. Even if it does, the appeal to how well off in prospect people who might not exist are relative to each other can only get us so far. Consider another case:

> *One of Two Possible Miseries*. No one exists, nor has existed, nor will exist, other than Colette and Darcie. Colette and Darcie do not yet exist, only one of them will exist, and there is a 99 percent chance that Colette will exist, and a 1 percent chance that Darcie will exist. If Colette existed, her life would be not worth living (as Parfit says, it would be *worse than nothing*). If Darcie existed, her life would be not worth living by a considerably greater margin than the life Colette's would be, if she existed. We can bring it about that, if Colette exists, she will receive a large benefit, or that, if Darcie exists, she will receive an equally large benefit. Other things are equal.

---

[185] It would not matter, for example, if it did not even *make sense* to compare how well off in prospect persons who might not exist are.

Here it seems plausible that we should bring about the 99 percent chance of the benefit for Colette. But suppose we ask: Which outcome would be morally better, the one in which Colette received the large benefit, or the one in which Darcie received an equally large benefit?

In this example, we can suppose that, while Darcie's life would be considerably worse than Colette's life would be, Colette is considerably worse off than Darcie in prospect – that is, suppose that a 99 percent chance of Colette's not worth living life is a worse prospect than a 1 percent chance of Darcie's considerably worse life (for example, .99 times -10 is less than .01 times -980).

It seems plausible that if Darcie received a large benefit this would be more morally important than if Colette received an equally large benefit. But we cannot capture this belief by appealing to the claim that Darcie is worse off relative to Colette. Darcie cannot be worse off relative to Colette, since in the outcome in which Darcie exists, Colette does not exist (nor does anyone else). And Darcie is not worse off relative to Colette in prospect. Insofar as it is plausible to compare the prospects of merely possible persons, Colette is worse off in prospect relative to Darcie. Therefore, to accommodate the belief that "if Darcie received a large benefit, this would be more morally important than if Colette received an equally large benefit," it seems we must appeal to AM. It seems to matter how well off Colette or Darcie would be in non-relative terms.

Another attempt to avoid appealing to AM involves instead appealing to the claim that in the outcome in which Darcie exists, she is worse off relative to someone who *could have* existed but never will exist, Colette. If it is implausible to extend RM to apply to persons who will certainly never exist, we will have to appeal to AM to accommodate our belief.[186] But suppose it is not. We can again revise the case, so that Colette never *could have* existed. In such a case, we cannot claim that Darcie is worse off relative to someone who could have existed. Consider:

> *My Morally Selfish Glee*. What I really care most about is how much moral value results from actions *I* perform. I do not care about how much moral value there is, or how much others bring about. I am in a position to benefit anyone who exists. Darcie exists, and Colette never could have existed. I am *glad* that Darcie exists, rather than Colette.

It does seem appropriate, relative to what I care about, for me to be glad that Darcie exists rather than Colette. This reflects the belief that benefiting Darcie is more morally valuable, or important, than benefiting Colette.

One might object to *My Morally Selfish Glee*, and claim that it is incoherent or inappropriate to be glad that X happened rather than Y, when Y could not have happened. But these claims seem false. For example, you can coherently be glad that *merely thinking about miserable people did not cause more miserable people to exist*

---

[186] Notice again that *AM*, on the other hand, would *not* have to apply to persons who certainly will never exist (but who could have existed), in order to capture the belief in question. In the outcome in which Darcie exists, we make no reference to anyone else in explaining the moral value of benefiting Darcie.

even though you believe that this could not have happened. It is also not

inappropriate for you to be glad about this. It makes perfect sense to be glad that

certain bad things could not have happened. And, for a morally selfish person who

cares only about how much moral value she promotes, it makes sense relative to what

she cares about to be glad that Darcie exists, rather than Colette. And considering a

case in which things are reversed, such that Colette exists and Darcie could never have

existed, it would make sense for this morally selfish person, relative to what she cares

about, to *wish* that Darcie could and will exist. (Compare: for those of us who believe

that suffering is bad, it makes sense for us to wish that no one had ever suffered, even

though we know that it is impossible to have our wish fulfilled).

A final attempt to avoid appealing to AM involves simply denying the belief

that it is more morally important to benefit Belda (in *Only One of Two Will Exist*) or

the belief that it is more morally important to benefit Darcie (in *One of Two Possible

Miseries* and *My Morally Selfish Glee*).

But first, notice that we cannot appeal to intrapersonal compensation as a

reason for being indifferent between the same-sized benefits in these cases, because

Adele and Belda, and Colette and Darcie, are separate persons.

And second, although these cases do not involve persons who are worse off

relative to others (at least not in any standard sense), they seem to me to elicit beliefs

which are no less plausible than those undergirding RM. It is hard to see how these

beliefs about these cases could be denied by someone who accepts those beliefs which

undergird RM. Therefore, it seems to me hard to defend RM *without also* accepting

AM.

Appendix C: Rational Prudence and Morality

Recall the Rational Prudence Constraint, which I mentioned at the end of Appendix A:

> *Rational Prudence Constraint*: we ought not to treat a person in ways in which she could not prudentially rationally choose to be treated, if it were only the well-being or interests of this person our act could affect, and other things were equal.

Again, *if* it is prudentially rational for one to choose, under risk, to maximize one's expected benefits, then the Rational Prudence Constraint will conflict with AM. This is because, as shown in cases like One Child, AM implies that we should sometimes provide a person with slightly smaller expected benefits rather than larger ones, even when it is only her interests that are at stake. But we might think that it is *not* prudentially rational for one to choose, under risk, to maximize one's expected benefits. If this were correct, then the Rational Prudence Constraint would not conflict with AM. It therefore seems worth inquiring about the viability of this non-standard conception of rational prudence.

We could deny that rational prudence *would* favor the treatment for the Slight Impairment in One Child; we could interpret AM as saying not just that benefits to the worse off *morally* matter more the worse off they are in absolute terms, but also that

benefits to the worse off *prudentially* matter more the worse off they are in absolute terms. Then the child in One Child, if rational and concerned solely with her own interests, would care more about the benefit for her life if she ended up with the Very Severe Impairment than she would care about the slightly larger benefit for her life if she ended up with Slight Impairment. She would, that is, rationally prudentially prefer the treatment for the Very Severe Impairment than the treatment for the Slight Impairment. Let us call the move articulated in this paragraph the *Appeal to Prudence*.[187]

Indeed, the Appeal to Prudence would make AM consistent with the Rational Prudence Constraint. But the Appeal to Prudence faces a separate problem. Roughly, the problem is that we may have no plausible way to measure the size of benefits other than by appealing to the recommendations of rational prudence under risk.

The orthodox measure of the size of benefits, or increases in utility, is the von Neumann-Morgenstern measure (henceforth *VN*):

> If and only if a person is in fact indifferent between equal chances of benefit B_1 and benefit B_2, then B_1 and B_2 are same-sized benefits for this person. If and only if a person in fact prefers an equal chance of B_1 over an equal chance of benefit B_2, then B_1 is a larger benefit than B_2 for this person.

---

[187] I am grateful to Dick Arneson for suggesting this idea.

I believe we should reject VN. It is a mistake to rely on a person's actual preferences, or the preferences she would have if perfectly informed, as the measure of the size of benefits to her. People can fail to prefer what would make them better off, or expectably better off (even if you disagree, the discussion below also applies to VN). More plausible than VN is the *Rational* von Neumann-Morgenstern measure (henceforth *RVN*):

> If and only if it would be prudentially rational for a person to be indifferent between equal chances of benefit B_1 and benefit B_2, then B_1 and B_2 are same-sized benefits for this person. If and only if it would be prudentially rational for a person to prefer an equal chance of B_1 over an equal chance of benefit B_2, then B_1 is a larger benefit than B_2 for this person.[188] In these biconditionals, the notion of *rational prudence* is the prior, unanalyzed one. The size of benefits are defined and measured in terms of this notion.

The problem, however, is that the Appeal to Prudence is inconsistent with RVN. In One Child, it was stipulated that the treatment for the Slight Impairment induces a slightly larger benefit than the treatment for the Very Severe Impairment. RVN thus implies that it would be prudentially rational for the child to prefer a 50 percent chance of the larger benefit over a 50 percent chance of the slightly smaller one. But the Appeal to Prudence insists that since benefits to worse off lives

---

[188] In proposing RVN, I am proposing a version of the *Sure-Thing Principle*. For an explanation and defense of this principle, see Broome 1991, especially chapters 4 and 5.

prudentially matter more, it would be prudentially rational for the child to prefer a 50 percent chance of the slightly smaller benefit over a 50 percent chance of the larger one. We therefore cannot accept both RVN and the Appeal to Prudence. We might wonder: why, then, can't defenders of the Appeal to Prudence simply reject RVN?

If we are to sensibly debate whether or not a benefit for a worse off life matters more than a same-sized benefit for a better off life, we require a definition of what it *means* for two benefits to be the same size. Without such a definition, our debate will get nowhere. Suppose that we are asked whether we should provide benefit B_1 for a badly off life or instead provide benefit B_2 for a well off life. And suppose we have the intuition that we should provide B_1 for the badly off life. Without a measure of the size of benefits, what, if anything, this intuition provides evidence for will be *ambiguous*. Should we provide B_1 because benefits to worse off lives matter more, independently of their size? Or should we provide B_1 because B_1 is bigger than B_2? We cannot say unless we know what it means for benefits to be larger, smaller, or the same size as, other benefits. This is what I call the *Ambiguity Problem*.

A defender of the Appeal to Prudence might claim that she can get on without a measure of the size of benefits. She might claim, for instance, that in One Child, it is simply stipulated that the benefit for the life with the Very Severe Impairment is slightly smaller than the benefit for the life with the Slight Impairment. She might then ask: despite the slightly greater size of the benefit for the life with the Slight Impairment, doesn't the benefit for the worse off life prudentially matter more?

But this will not do. Unless we know what it means for the two benefits to be the same size, it is left open that the size of the benefit for the worse off life is in part a function of the very fact that it is a benefit for a worse off life.

RVN provides a clear way to measure size of benefits. Without *any* measure of size of benefits, we cannot sensibly debate the question of whether benefits to the worse off matter more, since we cannot sensibly distinguish between the size of benefits and the further extent to which they matter independently of their size. And I am afraid I am not aware of a clear alternative to RVN that would adequately suit our purposes.

We might think that, instead of RVN, we can appeal to substantive theories of well-being to measure size of benefits. For example, suppose we believed:

> *Hedonism*: pleasure is the only thing that ultimately makes lives go well, and pain is the only thing that ultimately makes lives go badly.[189]

Given Hedonism, we might appeal to the following measure of size of benefits:

> *Hedons*: any two benefits, B_1 and B_2, are the same size if and only if they induce equal amounts of pleasure (minus pain). B_1 is a larger benefit than B_2 to the extent that it induces more pleasure (minus pain).

---

[189] For a contemporary statement of Hedonism, which is eminently clear and defensible, see Crisp 2006a.

We have, in a way, pushed back our problem. For now we require a measure of amounts of pleasure. Let us just assume that we have a perfect measure of pleasure. The Ambiguity Problem remains.

Suppose there are two possible lives – L_1 and L_2 – that a single person could live. L_1 contains a very small amount of pleasure. L_2 contains a very large amount of pleasure. A pleasure can be achieved for L_1 or instead an equally large pleasure can be achieved for L_2. Suppose we have the intuition that we should achieve the pleasure for L_1.

A defender of AM might be tempted to claim that this intuition counts as evidence in favor of her view. But what argument could she possibly give against an objector who claimed that the intuition was instead evidence *against* Hedons, or indeed against Hedonism? The objector might claim that we should achieve the pleasure for L_1 because this would result in a *greater benefit* than an equal pleasure for L_2. What this intuition provides evidence for is ambiguous, and so AM will fail to acquire unambiguous intuitive support. If versions of egalitarianism and prioritarianism that accept AM cannot answer this Ambiguity Problem, they will rest on unacceptably shaky ground.

It is not hard to see that if, instead of Hedons and Hedonism, we appeal to other substantive theories of well-being (e.g., Perfectionist or Objective List theories) in search of an alternative measure to RVN, we will not make any more progress on the Ambiguity Problem.

It is perhaps worth noting briefly: RVN is neutral about *what*, in particular,
well-being gains consist in – since it is neutral about *what*, in particular, it would be
prudentially rational to prefer under risk. It merely provides a definition of the size of
benefits, whatever they are. (It is therefore consistent with Hedonism, Desire-
Satisfaction, and Objective List theories of well-being).

Possibly the best response to the Ambiguity Problem, other than that provided
by RVN, is to simply appeal to our *intuitive beliefs about how big benefits are*,
whatever we think well-being, or increases in well-being, consists in. But this
response still might not avoid the Ambiguity Problem. When we judge it better to
benefit L_1 rather than provide *what we believed to be* a same-sized benefit for L_2, who
is worse off than L_1, there are two claims this judgment might support: (1) that what
we believed to be an equal benefit was in fact mistaken, and (2) that it is better to
benefit people who are worse off. Unless we were given strong reasons for (2) as
opposed to (1), or for the claim that we are not mistaken in our belief about the size of
the benefits in question, the ambiguity would remain.

One of the principal advantages of RVN is that it cleanly solves the Ambiguity
Problem. Suppose we adopt RVN, and find two benefits, B_1 and B_2, which are the
same size. Further suppose that we have the intuition that we should achieve B_1 for L_1
rather than achieve B_2 for L_2. As before, a defender of AM might claim that this
intuition counts as evidence in favor of her view. I believe that she would *now* be on

firm ground to claim this.[190] For now, unlike before, she would have a good argument

against someone who objected that the intuition that we should achieve B_1 for L_1

rather than achieve B_2 for L_2 is evidence (not for AM but) that B_1 for L_1 is a bigger

benefit than B_2 for L_2.

Her argument would be that, when she considered herself having a 50 percent

chance of B_1 for L_1 versus a 50 percent chance of B_2 for L_2, it was prudentially rational

for her to be indifferent between the two. The key is that RVN, in assessing the size

of B_1 and B_2, directly compares, and enables judgments about, the relative prudential

value of these benefits in the context of the lives they would be realized in (L_1 and L_2).

It thus allows us to rule out the possibility that the judgment that we should provide a

benefit for a worse off life rather than a benefit for a better off life is really evidence

for the view that benefits interact with lives such that they become bigger, the worse

off these lives are. This is how it solves the Ambiguity Problem.

Again, if upon adopting RVN, we find that it is prudentially rational to be

indifferent between B_1 for L_1 and B_2 for L_2, we cannot consistently claim that

achieving B_1 for L_1 is prudentially better than achieving B_2 for L_2. However, we

might have the intuition that achieving B_1 for L_1 is *morally* better than achieving B_2

for L_2. If so, then we will have found unambiguous intuitive support for AM. The

cost of adopting RVN to solve the Ambiguity Problem, of course, is that it makes it

impossible to offer the Appeal to Prudence and thereby make AM consistent with the

Rational Prudence Constraint.

---

[190] Matthew Adler, a defender of the Priority View, solves what I am calling the Ambiguity Problem in a similar if not the very same way in Adler 2011, chapter 3.

Finally, I am unaware of how to *decisively* answer the Ambiguity Problem without appealing to RVN, or something quite similar to it (e.g., VN). Perhaps there is an alternative solution. But until a clear alternative to RVN is proposed and defended, it looks like the Appeal to Prudence will render AM consistent with the Rational Prudence Constraint only at the cost of lacking a decisive answer to the Ambiguity Problem. But again maybe a plausible (though I don't think decisive) answer to this Problem is to simply appeal to our intuitive beliefs about the size of benefits. It is not clear to me how plausible an answer this is.

Finally, a distinct option for defenders of AM is to deny the Rational Prudence Constraint. But first, we might believe that this constraint is hard to deny. Secondly, this constraint seems to garner intuitive support which is similar to, and perhaps no weaker than, the intuitions opposed to levelling-down. Some versions of the Equality View could imply that sometimes we ought to treat a person in a way in which she could not prudentially rationally choose to be treated – even if her well-being were the only thing our action would affect. It could have this implication if this person were better off than most other people, since increasing her well-being would increase inequality of well-being across persons. Some defenders of AM claim that this is implausible; but is such levelling-down any more implausible than *their* denial of the Rational Prudence Constraint?

We might think that it is in *some cases* acceptable to deny the Rational Prudence Constraint, but that in others it is not. If we believe that we should level-

down well-being for the sake of reducing inequality, for example, then we believe we should do what will *actually* make some people worse off, and none better off. If instead, in One Child, we believe that we should provide the treatment for the Very Severe Impairment rather than the treatment for the Slight Impairment, then we believe we should do what will *expectably* make some people worse off, and none better off. In both kinds of case, the Rational Prudence Constraint is denied. But we might think that it is plausibly denied in the second kind of case, but not plausibility denied in the first.

It does initially seem odd that we should find the Rational Prudence Constraint plausible when dealing with cases in which what the outcomes will be is certain, but that we should find it implausible when dealing with cases in which what the outcomes will be is uncertain. But then again, it may be appropriate to treat these kinds of cases differently. Consider an analogy. We may find the claim that we should not do what is worse for everyone to be plausible in what Parfit calls *Same-People Choices*, but find it very implausible in *Different-People Choices* (see chapter 1 for definitions of these terms). Similarly, we might think that in *Certain-Outcome Choices* the Rational Prudence Constraint is plausible, but that in *Uncertain-Outcome Choices* it is not.

We might think these latter two kinds of cases should be treated differently if we have doubts about personal identity across possible worlds.[191] Or we might believe that the personal unity that matters about personal identity across time within a

---

[191] If Lewis 1971 about transworld personal identity were correct, then One Child may not be a genuine one-person case.

single world does *not* hold between a person in one world and the person in another possible world that she is identical with.[192] If either were true, we might not be able to appeal to the Intrapersonal Compensation Justification to justify, e.g., providing the treatment for the Slight Impairment in One Child, since this would then be analogous to giving one person a slightly larger benefit rather than giving a separate (worse off) person a slightly smaller one. Although these possibilities are worth exploring, I will not explore them here.

If neither of these possibilities provide a solid basis for maintaining a sharp moral distinction between Certain-Outcome Choices and Uncertain-Outcome Choices, then it would seem quite odd that the Rational Prudence Constraint should be plausible when the outcomes of our acts are 100 percent certain, but implausible when they are only 99.999 percent certain. Should such a small difference really morally matter so much? It seems to me that *if* we decide we can tolerate denying the Rational Prudence Constraint, we should tolerate denying it *both* in Certain-Outcome Choices as well as in Uncertain-Outcome Choices. Defenders of AM, however, could defensibly deny my claim. They might judge that it is, on reflection, plausible to violate the Rational Prudence Constraint in cases like One Child (and versions of the case involving different probabilities) but implausible to violate it in cases like Levelling-Down.

Nonetheless, I think that defenders of RM and AM alike should at least admit that it is *somewhat* implausible to deny the Rational Prudence Constraint. Views which violate this constraint seem to be less plausible than views which do not, *at*

---

[192] For a fascinating defense of this view, see Velleman 2008, 221-244.

least in one respect. Determining which of the views here discussed is *overall* most plausible, or least implausible, is a considerably harder task.

One final remark: we might think that we can defend RM without denying the Rational Prudence constraint. For example, perhaps we could defend a version of RM which denies the Equality View (and thereby avoids violating the Rational Prudence Constraint in Certain-Outcome Choices), *and* denies AM (and thereby avoids violating the Rational Prudence Constraint in Uncertain-Outcome Choices). This view may ultimately be plausible, but it does face a problem: as I argued in Appendix B, it seems hard to accept RM without also accepting AM. The bedrock beliefs which undergird RM seem no less plausible than those elicited by Only One of Two Will Exist and One of Two Possible Miseries. But to capture our beliefs about these cases, we have to appeal to AM.

Appendix D: Triage Priority Views and Diminishing Priority Views

Recall that, according to the Priority View, benefiting people matters more the worse off these people are. A very natural assumption is that this view, and the intuitions which support it, would apply to people however badly off they are. It is possible, however, to deny this. Indeed, defenders of the *Triage Priority View*[193] claim that:

> Benefiting people matters more the worse off these people are, except for people who are sufficiently badly off. Once people are sufficiently badly off, benefiting these people does not matter more the worse off they are.

When two people are sufficiently badly off, the Triage Priority View implies that it fails to be the case that it is better to achieve a slightly smaller benefit for the worse off of the two. If this view were true, it would provide us with a way of denying Parity and those Premises near P*n*. Such Premises are false, it could imply, because they involve people whose lives are sufficiently bad. The Triage Priority View, if it were plausible, thus seems to provide a way to defend the Priority View without fear of implying Priority Monster. Before turning to the question of whether it does provide a plausible way to avoid Priority Monster, it may be useful to explore some specific ways in which this view could be formulated or defended.

---

[193] I believe I first came across this variant of the Priority View in Brown 2007.

We can call the particular well-being threshold relevant to the Triage Priority View the *triage threshold*. That is, the Triage Priority View implies that benefiting people below the triage threshold does not matter more the worse off they are. I assume that there can be significant vagueness or indeterminacy about where the triage threshold lies, and do not regard this sort of vagueness as a problem for the Triage Priority View. But the triage threshold, it seems, should be lower than the level of well-being which Sufficiency Views regard as sufficient (recall the discussion of Sufficiency Views in 3.4.3). It seems very implausible that considerations of priority for the worse off would *only* apply to those who count, if anyone does, as sufficiently well off.[194]

Perhaps a nonarbitrary triage threshold is the point at which lives are no longer worth living. Or perhaps, on reflection, we will think there is a plausible triage threshold which is higher or lower than this. In any case, there are a few structural decisions which defenders of the Triage Priority View must make. They could accept either:

> *Harsh and Abrupt Triage*: benefits to people below the triage threshold matter
> in the way that the Equal Weight View claims that all benefits matter.

Or:

---

[194] Defenders of Sufficiency Views would no doubt agree, as they believe that considerations of priority for the worse off *only* apply *below* the sufficiency threshold (rather than that they *only* apply *above* it!).

Harsh and Gradual Triage: benefits to people below the triage threshold gradually matter less and less, approaching the in the way in which the Equal Weight View claims that all benefits matter.

Or:

Compassionate Triage: benefits to people below the triage threshold should receive equal priority weight.

We can illustrate these three views by comparing some simple lists of numbers, which represent people's well-being scores and the corresponding priority weights. Each lists moves from a well-being score (left) to a corresponding priority weight (right). And I here assume, for simplicity, that the triage threshold is at the point where life becomes worse than nothing.

Harsh & Abrupt:			*Harsh & Gradual*:			*Compassionate*:		
5	→	1	5	→	1	5	→	1
4	→	2	4	→	2	4	→	2
3	→	3	3	→	3	3	→	3
2	→	4	2	→	4	2	→	4
1	→	5	1	→	5	1	→	5
0	→	6	0	→	6	0	→	6

--------------------TRIAGE THRESHOLD---

-1	→	1	-1	→	5	-1	→	7
-2	→	1	-2	→	4	-2	→	7
-3	→	1	-3	→	3	-3	→	7
-4	→	1	-4	→	2	-4	→	7
-5	→	1	-5	→	1.5	-5	→	7
-6	→	1	-6	→	1.25	-6	→	7
-7	→	1	-7	→	1.125	-7	→	7

Figure 10: Triage Priority Views

Some defenders of the Triage Priority View might prefer one of the variants of

Harsh Triage over Compassionate Triage. Of the two variants of Harsh Triage, Harsh

and Gradual Triage seems more plausible. It seems hard to believe, as implied by

Harsh and Abrupt, that benefiting people who are just above the triage threshold

would matter *much* more than benefiting those who are worse off and just below the

triage threshold. Defenders of Harsh and Gradual avoid this problem, as their view

implies that benefiting those just above the triage threshold only matters *a little* more

than benefiting those who are just below the triage threshold. Defenders of Harsh and

Gradual might appeal to the following rationale for their view:

Higher Outcomes: benefits for people who are below the triage threshold matter more, the higher the well-being score with which they leave these people.

According to this view, moving a person from -3 to -2 would be better than moving a person from -99 to -98 (assuming both of these people are below the triage threshold). Though Higher Outcomes may appear inimical to the Priority View, it is important to remember that it is only applies to people who are below the triage threshold.

Compassionate Triage seems more plausible than either version of Harsh Triage. Both versions of Harsh Triage have the implausible implication that it would be better to achieve a slightly smaller benefit for someone well above the triage threshold (e.g., at 4) than to achieve a slightly larger benefit for someone well below it (e.g., at -7). On the other hand, Compassionate Triage is inconsistent with Higher Outcomes.

But Higher Outcomes at most only seems plausible near the triage threshold, e.g., in the comparison between moving a person from -3 to -2 and moving a person from -99 to -98; it does not seem plausible far from the threshold, e.g., in the comparison between moving a person from -1,000 to -995 and moving a person from -10,000 to -99,995. This suggests that we never really found Higher Outcomes plausible, but rather perhaps a view which claimed that it matters more to leave people

with a well-being score within a zone near the triage threshold than to leave them with a well-being score outside it.

Let us leave these structural questions about the Triage Priority View unanswered and ask whether, however they are answered, this view would enable us to avoid plausibly Priority Monster. There is a question prior to the question of whether the Triage Priority View would enable us to avoid *plausibly* Priority Monster: whether it would, plausibly or not, enable us to *avoid* Priority Monster.

Recall the Series from 3.2. As we move down the Series, we will eventually come upon lives below the triage threshold. When we do, the Triage Priority View will assign them less priority weight, making the corresponding Premises false. It would thus avoid Priority Monster.

We could reconstruct the Series such that it is restricted to persons with lives above the triage threshold. But then for it to remain true that each person in the Series could be made arbitrarily worse off than her predecessor, we would have to allow that the first person in the Series could be arbitrarily well off. Then if the Premises were true, they would only imply:

> *Less Scary Priority Monster*. There could be a person so badly off that, if she existed, it would be better to achieve an arbitrarily small benefit for her than to achieve an arbitrarily large benefit for an arbitrarily well off person.

Some people might still find Less Scary Priority Monster to be implausible. But it definitely seems less implausible than Priority Monster, and some people might not find it implausible at all. Indeed, many defenders of Sufficiency Views will not find it implausible it all, insofar as the arbitrarily well off person is far above the sufficiency threshold.

It is possible that we would find Less Scary Priority Monster to be more implausible if the arbitrarily well off person lived a very long life, much of which was filled with horrible suffering. We could suppose that, though this person suffered intensely and for hundreds or thousands of years, she is arbitrarily well off in virtue of the enormous amounts of others goods that (at other times) her life contains. And we could suppose that her successor in the Series is arbitrarily worse off than she is, and that the next person is arbitrarily worse off, and so on, and that at the end of the Series is a person just above the triage threshold.

In sum, the Triage Priority View, if true, would allow us to avoid Priority Monster. It threatens to imply Less Scary Priority Monster, but we might not find that implication to be implausible. We might not find Less Scary Priority Monster implausible even if the arbitrarily well off person's life also contained an arbitrarily great sum of suffering.

Now return to the question of whether the Triage Priority View would enable us to avoid *plausibly* Priority Monster (and recall the discussion of Parity in 3.3.2). I think we can grant, for the sake of argument, that some claims of form Ω, which pertain to persons relevantly near the triage threshold, are less intuitively plausible

than the most plausible claims of form **Ω**. But suppose we set up the Series such that

the first person in it is arbitrarily far below the triage threshold, and such that an

arbitrarily large benefit for this person would fail to bring her anywhere close to this

threshold. We can also suppose that none of the benefits in the Series would, if

received, make the recipient better off than her predecessor.

This version of the Series achieves ample distance from whatever

considerations involving proximity to the triage threshold may have tempted people to

accept the Triage Priority View in the first place. And this version of the Series is

arranged so that each Premise pertains to two alternatives: (1) achieve a benefit for an

arbitrarily badly off person, or (2) achieve a slightly smaller benefit for an arbitrarily

worse off person. Each of the Premises would say that (2) is better. On careful

reflection, each of these Premises seems at least as intuitive as any claim of form **Ω**.

This, I believe, shows that we should either reject the Triage Priority View, or

restrict it in a way so that it can accommodate these claims about this version of the

Series involving persons relevantly far below the triage threshold. And of course, if it

accommodates these claims, it implies Priority Monster.

Since the Triage Priority View fails to provide a plausible way to deny Parity

and the Premises, it fails to provide a plausible way to avoid Priority Monster.

The Triage Priority View may appear inimical to the Priority View. This is

because it says that once people are sufficiently badly off, benefiting these people does

not matter more the worse off they are. We could avoid Priority Monster in the way

the Triage Priority View does while perhaps remaining truer to the spirit of the

Priority View if we adopted the:

Diminishing Priority View: benefiting people matters more the worse off these

people are. But once people are sufficiently badly off, the degree to which this

matters more diminishes, the worse off people are.

There is much room for disagreement about what level of well-being would

count as relevantly sufficiently badly off, and it might be that there is no single level at

which people become relevantly sufficiently badly off, but instead a wide range or

zone. Moreover, defenders of the Diminishing Priority View can choose between:

Non-Asymptotically Diminishing: there is no limit on how much more it

matters to benefit people, the worse off they are.

And:

Asymptotically Diminishing: there is an asymptotic limit on how much more it

matters to benefit people, the worse off they are.[195]

---

[195] Defenders of the Diminishing Priority View *could* accept a view according to which there is a limit
on how much more it matters to benefit people, the worse off they are, without claiming that this limit is
asymptotically or gradually approached. It is not clear, however, what advantage this view would have
over Asymptotically Diminishing, and it would face the problem with Asymptotically Diminishing
noted below.

We can illustrate these two views by again comparing some simple lists of numbers, which represent people's well-being scores and the corresponding priority weights. (I assume, for simplicity, that there is single level at which people become relevantly sufficiently badly off, and that it is at the point where life becomes worse than nothing).

Non-Asymptotically Diminishing:

5	→	1.5
4	→	2
3	→	3
2	→	4
1	→	5
0	→	6

--------------------------SUFFICIENTLY BADLY OFF-

-1	→	7
-2	→	7.5
-3	→	7.833333333
-4	→	8.083333333
-5	→	8.283333333
-6	→	8.45
-7	→	8.59

Asymptotically Diminishing:

5	→	1.5
4	→	2
3	→	3
2	→	4
1	→	5
0	→	6

-1	→	7
-2	→	7.5
-3	→	7.75
-4	→	7.875
-5	→	7.9375
-6	→	7.96875
-7	→	7.984375

Figure 11: Diminishing Priority Views

Starting with a priority weight of 7 for a well-being score of -1, for each additional point drop in well-being score, the priority weights in the portrayed Non-Asymptotically Diminishing Priority View increase by 1/2, 1/3, 1/4, 1/5, 1/6, 1/7, etc. These priority weights approach no limit. Starting with a priority weight of 7 for a well-being score of -1, for each additional point drop in well-being score, the priority

weights in the portrayed Asymptotically Diminishing Priority View increase by 1/2,

1/4, 1/8, 1/16, 1/32, 1/64, etc. These priority weights approach a limit of 8. This

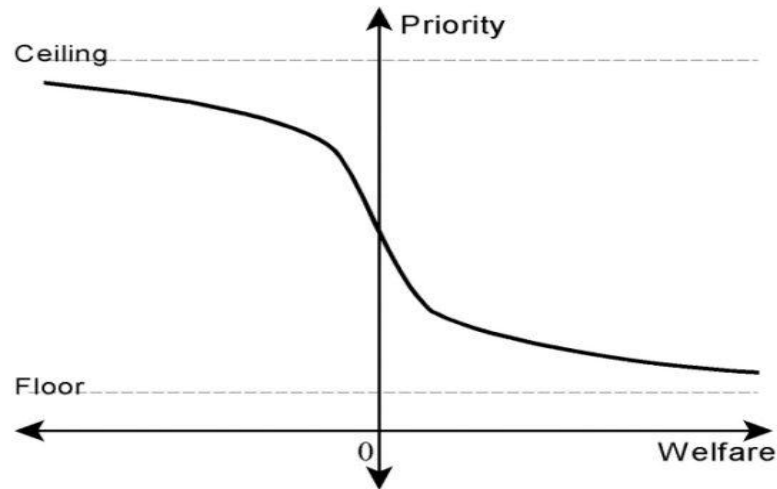Asymptotic View can be illustrated by a graph:[196]

Figure 12: Asymptotically Diminishing Priority View

The horizontal axis represents well-being scores, and the vertical axis

represents the priority weights that should be assigned to persons with those well-

being scores. We can suppose that the *Floor* for priority weight is 1, and that the

Ceiling is 8; these are the asymptotes. The Non-Asymptotic Diminishing Priority

View might have a similarly shaped curve, but it would not be bounded by a Ceiling.

It is important to distinguish between these two versions of the Diminishing

Priority View, because while Asymptotically Diminishing can avoid Priority Monster,

Non-Asymptotically Diminishing cannot. Non-Asymptotically Diminishing implies

---

[196] Thanks to Mike Huemer for making this graph.

that the priority weight an arbitrarily badly off person should receive is arbitrarily great, and this implies that it is better to achieve a very small benefit for such a person than to achieve a very large benefit for the currently worst off person. Asymptotically Diminishing, on the other hand, implies that an arbitrarily badly off person should not receive more than some finite amount of priority weight, and the product of this priority weight and the size of a very small benefit may be far less than the product of the priority weight assigned to the currently worst off person and the size of a very large benefit. Asymptotically Diminishing can, in this way, avoid Priority Monster.

Recall that, to avoid Priority Monster, we have to deny at least one of the Premises or deny Transitivity. If Asymptotically Diminishing were true, it would provide us with a way of denying Premises near Pn. Such Premises are false, it could imply, because they involve people whose lives are sufficiently bad. However small the difference between adjacent benefits in those Premises, and however great the difference in well-being level of the recipients of those benefits in those Premises, the priority weight assigned to the worse off person might fail to be sufficiently great to render it better to achieve the slightly smaller benefit for the worse off person. We might reply to Asymptotically Diminishing as follows:

It is true that, even for those Premises near Pn, the worse off person receives greater priority weight according to Asymptotically Diminishing. We can therefore stipulate that, for each of those Premises, the benefit the worse off

person would receive is slightly smaller *such that* Asymptotically Diminishing *would* imply that each of those Premises is true.

The problem with this reply is, roughly, Zeno's Paradox. If we adjust the Series so that the size difference between adjacent benefits decreases such that Asymptotically Diminishing would accept each of the Premises, the *benefit size* will asymptotically approach a limit, and thus fail to reach a "very small" or arbitrarily small benefit. For example, for Asymptotically Diminishing to accept each of the Premises, the Series might have to be adjusted so that the size of the benefits decreased by 1/2, 1/4, 1/8, 1/16, 1/32, 1/64, etc. But if we start with a benefit of just size 10 at the beginning of the Series, then decreasing the benefit size in this way will yield a Series with benefits of the following decreasing sizes: 9.5, 9.25, 9.125, 9.0625, 9.03125, and so on. The benefits in this Series, though they get smaller and smaller, will never get any smaller than size 9. We will get closer and closer to our Tortoise, but will never catch it.

Instead, we should construct the Series such that the size difference between adjacent benefits is an arbitrarily small but *constant* number. If it is a constant, then it is a plain mathematical fact that a sufficiently large number of such differences can take us from an arbitrarily large benefit at the beginning of the Series to an arbitrarily small benefit at the end of the Series. Thus, we avoid Zeno's Paradox.

But if the size difference between adjacent benefits is a constant, then defenders of Asymptotic Diminishing can claim that the priority weight assigned to

the worse off person in each Premise near Pn is only *barely* greater than the priority weight assigned to the better off person, such that it is worse to achieve the slightly smaller benefit for the worse off person. This is because, for any constant difference between adjacent benefits, Asymptotic Diminishing can claim that eventually the difference in priority weight assigned to adjacent persons will become relevantly smaller than it. So Asymptotic Diminishing can deny the Premises and thereby avoid Priority Monster.

However, the fact that Asymptotic Diminishing can avoid Priority Monster does not show that it is not implausible. Indeed, it is implausible for the reasons already given in support of Parity (in 3.3.2) and the Conditional (in 3.3.3). Assuming the Series is arranged so that each person is arbitrarily worse off than her predecessor and the size difference between adjacent benefits is an arbitrarily small but constant number, the Premises near Pn are just as intuitively plausible as any claim of form Ω. The Asymptotic Diminishing Priority View is implausible because it accepts some claims of form Ω on the basis of their intuitive plausibility, but rejects other claims of form Ω (those near Pn) which are no less intuitive and are plausibly the same kind of claim.[197] It thus violates the Conditional, which plausibly requires us to be consistent about our acceptance standards, and to treat like cases alike.

---

[197] Moreover, notice that if we, holding other things equal, replaced all the benefits in the Series with very large ones (suppose the benefit in Pn were arbitrarily large, and, working backwards to P1, the benefits got slightly larger and larger at each step), Asymptotic Diminishing would still imply that the Premises near Pn are false. Defenders of the Nontrivial Benefits Priority View might claim that the way in which Asymptotic Diminishing denies the Premises is implausible while maintaining that the way in which the Nontrivial Benefits Priority View does so is not implausible. But, as I claimed in 3.4.4, the latter is still implausible in virtue of its denial of Parity, and in any case it still faces *Two Priority Monsters with Nontrivial Benefits*.

Appendix E: The Infection Problem

In order to state the Conditional and its role in the Priority Monster Spectrum

Argument in a timely fashion, I had to omit discussion of a possibility which might

allow a defender of the Priority View to avoid Priority Monster without denying

Tradeoffs, Parity, or the Conditional. I will here discuss this possibility. First, recall

the Conditional: if we accept claim C1 on the basis of its intuitive plausibility, and if

claim C2 is at least as intuitively plausible as C1, and if other things are equal, then we

should also accept C2. In cases where we know the conjunction of C1 and C2 is false,

we can only avoid this false claim and satisfy the Conditional by accepting neither C1

nor C2. Consider the:

> *Infection Problem*. The implausibility of an implication (Priority Monster) of
>
> the conjunction of claims of form $\boldsymbol{\Omega}$, taken together with the fact there is
>
> equally good reason to accept or reject each of them (Parity), implies that there
>
> is at least some reason to reject each and every such claim (the Conditional).
>
> The implausibility of the conjunction of these claims "infects" each of the
>
> individual conjuncts.

It could be that claim C1 and one of the Premises, e.g., P*n*-3, are equally intuitively

plausible, and yet the Conditional would not imply that if we accept C1 we should also

accept P*n*-3. This could be because P*n*-3 faces the Infection Problem whereas C1 does

not, and it would fail to be true that other things are equal. Likewise, a defender of the

Priority View could defensibly claim that the Infection Problem does not impugn *all*

claims of form **Ω**. After all, it is *the Premises* that entail Priority Monster, rather than

other claims of form **Ω**. Other claims of form **Ω** are not essential to the derivation of

Priority Monster; they are merely "along for the ride" and do not *contribute* to this

implausible implication. A defender of the Priority View might argue that she can

avoid Priority Monster without denying Tradeoffs, Parity, or the Conditional, if she

simply gives up the Premises, but maintains other claims of form **Ω**. It is, I think,

worth discussing this possibility.

Remember that the Premises are based off of the Series, which begins with an

arbitrarily large benefit for the currently worst off person, and ends with an arbitrarily

small benefit for an arbitrarily badly off person. It may help to see all of this again.

Here is the Series of possible benefits:

(1): an arbitrarily large benefit for the currently worst off person.

(2): a benefit slightly smaller than the one in (1) for a person much worse off than the

one in (1).

(3): a benefit slightly smaller than the one in (2) for a person much worse off than the

one in (2).

And so on… all the way up to:

(*n*): an arbitrarily small benefit for an arbitrarily badly off person.

And remember that there are four variables that can be adjusted, in arriving at a precise formulation of the Series: the largeness of (1), the tininess of (*n*), what counts as "slightly smaller," and what counts as "much worse off." Just as (1) can be arbitrarily large and (*n*) can be arbitrarily small, the size difference between adjacent benefits can be arbitrarily small and each person can be arbitrarily worse off than her predecessor. And finally recall the Premises:

P1. Benefit (2) is better than benefit (1).

P2. Benefit (3) is better than benefit (2).

  And so on… all the way up to,

P*n*. Benefit (*n*) is better than benefit (*n*-1).

I have been assuming that we should formulate the Series in a way that renders the Premises as plausible as possible. That is, I have assumed that the four variables mentioned above are sufficiently and relevantly "turned up." Thus, if we should reject all of the Premises in their strongest form, we should reject all of the claims of form Ω which are merely more controversial, or less intuitive, *versions* of the Premises. (E.g.,

a more controversial, or less intuitive, version of P1 would still claim that it is better to

achieve a smaller benefit for a person who is worse off than to achieve a larger benefit

for the currently worst off person, but the size difference between these benefits would

be larger and the degree to which the worse off person is worse off would be smaller).

Claims which are less intuitive versions of the Premises simply belong to different

versions of the Premises, and these different versions of the Premises still imply

Priority Monster. So these less intuitive versions of the Premises still face the

Infection Problem.

Since the Series begins with the currently worst off person and moves toward

arbitrarily bad lives, all claims of form Ω which apply to persons with lives at least as

bad as that lived by the currently worst off person belong to some version of the

Premises or another, and thus face the Infection Problem. But we have not yet seen

how the Infection Problem would impugn claims of form Ω which apply to persons

with lives which are *better* than that lived by the currently worst off person. These

claims obviously do not collectively imply Priority Monster, and none of them is part

of a conjunction of claims that implies Priority Monster. Before proceeding, there are

two terminological issues that need to be cleared up:

(1) When I say that a claim of form Ω "applies to" people at a particular well-being

level, what I mean is that this claim says that, rather than achieve a larger benefit

for a person at this well-being level, it would be better to achieve a smaller benefit

for a worse off person. There is a sense which the claim obviously applies to the

worse off person. But I mean for the people at a particular well-being level to

which the claim of form Ω applies to be the better off people under comparison.

(2) Any claim could be included into a conjunction of claims which also includes the

Premises, but it would not thereby be a member of the *relevant* conjunction which

faces the Infection Problem. Suppose that A1 & A2 & A3 imply C1, and that C1

is implausible. If there is equal reason to accept A1, A2, and A3, and these claims

are essential to the derivation of C1, then these claims face the Infection Problem.

It might be that B1 & B2 & A1 & A2 & A3 imply C1, but B1 and B2 are merely

"along for the ride" and do not contribute to the derivation of C1. B1 and B2 are

not members of the relevant conjunction – the conjunction of claims which are

essential to entailing C1. For example, consider the conjunction which includes

the Premises and the claim "it is sunny outside." The Infection Problem impugns

the Premises, but not the claim "it is sunny outside." From here onward, when I

refer to conjunctions of claims which imply various conclusions, I am only

referring to the *relevant* conjunctions. The relevant conjunctions are those whose

members are essential to the derivation of the conclusion in question).

While claims of form Ω which apply to persons with lives which are better than that

lived by the currently worst off person do not belong to a conjunction of claims that

implies Priority Monster, they might nonetheless belong to conjunctions which have

other implications, which are implausible. For example, recall Less Scary Priority

Monster (from Appendix D). If Less Scary Priority Monster were implausible, then claims of form Ω which apply to persons with lives which are better than that lived by the currently worst off person would face the Infection Problem. This is because the version of the Premises which implies Less Scary Priority Monster begins with persons who are arbitrarily well off. Indeed, the most plausible claims of form Ω which apply to persons with lives at *any* level of well-being would belong to a conjunction which implies Less Scary Priority Monster. If we found Less Scary Priority Monster to be implausible, we might then conclude that all claims of form Ω face the Infection Problem.

But again, we might not find Less Scary Priority Monster to be implausible. To determine which claims of form Ω face the Infection Problem, we should consider:

> *Priority Monster X.* There could be a person so badly off that, if she existed, it would be better to achieve an arbitrarily small benefit for her than to achieve an arbitrarily large benefit for *a person at well-being level X*.

X is a variable. If X is arbitrarily large, then we have Less Scary Priority Monster; if X corresponds to the well-being level of the currently worst off person, then we have Priority Monster. If we find Priority Monster implausible, but Less Scary Priority Monster not implausible, then as X moves from an arbitrarily high well-being level to a lower and lower one, Priority Monster X becomes progressively less and less plausible until it becomes implausible. Perhaps the point at which it transitions from

not implausible to implausible is vague, or there are many specifications of X such that it would be unclear or indeterminate whether Priority Monster X (at those specifications) is implausible. But let us suppose, for the sake of argument, that when X is greater than or equal to **W**, the Priority Monster X is not implausible. **W** is a well-being level much greater than that of the currently worst off person (given that Priority Monster is implausible).

A defender of the Priority View might then claim that she only accepts those claims of form Ω which apply to persons with lives which are at well-being level **W** or higher. We can call this the *Restricted Priority View*. Defenders of this view would not accept any claims of form Ω which belong to conjunctions which have implausible implications, and would thus avoid the Infection Problem.

I have two replies. The first is that for Priority Monster X to be not implausible, I believe that **W** would have to be fairly large. Perhaps it would have to be so large that most of us would fall below the relevant level of well-being. If so, the Restricted Priority View would have far less scope than defenders of the Priority View would have hoped. Moreover, it would seem odd if the only people the Priority View did apply to were those who are better off than most of us. This is my first, and somewhat more rhetorical, reply.

The second reply is that, even though the Restricted Priority View avoids the Infection Problem, and can avoid Priority Monster (as well as other versions of Priority Monster X which are implausible) without denying Tradeoffs, Parity, or the Conditional, it is implausible for other reasons. It is important to notice the way in

which the Restricted Priority View is consistent with the Conditional. In particular, defenders of this view need not deny Parity. They can acknowledge that claims of form Ω which apply to people with well-being levels below **W** are just as intuitively plausible as claims of form Ω which apply to people with well-being levels above **W**. But they can correctly add that the Conditional does not imply that if they accept the latter claims they should also accept the former claims, and they can correctly add that this is because the Conditional's "other things equal" clause is not satisfied in this case. It is not satisfied because the former claims but not the latter claims face the Infection Problem. But the Restricted Priority View still seems implausible. It implies that:

> (1) it is better to achieve a slightly smaller benefit for a much worse off person than to achieve a slightly larger benefit for a person just (non-vaguely) above **W**.

But it does *not* imply:

> (2) it is better to achieve a slightly smaller benefit for an arbitrarily badly off person than to achieve a slightly larger benefit for a person just (non-vaguely) below **W**.

We can suppose that the distance between the well-being levels of the people in (2) is far greater than the distance between the well-being levels of the people in (1) (which we can make as large as we like).

Not only is (2) at least as intuitive as (1), but it seems that whatever considerations or factors could *make* (1) true would also make (2) true. In this sense, (1) and (2) are the same *kind* of claim. What consideration or factor would make (1) true? The intuitively plausible answer is simply that it is better to achieve a slightly smaller benefit for a much worse off person than achieve a slightly bigger benefit for someone better off. It does not seem plausible that the factor that would make (1) true is that it is better to achieve a slightly smaller benefit for worse off person than achieve a slightly bigger benefit for someone better off *and above **W***. The latter italicized bit does *not* seem to be an essential part of what, if anything, makes (1) true. And if not, then what does seem to make (1) true would also make (2) true.

The Conditional is a very modest claim. We could revise it so that it is slightly bolder:

> *Conditional**: if we accept claim C1 on the basis of its intuitive plausibility, and if claim C2 is at least as intuitively plausible as C1, and if these two claims are *relevantly of the same kind*, then we should also accept C2.

This revised version of the Conditional replaces the "other things equal" clause with the "relevantly of the same kind" clause. The Conditional* applies even when other

things are not equal. Other things might not be equal because, e.g., C1 faces the Infection Problem, whereas C2 might not. But for all that, C1 and C2 might be claims which are relevantly of the same kind. That is, it might be that whatever considerations or factors could *make* C1 true would also make C2 true. If so, the Conditional* would imply that if C1 is accepted on the basis of its intuitive plausibility, then C2 should also be accepted if it is at least as intuitively plausible. And so in cases where we know C2 is false, we would be required not to accept C1.

Earlier I said that the idea behind the Conditional is that we should treat like cases alike, and be consistent about our standards of acceptance. But the Conditional only rules out the most egregious failures to treat like cases alike. It does not rule out the Restricted Priority View. However, the Restricted Priority View is ruled about by a somewhat bolder requirement, the Conditional*. Perhaps the Restricted Priority View is not guilty of the most egregious sort of failure to treat like cases alike. But it still does seem to fail, in an important way, to treat like cases alike. And so it seems implausible.

Lastly, recall that since defenders of the Equal Weight View do not accept any claim of form Ω, the Conditional does not imply that they should accept the Premises. It is not a failure to treat like cases alike to, e.g., accept the claim that other things equal bigger benefits are better on the basis of its intuitiveness, but not to accept claims of form Ω on the basis of their intuitiveness. We can now note that this would not be a failure to treat like cases alike *not only* because the claim that other things equal bigger benefits are better avoids the Infection Problem, *but also* because this

claim is a relevantly different kind of claim than claims of form **Ω**. The factors or considerations which would make it true seem importantly distinct from the factors or considerations which would make claims of form **Ω** true. So the Conditional* would not imply that people who accept the claim that other things equal bigger benefits are better on the basis of its intuitiveness should also accept claims of form **Ω**, even if the latter claims were no less intuitively plausible.

For purely pragmatic reasons, I will continue to say that the Priority View cannot avoid Priority Monster unless it denies either Tradeoffs, Parity, or the Conditional. However, the truth is that the Restricted Priority View can avoid Priority Monster without denying any of these three latter claims. But, as I just explained, the Restricted Priority View does deny the Conditional*, and it seems implausible to do so.

References

Adler, Matthew. 2009. "Future Generations: A Prioritarian View." *George Washington Law Review* 77: 1478–1520.

———. 2011. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.

Arneson, Richard. 1999. "What, If Anything, Renders All Humans Morally Equal?" In *Peter Singer and His Critics*, ed. by Dale Jamieson. Oxford: Blackwell.

———. 2000. "Luck Egalitarianism and Prioritarianism." *Ethics* 110: 339–49.

———. 2007. "Desert and Equality." In Nils Holtug and Kasper Lippert-Rasmussen, eds, *Egalitarianism: New Essays on the Nature and Value of Equality*, 262–293. Oxford: Oxford University Press.

———. 2010. "Good, Period." *Analysis* 70: 731–44.

Arntzenius, Frank, Elga, Adam, and Hawthorne, John. 2004. "Bayesianism, Infinite Decisions, and Binding." *Mind* 113: 251–283.

Arrhenius, Gustaf. 2012. *Population Ethics*. Oxford: Oxford University Press.

Audi, Robert. 2004. *The Good in the Right*. Princeton: Princeton University Press.

Berker, Selim. 2009. "The Normative Insignificance of Neuroscience." *Philosophy and Public Affairs* 37: 293–329.

Brink, David. 1993. "The Separateness of Persons, Distributive Norms, and Moral Theory." In *Value, Welfare and Morality*, ed. R. Frey and C. Morris, 252–289. New York: Cambridge University Press.

———. 1997. "Rational Egoism and the Separateness of Persons." In *Reading Parfit*, ed. J. Dancy. Oxford: Blackwell.

———. 2011. "Prospects for Temporal Neutrality." In *The Oxford Handbook of Time*, ed. C. Callender. New York: Oxford University Press.

Broome, John. 1991. *Weighing Goods*. Oxford: Basil Blackwell.

———. 1992. *Counting the Cost of Global* Warming. White Horse Press.

———. 2002. "Respects and levelling down." Presented at the European Congress of Analytical Philosophy, Lund.

———. 2004. *Weighing Lives*. Oxford: Oxford University Press.

———. 2012. *Climate Matters*. Norton.

Brown, Campbell. 2005. PhD Dissertation. Australian National University.

———. 2007. "Prioritarianism for Variable Populations." *Philosophical Studies* 134: 325–361.

Carlson, Erik. 2000. "Aggregating Harms – Should We Kill to Avoid Headaches?" *Theoria* 66: 246–55.

———. 2005. "Intransitivity without Zeno's Paradox." In *Recent Work on Intrinsic Value*, M. Zimmerman and T. Rønnow-Rasmussen (Eds.), Dordrecht: Springer.

Chang, Ruth. 1997. *Incommensurability, Incomparability, and Practical Reason*. Cambridge: Harvard University Press.

———. 2002. "The Possibility of Parity." *Ethics* 112: 659–88.

Christensen, David. 2007. "Epistemology of Disagreement: The Good News." *Philosophical Review*. 116: 187–217.

Cooper, Neil. 1995. "Paradox lost: Understanding vague predicates." *International Journal of Philosophical Studies* 3: 244–269.

Crisp, Roger. 2003. "Equality, Priority, and Compassion." *Ethics* 113: 745–63.

———. 2006a. "Hedonism Reconsidered." *Philosophy and Phenomenological Research* 3: 619-45.

———. 2006b. *Reasons and the Good*. Oxford: Oxford University Press.

———. 2011. "In Defence of the Priority View: A Response to Otsuka and Voorhoeve." *Utilitas* 23: 105–8.

Cureton, Adam. 2009. "Moral Intuitions about Large Numbers." Presented at the Southern Society for Philosophy and Psychology Conference, Savannah, GA.

Dancy, Jonathan. 1993. *Moral Reasons*. Oxford: Blackwell.

Davidson, Donald, McKinsey, J. C. C., and Suppes, Patrick. 1955. "Outlines of a Formal Theory of Value." *Philosophy of Science* 22: 140–60.

DeGrazia, David. 1996. *Taking Animals Seriously*. Cambridge: Cambridge University Press.

Dreier, James. 1993. "Structures of Normative Theories." *The Monist* 76: 22–40.

Feldman, Fred. 1997. *Utilitarianism, Hedonism, and Desert: Essays in Moral Philosophy*. New York: Cambridge University Press.

Fischer, John. 2012. "Immortality." In *The Oxford Handbook of Philosophy of Death*, edited by Ben Bradley, Fred Feldman, and Jens Johansson. Oxford: Oxford University Press.

Foot, Philippa. 2002. *Moral Dilemmas*. Oxford: Oxford University Press.

Frankfurt, Harry. 1987. "Equality as a Moral Ideal." *Ethics* 98: 21–42.

Goodman, Nelson. 1983. *Fact, fiction, and forecast*. Cambridge: Harvard University Press.

Greene, Joshua. 2001. "A psychological perspective on Nozick's experience machine and Parfit's repugnant conclusion." Presentation at the Society for Philosophy and Psychology, Annual Meeting, Cincinnati, OH.

———. n.d. "Notes on 'The Normative Insignificance of Neuroscience' by Selim Berker."

Greene, Joshua, and Baron, Jonathan. 2001. "Intuitions about declining marginal utility." *Journal of Behavioral Decision Making* 14: 243–255.

Griffin, James. 1986. *Well-being*. Oxford: Clarendon Press.

Hare, Caspar. 2008. "A Puzzle About Other-Directed Time Bias." *Australasian Journal of Philosophy* 86: 269–277.

———. 2012. "Obligations to Merely Statistical People." *Journal of Philosophy* 109: 378–390.

Hare, R. M. 1981. *Moral Thinking*. Oxford: Oxford University Press.

Harsanyi, John. 1955. "Cardinal Welfare, Individual Ethics and Interpersonal

Comparisons of Utility." *Journal of Political Economics* 63: 309–321.

Holtug, Nils. "Utility, Priority, and Possible People." *Utilitas* 11: 16–36.

———. 2010. *Persons, Interests, and Justice*. Oxford: Oxford University Press.

Hooker, Brad. 2000. *Ideal Code, Real World*. Oxford: Oxford University Press.

Huemer, Michael. 2005. *Ethical Intuitionism*. Palgrave Macmillan.

———. 2008a. "Revisionary Intuitionism." *Social Philosophy and Policy* 25: 368–92.

———. 2008b. "In Defence of Repugnance." *Mind* 117: 899–933.

———. 2013. "Transitivity, Comparative Value, and the Methods of Ethics." *Ethics* 123: 318–45.

Hurka, Thomas. 1993. *Perfectionism*. Oxford: Oxford University Press.

Kagan, Shelly. 1988. "The Additive Fallacy." *Ethics* 99: 5–31.

———. 1989. *The Limits of Morality*. Oxford: Oxford University Press.

———. 1998. *Normative Ethics.* Westview.

———. 1999. "Equality and Desert." In Louis Pojman and Owen McLeod, eds. *What Do We Deserve? A Reader on Justice and Desert*, 298–314. New York: Oxford University Press.

———. 2012. *The Geometry of Desert*. New York: Oxford University Press.

Kamm, Frances. 1983. "Killing and Letting Die: Methodological and Substantive Issues." *Pacific Philosophical Quarterly* 64: 297–312.

———. 2007. *Intricate Ethics*. New York: Oxford University Press.

———. 2009. "Disability, discrimination, and irrelevant goods." In K. Brownlee & A. Cureton (eds) *Disability and Disadvantage*, 260–318. NewYork: Oxford University Press.

Keefe, Rosanna. 2000. *Theories of Vagueness*. Cambridge: Cambridge University Press.

Lewis, David. 1971. "Counterparts of Persons and Their Bodies." *Journal of Philosophy* 68: 203–11.

———. 1976. "Survival and Identity." In *The Identities of Persons*, ed. Amélie Oksenberg Rorty, Berkeley, University of California Press.

McKerlie, Dennis. 1989. "Equality and Time." *Ethics* 99: 274–96.

———. 1997. "Priority and Time." *Canadian Journal of Philosophy* 27: 287–309.

McMahan, Jeff. 1996. "Cognitive Disability, Misfortune, and Justice." *Philosophy and Public Affairs*, 25: 3–35.

———. 2000. "Moral Intuition." In the *Blackwell Guide to Ethical Theory*, H. LaFollette (ed.). Oxford: Blackwell.

———. 2002. *The Ethics of Killing*. New York: Oxford University Press.

McTaggart, J. M. E. 1921. *The Nature of Existence*, Volume One. Cambridge: Cambridge University Press.

Mill, John Stuart. 1998 [1861]. *Utilitarianism*, ed. Roger Crisp. Oxford: Oxford University Press.

Moore, G. E. 1988 [1903]. *Principia Ethica*. Amherst, New York: Prometheus Books.

Norcross, Alastair. 1997. "Comparing Harms: Headaches and Human Lives." *Philosophy and Public Affairs* 26: 135–67.

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

———. 1993. *The Nature of Rationality*. Princeton: Princeton University Press.

Otsuka, Michael. 2004. "Skepticism about Saving the Greater Number." *Philosophy and Public Affairs* 32: 413–26.

Otsuka, Michael and Voorhoeve, Alex. 2009. "Why It Matters That Some Are Worse Off Than Others: An Argument Against the Priority View." *Philosophy and Public Affairs* 37: 171–99.

———. 2011. "Reply to Crisp." *Utilitas* 23: 109–114.

Parfit, Derek. 1984. *Reasons and Persons*. Oxford: Oxford University Press.

———. 1986a. "Comments." *Ethics* 96: 832–872.

———. 1986b. "Overpopulation and the quality of life." In *Applied Ethics*, ed. Peter Singer, 145–164. Oxford: Oxford University Press.

———. 1991. "Equality or Priority?" *The Lindley Lecture*. Lawrence: University of Kansas; reprinted in 2000 in *The Ideal of Equality*, ed. Matthew Clayton and Andrew Williams, 81–125. London: Macmillan.

———. 2011. *On What Matters*, Volume One. Oxford: Oxford University Press.

———. 2012. "Another Defence of the Priority View." *Utilitas* 24: 399–440.

———. n.d. "Towards Theory X."

Pollock, John and Cruz, Joseph. 1999. *Contemporary Theories of Knowledge,* 2d ed. Lanham, MD: Rowman and Littlefield.

Portmore, Douglas. 2007. "Consequentializing Moral Theories." *Pacific Philosophical Quarterly* 88: 39–73.

Prichard, H. A. 1957. "Does Moral Philosophy Rest on a Mistake?" In Prichard, *Moral Obligation.* Oxford: Clarendon Press.

Pummer, Theron. 2013. "Intuitions about Large Number Cases." *Analysis* 73: 37–46.

Qizilbash, Mozaffar. 2005. "Transitivity and Vagueness." *Economics and Philosophy* 21: 109–131.

Quinn, Warren. 1990. "The Puzzle of the Self-Torturer." *Philosophical Studies* 59: 79–90.

Rabinowicz, Wlodek. 2000. "Money Pump with Foresight." In Michael J. Almeida (ed.), *Imperceptible Harms and Benefits*, 123–154. Dordrecht: Kluwer.

———. 2002. "Prioritarianism for Prospects." *Utilitas*: 2–21.

Rabinowicz, Wlodek and Broome, John. 1999. "Backwards Induction in the Centipede Game." *Analysis* 59: 237–242.

Rachels, Stuart. 1998. "Counterexamples to the Transitivity of Better Than." *Australasian Journal of Philosophy* 76: 71–83.

———. 2004. "Six Theses about Pleasure." *Philosophical Perspectives* 18: 247–267.

Raffman, Diana. 1994. "Vagueness Without Paradox." *Philosophical Review* 103: 41–74.

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

Raz, Joseph. 1986. *The Morality of Freedom*. Oxford: Clarendon Press.

Rosati, Connie. 2012. "The Makropulos Case Revisited: Reflections on Immortality and Agency." In *The Oxford Handbook of Philosophy of Death*, edited by Ben Bradley, Fred Feldman, and Jens Johansson. Oxford: Oxford University Press.

Ross, W. D. 1988. *The Right and the Good*. Indianapolis, IN: Hackett.

Ryberg, J. and T. Tännsjö, (eds.). 2004. *The Repugnant Conclusion: Essays on Population Ethics*. Kluwer Academic Publishers.

Schroeder, Mark. 2007. "Teleology, Agent-Relative Value, and 'Good'." *Ethics* 117: 265–295.

Sen, Amartya. 1980 [1979]. "Equality of what?" in *The Tanner Lectures on Human Values*, Vol. 1, ed. Sterling McMurrin, 197–220. Cambridge: Cambridge University Press.

———. 1979. "Utilitarianism and Welfarism." *Journal of Philosophy* 76.

Sidgwick, Henry. 1907. *The Methods of Ethics*, seventh edition. London: Macmillan.

Singer, Peter. 1975. *Animal Liberation*. New York Review/Random House.

———. 2011. *Practical Ethics*, third edition. New York: Cambridge University Press.

Steinbock, Bonnie. 1978. "Speciesism and the Idea of Equality." *Philosophy* 53: 247–256.

Sumner, Wayne. 1996. *Welfare, Happiness, and Ethics*. New York: Oxford University Press.

Temkin, Larry. 1983. PhD Dissertation. Princeton University.

———. 1993. *Inequality*. New York: Oxford University Press.

———. 1996. "A Continuum Argument for Intransitivity." *Philosophy and Public Affairs* 23: 175–210.

———. 2000. "Equality, Priority, and the Levelling Down Objection." In *The Ideal of Equality*, ed. Matthew Clayton and Andrew Williams, 126–61. London: Macmillan.

———. 2012. *Rethinking the Good*: *Moral Ideals and the Nature of Practical Reasoning*. New York: Oxford University Press.

Thomson, Judith. 1971. "In Defense of Abortion." *Philosophy and Public Affairs* 1: 47–66.

———. 1992. *The Realm of Rights*. Cambridge: Harvard University Press.

———. 2001. *Goodness and Advice*. Princeton: Princeton University Press.

———. 2008. *Normativity*. Chicago and La Salle, Illinois: Open Court.

Tooley, Michael. 1972. "Abortion and Infanticide." *Philosophy and Public Affairs* 2: 37–65.

Unger, Peter. 1996. *Living High and Letting Die*. New York: Oxford University Press.

Vallentyne, Peter. 2005. "Of Mice and Men: Equality and Animals." *Journal of Ethics*, 9: 403–433.

Velleman, David. 1991. "Well-Being and Time." *Pacific Philosophical Quarterly* 72: 48–77.

———. 2008. "Persons in Prospect." *Philosophy and Public Affairs* 36: 221–288.

Voorhoeve, Alex. 2008. "Heuristics and Biases in a Purported Counterexample to the Acyclicity of 'Better Than'." *Politics, Philosophy and Economics* 7: 285–99.

———. 2009. *Conversations on Ethics*. Oxford: Oxford University Press.

Voorhoeve, Alex and Binmore, Ken. 2003. "Defending Transitivity Against Zeno's

Paradox." *Philosophy and Public Affairs* 31: 272–9.

Williams, Bernard. 1978. "The Makropulos Case: Reflections on the Tedium of Immortality." In *Language, Metaphysics, and Death*, edited by John Donnelly, 229–242. New York: Fordham University Press.

Williamson, Timothy. 1994. *Vagueness*. New York: Routledge.