'Making People Happy, Not Making Happy People': A Defense of the Asymmetry Intuition in Population Ethics

Citation

Frick, Johann David. 2014. 'Making People Happy, Not Making Happy People': A Defense of the Asymmetry Intuition in Population Ethics. Doctoral dissertation, Harvard University.

Permanent link

http://nrs.harvard.edu/urn-3:HUL.InstRepos:13064981

Terms of Use

Share Your Story

'Making People Happy, Not Making Happy People':

A Defense of the Asymmetry Intuition in Population Ethics

A dissertation presented

by

Johann David Anand Frick

to

The Department of Philosophy

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Philosophy

Harvard University

Cambridge, Massachusetts

September 2014

Dissertation Advisors: Professor T.M. Scanlon   Author: Johann Frick
                    Professor Frances Kamm

'Making People Happy, Not Making Happy People':
A Defense of the Asymmetry Intuition in Population Ethics

Abstract

This dissertation provides a defense of the normative intuition known as the

Procreation Asymmetry, according to which there is a strong moral reason *not* to create

a life that will foreseeably not be worth living, but there is no moral reason to create a

life *just because* it would foreseeably be worth living.

Chapter 1 investigates how to reconcile the Procreation Asymmetry with our

intuitions about another recalcitrant problem case in population ethics: Derek Parfit's

Non-Identity Problem. I show that what has prevented philosophers from developing a

theory that gives a satisfactory account of *both* these problems is their tacit commitment

to a teleological conception of well-being, as something to be 'promoted'. Replacing this

picture with one according to which our reasons to confer well-being on people are

conditional on their existence allows me to do better. It also enables us to understand

some of the deep structural parallels between seemingly disparate normative

phenomena such as procreating and promising.

Chapter 2 attempts to connect my defense of the Procreation Asymmetry to

corresponding evaluative claims about the goodness of the outcomes produced by

procreative decisions. I propose a view, the 'biconditional buck-passing view of outcome

betterness', according to which facts about the comparative goodness of outcomes are a function of our reasons for bringing about one outcome rather than another under certain conditions. This enables me to derive an Evaluative Procreation Asymmetry from the corresponding normative claims established in Chapter 1. The biconditional buck-passing view also provides me with a principled basis for challenging a version of the Independence of Irrelevant Alternatives Principle. This, in turn, permits me to provide a novel solution to another famous problem in population ethics: Parfit's Mere Addition Paradox.

Finally, in Chapter 3, I rebut some key objections to the Procreation Asymmetry by showing that upholding it does not commit us to anti-natalism and that it is compatible with a moral concern for the long-term survival of humanity.

Table of Contents

Acknowledgements

My first and foremost thanks goes to the three members of my dissertation committee: Tim Scanlon, Frances Kamm, and Derek Parfit. An aspiring moral philosopher could not dream of a better and more devoted team of mentors than I have been privileged to work with over the past years. I am deeply grateful for their guidance and encouragement, their patience and constructive criticism, but most of all for setting, in their own work, an example of integrity and rigor, of imagination and wisdom that I can only strive to emulate in the future.

For illuminating conversations on the subject matter of this thesis, as well as for invaluable feedback on my own ideas at various stages of the dissertation process, I thank Gustaf Arrhenius, Ralf Bader, Nick Beckstead, Elizabeth Harman, Iwao Hirose, Christine Korsgaard, Jeff McMahan, Melinda Roberts, and Larry Temkin.

Eric Beerbohm, Selim Berker, Nir Eyal, David Langlois, Doug Lavin, Michael Otsuka, Julian Savulescu, Michael Smith, Gerard Vong, Ralph Wedgwood, and Gabriel Wollner, all gave me extremely helpful comments on earlier papers that grew into this dissertation.

I am greatly indebted to the students in my 2014 graduate seminar on population ethics at Princeton for their many excellent questions and comments, and for pushing me to clarify my ideas on many issues. In particular, I thank Annie Fang, Rishi Joshi, Adam Lerner, Nat Tabris, and Daniel Wodak.

Chapter 1: Conditional Reasons and the Procreation Asymmetry

1. Introduction

This essay sketches a theory of the reason-giving force of well-being that promises to reconcile our intuitions about two of the most recalcitrant problem cases in population ethics: Jan Narveson's *Procreation Asymmetry*[1] and Derek Parfit's *Non-Identity Problem*.[2] I show that what has prevented philosophers from developing a theory that gives a satisfactory account of *both* these problems is their tacit commitment to a teleological conception of well-being, as something to be 'promoted'. Replacing this picture with one according to which our reasons to confer well-being on people are *conditional* on their existence allows me to do better. It also enables us to understand some of the deep structural parallels between seemingly disparate normative phenomena such as procreating and promising.

---

[1] The Asymmetry was first discussed by Jan Narveson in "Utilitarianism and New Generations", *Mind* 76 (1967), pp. 62–72. The label is due to Jeff McMahan, "Problems of Population Theory", *Ethics*, 92 (1981), pp. 96-127.

[2] Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), Ch. 16.

2. The Procreation Asymmetry

Many of us hold pre-theoretical views about the morality of procreation that are, in an important sense, asymmetrical. Suppose you can foresee that any child you could create would live a life so full of uncompensated suffering as to be not worth living.[3] Most would agree that – exceptional circumstances aside – it would constitute a serious moral wrong to bring such a person into existence. Next, imagine that the child you could create would have a life that would be well worth living. Many have the intuition that it is morally *indifferent* whether you decide to create this person or not. All else equal, it seems permissible for you to have the child if you wish; but we don't believe that you act contrary to strong moral reasons, let alone a moral obligation, if you decide not to.

Let us distinguish two different ways of fleshing out this Asymmetry intuition. To my mind, the most natural way of formulating the Asymmetry is in normative terms. Thus, according to the

> *Normative Procreation Asymmetry*:
> (1a) If a future person would foreseeably have a life that is not worth living, this in itself gives us a strong moral reason to refrain from bringing this person into existence. By contrast, (2a) there is no moral

---

[3] More precisely, imagine that such a child would have a life that is, in Derek Parfit's phrase, "worth not living", since its life would be *worse* than a life spent in a permanent coma (which would also be not worth living). For stylistic reasons, I will continue to use the former locution. However, you may assume throughout that when I refer to a life as "not worth living", this is also a life that is "worth not living", in Parfit's sense.

reason to create a person whose life would foreseeably be worth living, just *because* her life would be worth living.[4]

However, much of the literature in population ethics is written by consequentialists or those friendly to the consequentialist perspective, who regard evaluative categories such as *good*, *value*, *better*, *worse*, etc. as *prior* to normative categories such as *reason*, *ought*, *right*, *obligation*, etc. In discussing these thinkers, it will sometimes be convenient to characterize the Asymmetry in evaluative terms, as a claim about the goodness of the outcome that results from an act of procreation. Thus, according to the

> *Evaluative Procreation Asymmetry*:
> (1b) It makes the world go worse, all else equal, to create a life not worth living. By contrast, (2b) it does not make the world go better, all else equal, to create a life worth living.

My aim in this chapter is to provide a defense of the Normative Procreation Asymmetry. In Chapter 2 of this dissertation, I will argue in support of the Evaluative Procreation Asymmetry, and claim that it is true *because* the Normative Asymmetry is true.[5]

---

[4] In the essay I do not take a stand on what, exactly, makes a life worth living or not worth living for a person. That is, I shall remain agnostic with regard to the correct theory of well-being (which I understand as that which makes a person's life go well or, at least, worth living). The Asymmetry, I argue, is true whichever theory of well-being we have most reason to endorse. In keeping with established practice in the literature on the Asymmetry, I will sometimes employ the phrases "a happy life" and "a miserable life" as synonyms for "a life worth living" and "a life not worth living". But, again, "happy" and "miserable" should here be understood in a sense that is agnostic with regard to the true theory of well-being.

[5] There are various ways of qualifying these two Asymmetry claims. For instance, we may think that while we do have pro tanto moral reasons to create a life worth living, these reasons are *much weaker* than our reasons both against creating a life worth not living, and for conferring equivalent benefits on already existing people (for example by extending their lives). Call this the *Weak Normative Asymmetry*. (The

With few exceptions, those who reject either the Normative or the Evaluative

Asymmetry accept the first conjunct of these claims – (1a) or (1b) – but deny the second –

(2a) or (2b). Instead, they affirm (normative version) that we *do* have significant moral

reasons to create new people whose lives would be worth living *because* they would be

worth living, or (evaluative version) that by creating new lives worth living, we *do* make

the world go better, all else equal. I shall refer to these opponents of the Asymmetry

intuition as holding a *symmetry* view. The paradigmatic proponents of a symmetry view

are totalist act-utilitarians, who believe that in each of our actions, we have decisive

moral reason to do what makes the world go best, namely to produce the greatest

possible net aggregate of well-being over ill-being.

The aim of this chapter is not to present a knock-down argument against such

symmetry views. Given the very fundamental nature of this disagreement, such an

argument would be hard to come by. More importantly, in one sense such an argument

is not *needed*. The Asymmetry strikes many people – even some of those who have

---

Evaluative Asymmetry can be similarly weakened). The Asymmetry can also be restricted in *scope*. Instead of applying to all lives worth living, we may think the Asymmetry claims hold only for lives that are worth living, but whose level of lifetime well-being falls within a certain *neutral range* that is upwardly bounded. If we could create a person whose level of lifetime well-being *exceeds* the neutral range, we would have pro tanto moral reason to bring this person into existence (and doing so would make the world better, all else equal). In this dissertation, I shall concern myself primarily with the unqualified version of the Normative and Evaluative Asymmetry. The defense I offer of these claims could be adapted to cover the qualified versions as well.

opposed it in print – as intuitively highly plausible. For instance, John Broome, after almost fifteen years of arguing against the intuition, confesses to still being gripped by it:

> We [intuitively] care about the well-being of people who exist; we want their well-being to be increased. If it is increased, an effect will be that there will be more well-being in the world. But we do not want to increase the amount of well-being in the world for its own sake. A different way of achieving that result would be to have more people in the world, but most of us are not in favor of that. We are not against it either; we are neutral about the number of people.[6]

Furthermore, common-sense ethical thought rejects many of the strongly revisionary implications that symmetry views would appear to have. Again I quote from Broome:

> When people's lives are saved, by making roads safer or in other ways, the well-being of the people who are saved is generally small in comparison to the well-being of all the new people, their descendants, who come into existence as a result. This is perfectly predictable. If all the descendants' well-being had to be counted too, that would enormously alter the value we attach to saving people's lives. But actually, in judging the value of safety on the roads, we routinely ignore all this well-being.[7]

Finally, upholding the Asymmetry intuition also has important theoretical payoffs within the field of population ethics, since rejecting (2a) or (2b) invites Parfit's *Repugnant Conclusion*, insofar as this opens up the possibility that adding sufficiently many lives that are *barely* worth living to a world can morally outweigh a reduction in the well-being of an original population, in which everyone was very well off.

---

[6] John Broome, *Climate Matters: Ethics for a Warming World* (New York: Norton, 2012), p. 171.

[7] "Should We Value Population?", *The Journal of Political Philosophy*, Vol. 13, No. 4 (2005): 399-413, p. 402.

What has stood in the way of more widespread acceptance of the Asymmetry amongst population ethicists is that, to date, all attempts at fleshing out the intuition have encountered serious difficulties. As we shall see in the following section, one of the biggest stumbling blocks is offering an account of the Asymmetry that does not commit us to an implausible position regarding Derek Parfit's Non-Identity Problem.

Besides avoiding such counterintuitive implications, I submit that a successful account of the Asymmetry Intuition must, at the least, offer a non-question begging response to two basic challenges. Consider first what I shall term the *Objection from Benefit:* We all recognize that we have pro tanto moral reasons to benefit *existing* people by doing what is good for them – for instance saving their lives, which allows them to live longer. Given this, shouldn't we recognize a corresponding moral reason to benefit future people, by bringing them into existence with a life worth living?

Note that proponents of this Objection from Benefit need not be committed to the claim that it is *better* (or worse) *for* a person to be brought into existence. This claim has struck many philosophers as incoherent, since, they contend, it implies that it would have been *worse for* that person never to have existed. But this, it is argued, cannot be the case. If the individual in question never exists, there *is* no person *for whom* non-existence is worse, and consequently no-one for whom existence would have been better. The

comparandum lacks a subject.[8] (Recall the old Yiddish joke: "Life is so terrible, it would be better never to have been born." Response: "Who is so lucky? Not one in ten thousand!" The joke works, because people who never exist can be neither the subjects of fortune or misfortune).

There is, however, a way of putting the Objection from Benefit that does not encounter such conceptual difficulties. As Jeff McMahan points out, it is both coherent and plausible that being caused to exist with a life worth living can be good for a person in a *non-comparative* sense, namely insofar as "the intrinsically good elements of the person's life more than compensate for the intrinsically bad elements."[9] This can be true despite the fact that the outcome in which he does not exist would not have been bad, or worse, for him. *Mutatis mutandis*, it can be non-comparatively bad for a person to be brought into existence with a life that is overall not worth living, despite the fact that never existing would not have been good, or better, for him. Because bringing a person into existence can be good or bad for her in this non-comparative sense, it is also plausible to speak of acts of procreation as "benefiting" and "harming" those whom they create. (McMahan calls these "existential" benefits and harms, in contrast with

[8] For versions of this argument, see Krister Bykvist, "The Benefits of Coming into Existence", *Philosophical Studies* (2007), pp. 95-115 and Jeff McMahan, "Causing People to Exist and Saving People's Lives", *The Journal of Ethics* (2013), pp. 5-35. For a dissenting view, see Nils Holtug, "On the Value of Coming into Existence", *Journal of Ethics* (2001), pp. 361-384.

[9] "Causing People to Exist and Saving People's Lives", p. 6.

"ordinary" benefits and harms, "which are bestowed or inflicted on existing people, or on future individuals whose existence is independent of the act that causes or constitutes the benefit or harm."[10]).

The question raised by the Objection from Benefit, then, is this: If what is good for existing people (ordinary benefits) makes a moral claim on us, why not also what is good for potential future persons (existential benefits)?

Fully grasping the Objection from Benefit will immediately block one mistaken way of thinking about the Asymmetry, namely to view it as just an instance of a more general moral asymmetry between the strength of our reasons not to harm and the strength of our reasons to benefit other people. Those who accept the Asymmetry believe, to paraphrase Jan Narveson's famous dictum, that while there are often weighty moral reasons to make (existing) people happy, there is no corresponding moral reason to make happy people. By contrast, note that there is *no* such intuitive asymmetry when speaking of *miserable* lives: Our moral reasons against creating new lives that are miserable seem just as weighty as our reasons against making existing lives miserable. Thus, even if there is a general moral asymmetry between the strength of our reasons not to harm and the strength of our reasons to benefit other people, this couldn't explain

---

[10] Ibid., p. 7.

why our moral reasons to benefit people by *creating* them are intuitively so much weaker than our reasons to provide ordinary benefits, or indeed non-existent.

Second, there is what I call the *Objection from Symmetry*: We accept that if causing a person to exist would foreseeably be *bad for* that person, because her life would be not worth living, this gives us moral reason *against* bringing her into being. By symmetry of reasoning, why does the fact that causing a person to exist would foreseeably be *good for* that person, because her life would be worth living, not give us moral reason *for* bringing her into being?

Neither the Objection from Benefit nor the Objection from Symmetry are, I believe, insuperable. Indeed, they should be understood, not as suggesting that a satisfactory account of the Asymmetry is impossible, but as defining minimal adequacy condition for such an account. On pain of appearing *ad hoc*, any plausible defense of the Asymmetry must be capable of responding to at least these two objections.

3. The Asymmetry Versus the Non-Identity Problem

Totalist utilitarianism, as the name suggests, holds that that the world goes better, the greater the total sum of positive minus negative well-being it contains. In pursuing this goal, the potential well-being of existing people and of possible future people matters equally and for the same reason: Our moral reason to add X units of positive well-being to the world by creating a new life is just as strong as our reason to increase the well-being of an already existing life by X units; after all, both actions will have the morally

equivalent result of increasing the total sum of well-being in the world by the same amount.

The past three decades have seen a plethora of work in population ethics aimed at replacing totalism with a more plausible view. None of these attempts have proved fully satisfactory, however, because they end up with commitments that contradict either one of the conjuncts of the Asymmetry Intuition or the intuitively correct verdict in Derek Parfit's Non-Identity Problem. Consider the following illustration of this problem:

> *Non-Identity Case*
> A woman has decided to have a child. Depending on when exactly she procreates, she knows that she will either
> (A) create person X, with a moderately happy life
> or
> (B) create a numerically non-identical person Y, with a very happy life. Assume that it is easy and costless for the woman to choose option (B) instead of option (A).

The judgment that almost everyone has about this case is that the woman has a strong moral reason to choose (B) over (A). This is true, despite the fact that picking (A) would not be bad for, or worse, for X, nor picking (B) – for reasons discussed in the previous section – better for Y. Indeed, the woman's reasons for choosing (B) over (A), we think, are comparable in strength to her reasons for picking option (C) over option (A) in the following *Same-Person Case*, where doing (A) *would* be worse for some person[11]:

---

[11] Parfit endorses a stronger claim, the *No Difference View*, according to which there is *no morally relevant difference* between a same-person choice, in which choosing option (A) is worse for X, and a different person

Same-Person Case
A woman has conceived a child. Depending on whether she undergoes a simple and costless procedure during her pregnancy, she knows that she will either

    (A) create person X, with a moderately happy life

or

    (C) create person X, with a very happy life.

Space constraints do not permit me to undertake a comprehensive survey of the alternatives to totalism in the literature. Three representative examples will have to suffice. What I am trying to bring out in discussing these examples is that, despite their differences, many existing alternatives to totalism share the same basic strategy: Totalism holds that the potential well-being of *all* possible people is reason-giving. By contrast, these alternative views divide up possible people according to their temporal location or their modal status, and assign different reason-giving force to the well-being of members in these various groups. *Presentists* draw a distinction between presently existing and presently non-existing people, and claim that only the well-being of presently existing people provides us with any moral reasons; *necessitarians* distinguish between people (present or future) who exist or will exist no matter how we decide to

---

choice, like *Non-Identity Case*, in which choosing (A) is not worse for anybody, provided the two actions have equivalent overall effects. (*Reasons and Persons*, pp. 366-369). We need not, however, accept the No Difference View. Thus, it might be argued that one difference between the two cases is that while in Same Person Case, choosing option (A) *wrongs* a person, choosing (A) in Non-Identity case wrongs no-one, and that this difference affects the comparative strength of our reasons against choosing option (A) in either case. For a development of this point, see my "Future Persons and Victimless Wrongdoing" (ms).

act, and people whose existence is contingent on our decisions; *actualists* separate people who exist or who are going to exist in the actual world, on the one hand, from people who don't, and won't exist, on the other. In each instance, we will see how the very feature that allows each of these accounts to capture one or both conjuncts of the Asymmetry precludes it from giving a satisfactory of the Non-Identity Problem.

Consider first the 'generocentric' view of David Heyd. This view

> (…) takes the present generation, viz. that making the demographic choice, as the only relevant group to which moral considerations are applicable. (…) The generocentric approach grants a moral standing only to those who 'generate' population growth, excluding such a standing from those who are 'generated', despite the fact that once they are 'generated' beings of this type are the same as their 'generators' and accordingly enjoy a moral status. In other words, decisions to enlarge the moral community are taken only 'from within' and in the light of the rights, welfare, and interests of the original community.[12]

Heyd's view in this passage is a form of presentism, since it draws a moral distinction between presently existing people and possible people who do not presently exist.[13] According to Heyd, only the well-being of presently existing people gives us moral reasons of any kind. While this view captures the second conjunct of the Asymmetry intuition – that we have no reason to bring a possible person into existence just because

---

[12] David Heyd, "Procreation and Value: Can Ethics Deal With Futurity Problems?", *Philosophia* 18 (1988): 151-170, pp. 158-9.

[13] There are other passages in which Heyd seems to endorse not presentism but necessitarianism. The problems with necessitarianism are discussed below.

her life would be worth living for her – it runs counter to the first. According to presentism, there is no welfare-related reason not to create a new person whose life would be miserable, since moral considerations stemming from people's welfare apply only to those who presently exist. For the same reason, we have no moral reason to choose (B) over (A), nor (C) over (A) in the *Non-Identity* and *Same-Person Cases* above. This is not a plausible view.

Consider next a proposal presented by Derek Parfit in *Reasons and Persons* as the best explanation of the Asymmetry. (Although Parfit thinks we can *explain* the Asymmetry in this way, he does not himself endorse this proposal. In fact he rejects it, for the reasons I elaborate below.[14]). According to the proposal,

> [I]t is wrong, if other things are equal, to do what would be either bad for, or worse for, the people who ever live. It is therefore wrong to have the Wretched Child. Since his life is worse than nothing, having this child is bad for him. But it is in no way wrong to fail to have the Happy Child, whose life would be well worth living. True, if the couple had this child, this would be good for him. But if they do not have this child, this would not be bad for him.[15]

Parfit's proposal is a form of *actualism*. The only persons with regard to whom we have well-being-related reasons of any kind are *actual* people, i.e. the people who ever live, *given* our action. Other things being equal, it is wrong to do what is bad for, or worse for,

---

[14] Parfit's proposal echoes a similar argument of Jan Narveson's, who, unlike Parfit, does appear to embrace moral actualism. See Jan Narveson, "Moral Problems of Population", *The Monist* (1973), p. 73.

[15] Parfit (1984), p. 150.

these people. By contrast, it is in no way wrong to fail to create a life that would have been worth living, since there are no moral reasons related to the well-being of people who, given our choice of action, will never live. [16]

While it captures the Asymmetry intuition, the actualist proposal is subject to three serious objections, which give us reason to reject it.

First, it seems extremely plausible to embrace

> *Normative Invariance:* An action's normative status – whether it is right or wrong – does not depend on whether or not it is performed.

Actualism, however, violates Normative Invariance. After all, who the "actual" people *are*, whose interests we must take into account in acting, will often depend on which action we perform. For this reason, actualism will often fail to be action-*guiding* in different-person choices, since there is no morally correct answer to the question "what is the right thing to do?" prior to acting. In order to know how we *ought* to act, we would already know how we will act (which, of course, makes moral deliberation otiose).

Second, actualism has the consequence that actions can be "self-condemning", in the sense that, by performing the action, I make it the case that it was wrong to perform the

---

[16] Actualism thus corresponds to the so-called *narrow person-affecting principle*, according to which something is bad only if it is bad for, or worse for, someone who ever lives.

action.[17] This can give rise to particularly vicious kinds of moral dilemmas, as the

following case illustrates:

> *Actualist's Dilemma*
> Suppose that we must choose between causing either person J or person
> K to exist with a life that is not worth living. If we choose Option 1, J will
> live in agony for 25 years, and K will never exist. If we choose Option 2, K
> will live in agony for 50 years, and J will never exist.

Actualism implies that you cannot but act wrongly in this case (whereas, intuitively, we

might think that, although neither option is *attractive*, it is nevertheless right to select

Option 1, since J will suffer less than K would if we picked Option 2). However, unlike

in an *ordinary* moral dilemma, in which *all* available actions are assumed to be wrong,

actualism implies that, however you act in this case, it will always be true that the

available alternative would have been *right*: Thus, if I choose Option 1 and create J, J's

interests matter morally, but not K's, since K will never exist. Given this, it was wrong to

choose Option 1; I should have chosen Option 2 instead, since this would not have been

bad for J. Unfortunately, if I choose Option 2, it is now true that Option 2 is wrong

---

[17] Similarly, actualism implies that actions can be "self-requiring": If I create a new person with a life worth
living, it will then be true that it would have been wrong not to create him. This is so, because, being actual,
it matters morally that I do what is good for this person. By contrast, if I do not create this person, he is not
actual, and hence I do not act wrongly in failing to create him.

(because it is bad for K and now only K's interests matter); instead, choosing Option 1 would have been morally right.[18]

Third, like presentism, actualism lacks the resources to generate the correct answer in *Non-Identity Case*. Suppose the woman in our example chooses option (A) and creates child X, who will have a moderately happy life, instead of child Y, who would have had a very happy life. Intuitively, this is the wrong thing to do. Actualism, however, is committed to the opposite conclusion. Given the woman's choice of action, child Y never exists, so the fact that this person would have enjoyed a significantly better life than X is morally irrelevant, according to actualism, and provides us with no grounds for criticizing the woman's action. Nor can it be said that creating X is bad, or worse, for anyone who ever lives. By assumption, child X has a life that is well worth living for him. And, unlike in *Same-Person Case*, there is no possible outcome in the Non-Identity Case that would have been *better* for X. All told, it seems that the very feature of actualism that allows it to capture the Asymmetry intuition – namely that it allows us to discount in our moral deliberation the potential well-being of persons who never will exist, given our choice of action – prevents it from giving the right answer in the Non-Identity Case.

---

[18] For a more detailed discussion of a similar case, see Caspar Hare, "Voices From Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?", *Ethics* (2007), pp. 498-523.

Finally, consider *asymmetrical necessitarianism*, a view recently discussed (but again, not endorsed) by Ben Bradley. Bradley writes:

> Let us introduce a bit of jargon to state the view. Let us say that the N-utility of an act = (the total positive welfare for necessary people [i.e. those people who will exist no matter which alternative is performed] produced by the act) (…) minus (the total negative welfare produced by the act). Now we can state Asymmetrical Necessitarianism as follows:
>   *AN*: the welfare-related reason to do an act is proportional to the N-utility of the act.[19]

Like the actualist proposal, asymmetrical necessitarianism manages to capture both conjuncts of the Asymmetry Intuition. Under AN, the negative well-being produced if we create a life not worth living gives us a welfare-related reason against bringing such a person into existence. By contrast, positive welfare matters only if it accrues to a necessary person, i.e. to someone who will exist no matter how we choose to act. This isn't the case here, so there is no reason to create a new person just because her life would be worth living.

Once again, however, the proposal founders on the shoals of the Non-Identity Problem. Since only the well-being of *necessary* people gives us any welfare-related reasons, there is no reason to pick (B) over (A) in *Non-Identity Case*, since neither X nor Y will exist *no matter* how the woman chooses to act. As was the case for presentism and actualism, the very feature that allows asymmetrical necessitarianism to capture the

---

[19] Ben Bradley, "Asymmetries in Benefiting, Harming, and Creating", *The Journal of Ethics* 17 (2013): 37-49.

second conjunct of the Normative Asymmetry intuition prevents it from rendering the correct verdict in *Non-Identity Case*.

Presentism, actualism, and neccessitarianism are all attempts to leave the totalist paradigm, and to find a more plausible basis for population ethics. The reason they fail, I believe, is that they all focus on the wrong *aspect* of totalism, while letting its crucial assumption go unchallenged. As I shall explain in greater detail in the following section, the crucial assumption behind totalism is about the *kind* of welfare-related reasons that we have. According to the totalist, potential well-being matters in exactly one way: it provides us with an *unconditional* (or *categorical*) reason to *bring it about* (be it by benefiting an existing person, or by creating a new person with a life worth living). Presentism, actualism, necessitarianism, and other views of this kind leave this crucial assumption largely unchallenged. They agree that, *if* a person's well-being matters, it can matter only as an unconditional reason to bring about the state of affairs in which this well-being exists. They only depart from totalism by circumscribing the class of persons *whose* well-being matters (only the present people's, only the actual people's, only the necessary people's).

As we have seen, while the proposed modifications to totalism allows these alternative views to capture the second conjunct of the Normative Asymmetry, they simultaneously deprive them of the resources to explain our intuition in *Non-Identity Case*. Try to solve the Non-Identity Problem instead, by affirming totalism, and the solution to the Asymmetry collapses.

18

It is time to try a new approach. In the remainder of this chapter, I argue that the Asymmetry must be explained, not by challenging totalism on *whose* well-being matters, but on *how* well-being matters, i.e. on the *kinds* of welfare-related reasons that we have in procreative contexts.[20]

4. Teleology and its Discontents

Totalist consequentialism is a *teleological* moral theory. What is characteristic of the teleological perspective in modern-day moral philosophy is not just that it takes evaluative notions such as "value" or "good" to be prior to the right. More importantly, what marks a moral view out as distinctly teleological are its claims about the *kinds* of reasons we have with regard to that which is good or valuable. According to the teleologist, the unique appropriate response to what is good or valuable is to *promote* it, ensuring that as much of it exists as possible; the proper response to disvalue is to prevent it, or to ensure that as little of it exists as possible.

Some teleological thinkers, such as G.E. Moore, see such a close connection between goodness and its promotion that Moore often characterizes the good in terms of "what

---

[20] One philosopher who shares my diagnosis that it won't be possible to explain the Asymmetry by dividing up potential persons according to their temporal or modal status into those whose well-being matters and those whose well-being doesn't matter is Melinda Roberts. (See, for instance, Melinda Roberts "The Asymmetry: A Solution", *Theoria* 77 (2011), pp. 333-367). Since Roberts' approach has certain commonalities with the proposal I offer in this paper, it is worth giving a more detailed account of why I regard her proposed solution to the Asymmetry as unsatisfactory. I do so in the Appendix to Chapter 1.

ought to exist". For Moore, ethical questions can be divided into two kinds. The first

concerns the good: what things "ought to exist for their own sakes?"[21] The second is

about the right: "What kind of actions ought we to perform?"[22] One of Moore's central

claims is that latter kind of question can be reduced to the former. What action it is right

to perform in a given situation reduces to the question which available action would

produce the most good: "To assert that a certain line of conduct is, at a given time,

absolutely right or obligatory is obviously to assert that more good or less evil will exist

in the world, if it be adopted than if anything else be done instead".[23] Thus, for Moore

and other teleologists, our moral reasons are *state-regarding* reasons, since they are

reasons to cause what is valuable to exist and what is disvaluable not to exist.[24]

Next, note that viewing some value F as to be promoted implies that there is no deep

moral distinction between increasing the degree to which F is realized amongst *existing*

potential bearers of that value, and creating *new* bearers of that value. These are both just

---

[21] G. E. Moore, *Principia Ethica,* ed., with an introduction by Thomas Baldwin, rev. ed., with the preface to the (projected) 2d ed. and other papers (Cambridge: Cambridge University Press, 1993), p. 33.

[22] Ibid., p. 33.

[23] Ibid., p. 77.

[24] Even paradigmatically non-consequentialist ethical theories like W.D. Ross's can have teleological *elements*. Thus, in *The Right and the Good* Ross argues that goodness is "an intrinsic quality of certain things," such as pleasure, knowledge, and virtue. "What we ought to do," he says, "depends to a large extent (…) on the goodness or the badness of the things we can in our acts bring into being."

ways of making it the case that more of what is valuable or good exists in the world.

That is, someone who views a value *F* as to be promoted affirms the following

> *Transfer Thesis*: If there is reason to increase the extent to which *F* is instantiated amongst existing potential bearers, there is also reason to increase the extent to which *F* is instantiated by *creating* new bearers of *F.*

This thought is at the root of the totalist utilitarian's rejection of the Procreation Asymmetry: if well-being is good or valuable, as witnessed by the fact that we want the lives of existing people to contain as much well-being as possible, then surely the fact that the lives of potential new people would also contain well-being must constitute a reason for creating these people.[25]

If we accept the thought that the unique appropriate response to what is good or valuable is to promote it, this also has implications for the kinds of things that we can think of as ultimately valuable. For only certain kinds of things *can* be promoted: Specifically, note that promoting is not really a response that it is possible to have towards *particular* concrete entities, such as particular persons or animals. What could it

---

[25] Teleologists about well-being need not endorse a further controversial claim sometimes attributed to Moore, namely that what is *good for* people is so only if, *and because*, it is good *impersonally*, "from the point of view of the universe". (For a contemporary defense of this view, see Donald Regan, "Why Am I my Brother's Keeper?" in Wallace, Pettit, Scheffler and Smith, eds., *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 202-230). For some central components of well-being, such as sensory pleasure, it is far more plausible to think of the dependence relationship between "good impersonally" and "good for people" as being the other way around. See T. M. Scanlon, "Ideas of the Good in Moral and Political Philosophy" (ms.).

mean to "promote" Tim Scanlon, or Baloo the bear? Rather, what can be promoted are *abstracta*, such as properties (well-being; wisdom) or universals (bears), which we can cause to be realized or instantiated to a greater or lesser extent in a state of affairs.

Much of what makes utilitarianism unattractive to many people has its root in this focus on abstracta over particular beings and entities: For one thing, a focus on promoting as the unique response to what is good or valuable sidelines a whole range of valuing attitudes that we have specifically (indeed exclusively) towards particulars: cherishing, respecting, loving, caring for, honoring, etc. For another, it feeds a common criticism of utilitarianism, namely that it treats people as *fungible* and views them in a *quasi-instrumental* fashion. Instrumental valuing is an attitude that we have towards particulars. However, to value something instrumentally is to value it, in essence, for its causal properties. But these same causal properties could just as well be instantiated by some other particular thing. Hence, insofar as a particular entity is valued only instrumentally, it is regarded as fungible. Similarly, a teleological view which regards our welfare-related reasons as purely state-regarding can be accused of taking a quasi-instrumental approach towards people. It views them as fungible receptacles for well-being, not as mattering qua individuals.[26] Totalist utilitarianism, it is often said, does not

---

[26] Note that a modified utilitarian view, according to which we have reason to promote not well-being *per se*, but *happy lives*, would face similar objections. Although such a view could not be accused of regarding people as mere *receptacles* or *loci* of what ultimately matters (namely well-being), it would still regard

take persons sufficiently seriously.[27] By treating the moral significance of persons and their well-being as *derivative* of their contribution to valuable states of affairs, it *reverses* what strikes most of us as the correct order of explanation. Human wellbeing matters because people matter – *not vice versa*.

I have focused on totalist utilitarianism in this section, not because I believe that it is the only, or indeed the most plausible, basis on which philosophers tend to reject the Asymmetry and endorse a symmetry view. Rather, reflecting on the shortcomings of totalist utilitarianism is the easiest way to recognize the attractions of an alternate way of conceiving the reason-giving force of well-being. According to my proposal, which I sketch in the next section, whatever moral reasons we have to confer well-being on people are not state-regarding but what I call *bearer-dependent* reasons. As we shall see, if *this* is the correct way to think of the reason-giving force of well-being, then the Asymmetry becomes not just relatively straightforward to explain, but indeed hard to resist.[28]

---

individual happy people as fungible constitutive means towards achieving our ultimate goal, namely to bring about a state of affairs in which as many happy lives as possible are lived.

[27] For a penetrating discussion of this point, to which I am indebted, see Ralf Bader, "Aggregating v. Balancing" (ms).

[28] One alternative to totalist utilitarianism, which I return to in the following section, is a version of the *wide person-affecting principle*, first discussed by Parfit in Chapter 18 of *Reasons and Persons*. According to the *Wide Total Principle*, we have moral reason to bring about that outcome which gives to people the greatest total net sum of benefits, where these include the *existential* benefits to people of being brought into existence. Like totalist utilitarianism, the Wide Total Principle rejects the Asymmetry, but does so for a different reason: all else equal, it is claimed, we have reason to create new happy lives because doing so will benefit

5. Bearer-Dependent Reasons and the Second Conjunct of the Asymmetry

My strategy for defending the Procreative Asymmetry begins by locating it in a wider normative phenomenon. It is striking that a teleological approach, according to which our moral reasons relating to some value F must be state-regarding reasons to *promote* F, seems even more problematic for moral values *other than* well-being.

Consider, for instance, the value of justice: The thought that it is good to achieve justice is not a free-floating claim about valuable states of affairs. Rather, we believe, the demands of justice have their source in *other persons*, as beings that are capable of having and responding to reasons, and of choosing and revising their ends. As such, they have the standing to demand of us certain appropriate attitudes and behaviors, amongst which is a reciprocal willingness to structure our shared institutions and social interactions in a manner that is justifiable to all. Given that we are surrounded by such beings, social interaction with which gives rise to demands of justice, we also have reason to think that the world is better, all things equal, if we succeed in treating one another justly. But this thought is a *derivative* one, which follows from the normative reasons we have to structure our institutions and social interactions in a way that is

---

the *particular* people we create, not because we have state-regarding reason to bring about a state of affairs that is better on account of containing more well-being. While avoiding some of the problems with totalist utilitarianism, I will argue that the Wide Total Principle nonetheless rests on a mistaken conception of our reasons for benefiting other people.

justifiable to all. It does not flow from the belief that justice is a value that ought to be maximally instantiated. Indeed, it would plainly be absurd to think of justice as a value to be promoted in the sense of the Transfer Thesis (such that we could have moral reason to create new persons *just in order* that they may treat one another justly). If that were the case, the claims of justice would be limitless. It would be impossible *in principle* to achieve a situation that is 'perfect' from the point of view of justice; for we could always promote this value further by creating new people whose relationships with other persons also instantiate the value of justice.

Similar remarks apply, *mutatis mutandis*, to many other values: liberty, equality, fairness, honesty, fidelity, loyalty, promise keeping, gratitude, charity, health, safety, etc. None of these values appear even remotely plausible as candidates for "promotion" in the sense of the Transfer Thesis. For instance, while we recognize strong moral reasons to make people free and equal, freedom and equality clearly do not require us to create *new* people so that they, too, may instantiate these values.

Indeed, it is striking that in thinking about these other values, we soon notice asymmetries that are structurally analogous to the claims of the Procreation Asymmetry. Consider the case of promise-keeping: Most of us believe that we have a moral reason not to make a promise that we won't be able to keep. (Compare: we have a moral reason not to create a life that will unavoidably be not worth living). By contrast, we do not think that we have a reason *to make* a promise *just because* we will be able to keep it (Compare: we do not think we have a moral reason to create a new life, *just because* that

25

life will be worth living). Keeping a promise does not seem to add any moral value to the world that must be taken into account when deciding whether to make that promise.

My contention is that there is indeed a common moral phenomenon which explains why all moral values, including the value of human well-being, are prone to exhibit intuitive asymmetries similar to the one that we observe in the case of procreation. But, for simplicity of exposition, the rest of this section will concentrate on exploring only the parallel between the procreation and promising asymmetries.

In the case of promising, it is not hard to see why I have no reason to make a promise, just because I can keep it: Any reasons to keep our promises are *conditional* reasons, namely conditional on the promise having been made. Making a promise, on the account I favor, involves the promisor giving the promisee a claim-right to a certain future action on the part of the promisor. In language that will be helpful when comparing promising to procreation, we could say that the act of promising involves "creating" a promisee, i.e. creating a *bearer* of a promissory claim-right, and that any reasons to keep a promise are conditional on the existence of such a promisee. If this is correct, it is plain to see why there could not be an unconditional *promissory* reason to (make and keep a promise).[29] For any reason to keep a promise is *conditional* on the

---

[29] Though there could be other sorts of reasons for doing so, for example, to prove one's trustworthiness.

existence of the bearer of a promissory claim-right. Our reasons to keep promises are, to coin a new term, *bearer-dependent* reasons.

My central claim in this section is going to be that any moral reasons we have to confer benefits on a person P are likewise bearer-dependent, in the sense that they are *conditional* on P's existence. If this is correct, it will explain the second conjunct of the Normative Asymmetry. There is no unconditional moral reason to create a person, just on account of the benefits we would thereby confer on her, since any welfare-related reasons to confer benefits on a person are conditional on her existence.

Here is how I propose to establish this claim. According to the totalist utilitarian view that I criticized in the previous section, people and their well-being matter *in virtue of* contributing to good *states of affairs*. This view, I contended, *reverses* the true order of explanation, taking the significance of persons and their well-being to be *derivative* of their contribution to good states of affairs.

The truth is almost exactly the opposite. I affirm:

> *Claim 1*: It matters that people have well-being just if, *and because,* people matter.

Claim 1 employs the term "mattering" in two distinct senses that we should distinguish: First, there is what we may call the *state-regarding sense* of 'mattering'. It is the sense involved in statements of the form "It matters that p", where p is a proposition that describes some state of affairs or way the world can go. To affirm that "it matters that p", in this state-regarding sense, is roughly to assert that it is not morally indifferent

27

whether p is true or not. There are reasons to make it the case that p and/or to have a pro-attitude towards p's being the case.

Second, there is what we can call the *bearer-regarding sense* of 'mattering'. This is the sense involved in statements of the form "S matters", where S is not some state of affairs, but a particular being or thing. In the case where S is a person, to affirm that "S matters" is, very roughly, to assert that that it is not morally indifferent how S fares and is treated. We have (moral) reasons to regulate our conduct and attitudes in certain ways out of consideration for S's interests, rights, and claims. The generic plural statement "S's matter" (e.g. "people matter") can be read as affirming: "It is true for any S that S matters in the bearer-regarding sense". [30]

Claim 1 can thus be read as affirming that any moral reasons that we have to make it the case that people have well-being (or to have a pro-attitude towards the obtaining of this state of affairs) are true in virtue of, and derivative of, the fact that individual people matter in the bearer-regarding sense.[31]

---

[30] We can distinguish further senses of mattering, besides the state-regarding and the bearer-regarding. For instance, there is what we can call the *prudential sense* of 'matters' (as in "health matters"), where to say that "F matters" means, roughly, that there are prudential reasons for some bearer S to be or have F. There is also the *ideal-regarding sense* of 'matters' (e.g. "loyalty matters", "justice matters"), where to say that "I matters" is to say, roughly, that we have reason to regulate our conduct and attitudes in line with I.

[31] A number of philosophers affirm views that are similar to proposition (1). See in particular Stephen Darwall, *Welfare and Rational Care* (Princeton, NJ: Princeton University Press, 2002), Chapter 1 and Elizabeth Anderson, *Value in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1993). For Anderson, in particular, something like (1) falls out of a broader view, according to which states of affairs *in general* are valuable only "extrinsically". By this she means that the fact that certain states of affairs matter, in the state-regarding sense, is *derivative* of the fact that some things matter in the bearer-regarding sense. More

Next, let us ask: Why do individual people matter in the bearer-regarding sense? To give a detailed response to this question is beyond the scope of this paper, but for my present purposes a schematic answer is sufficient. I believe that most moral philosophers would subscribe to something like the following:

> *Claim 2*: A person S matters, in the bearer-regarding sense, just if, and because, S has moral status, which in turn is grounded in various properties of S.

There are many competing views about what properties of a person ground her moral status, but my argument does not depend on any particular view about the grounds of moral status. I will just mention my own view, which is that a person S's moral status is grounded in the fact that S has a good and is capable of having reasons and choosing ends. But, to repeat, nothing rides on this being the correct account of the grounds of moral status for persons.

---

precisely, those states of affairs that matter, in the state-regarding sense, are precisely those states of affairs that we have reason to bring about *because* some being or thing S matters in the bearer-regarding sense. As she writes:

> (…) states of affairs, whether they be final aims or mere means, are for the most part only *extrinsically* valuable. It makes sense for a person to value most of them only *because* it makes sense for a person to care about the people, animals, communities, and things concerned with them. (*Value in Ethics and Economics*, p. 20)

While I am in sympathy with Anderson's thesis that states of affairs in general are only extrinsically valuable, I need not affirm anything as broad as Anderson's thesis in mounting my defense of the Procreation Asymmetry.

I will now show that, given Claims 1 and 2, there is an argument for the second conjunct of the Normative Asymmetry that has not, to my knowledge, been made before.

Let w be a world such that the following is true:

(A) It is not the case that S is happy in w.

When, and why, is it a matter of moral concern that (A) is true? It seems that

Only if S's well-being matters, in the state-regarding sense, is it a matter of moral concern that (A) is true.

However, according to Claim 1,

S's well-being matters, in the state-regarding sense, just if, and because, S matters, in the bearer-regarding sense.

And, according to Claim 2,

S matters, in the bearer-regarding sense, just if, and because, S has moral status, which in turn is grounded in various properties of S [such as the fact that S has a good and is capable of having reasons and choosing ends].

However,

Only if S exists in w does S possess *any* properties, including those properties that ground her moral status.

Therefore,

The truth of (A) is of moral concern *only if S exists in w*.

By contrast, if S does not exist in w, it will also be true that "it is not the case that S is happy in w." But in that case, the fact that (A) is true is *not* a matter of moral concern. For in that case, there exists no person whose moral status gives us reasons to care about

his potential happiness (or lack thereof). Therefore, if by failing to create a person S who would have been happy in w, we make it the case that (A) is true but S does not exist in w, this is not a matter of moral concern. To paraphrase Jonathan Bennett: while we have reason to deplore the situation where a person lacks happiness, there is no reason to deplore a situation where happiness lacks a person.[32]

But if it is not a matter of moral concern that we fail to create a person who could have had a happy life, this means that there is no moral reason to create a person, just because that person could have a happy life. For if there were such a reason, the failure to comply with this reason *would* be a matter of moral concern.

From the fact that creating S would constitute an existential benefit to S, we therefore cannot infer that there is a moral reason of beneficence to create S[33]. Reasons to benefit a person S, I have argued, obtain only *conditional* on S's existence. They do not give us *unconditional* reasons to bring S into existence.

If this argument goes through, I have shown how jettisoning the totalist utilitarian's view of well-being as something to be promoted in favor of a more attractive view according to which a person's well-being matters, just if, and because, the person matters, lends support to the second conjunct of the Normative Asymmetry. Moreover it

[32] Jonathan Bennett, "On Maximizing Happiness" in R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations* (Temple University Press, 1978), pp. 61–73.

[33] As would a proponent of the Wide Total Principle - see footnote 28 above.

does so in a way that offers a principled response to the Objection from Benefit: if our reasons to benefit other people are *conditional* on the fact of their existence, this explains why we often have weighty moral reasons to provide *ordinary* benefits (which accrue to people who either already exist or will exist independently of our action), but no moral reasons to provide *existential* benefits to people by bringing them into existence.

6. Explaining the First Conjunct: Moral Standards

I have yet to provide much detail about the content of our welfare-related reasons with regard to procreation. Most importantly, I need to answer the following question: Against the teleologist, I have argued that any reasons to *confer* well-being on a person are *conditional* on the fact of her existence. Thus, they are not also reasons to bring her into existence just in case her life would contain sufficient well-being. Given this, how can we nonetheless have *unconditional* reasons *against* bringing a person into existence whose life would not be worth living – as the first conjunct of the Normative Asymmetry affirms?

In this section, I introduce a new normative notion, that of *normative standard*. Characterizing our welfare-related reasons with regard to procreation in terms of normative standards allows us to explain both conjuncts of the Asymmetry Intuition. The concept of normative standard also enables us to capture structurally analogous asymmetries, such as the promissory asymmetry that we considered above. In the

following section, I then argue that our "standard-regarding" reasons are best expressed in the form of *wide-scope* conditional reasons.

A normative standard, as I shall use the term, is a normative criterion that applies to those outcomes of an agent's actions that fall within the scope of the standard. An outcome can either *fail* to satisfy the standard, in which case there are "standard-regarding" reasons to avoid this outcome, or it can *pass* the standard, in which case there are no standard-regarding reasons against bringing it about. That in virtue of which a standard has normative force, I call its *ground*. The scope of the standard is coextensive with those outcomes in which its ground ever exists. If the ground never exists in some outcome, the standard does not apply to that outcome, and consequently we can have no standard-regarding reasons against bringing it about.

Let us familiarize ourselves with this notion by considering some examples. Consider, first, the case of promising.

Suppose I am deliberating whether to promise you to walk your dog tomorrow. (See *Figure 1.1*). I know that, if I make you this promise, my subsequent actions will be subject to a new moral standard. I can satisfy this standard (by walking your dog tomorrow) or fail it (by breaking my promise and not walking your dog). Thus, conditional on making the promise today, I have a standard-regarding reason to walk your dog tomorrow. What grounds the moral force of this standard is that in both outcomes that I can bring about, conditional on having promised to walk your dog, you are the bearer of a promissory claim-right concerning my actions tomorrow. By contrast,

if I do not make you this promise, I remain free to do as I please, at least as far as the promissory standard is concerned. For in this case, there is nothing that grounds a promissory reason to behave in any particular way tomorrow. My actions fall outside the scope of the promissory standard.[34]
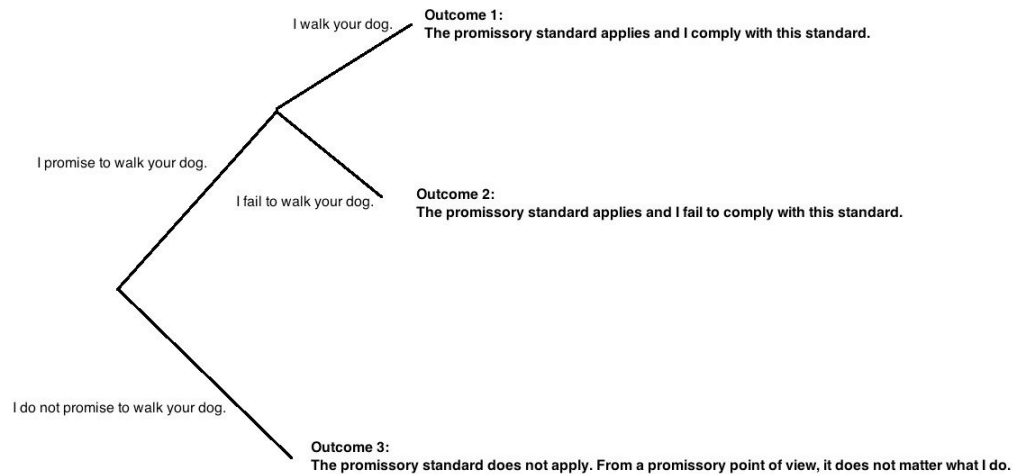
Figure 1.1

Suppose, next, that I can foresee that, having made you the promise, I will be unable to keep it. That is, in *Figure 1.1*, if I make the promise, I will unavoidably bring about

---

[34] The case of promising can also be used to explain why I wrote above that "the scope of [a normative] standard is coextensive with those outcomes in which its ground *ever* exists." This "timeless" formulation is needed in order to capture the fact that we can have moral reasons to keep promises that we made to people who no longer exist.

Outcome 2, in which I fail the promissory standard. Knowing this, I have a standard-regarding reason to *avoid making* the promise in the first place. More generally, there is a standard-regarding reason to avoid bringing about an outcome to which a normative standard applies and in which I am unable to comply with that standard.

By contrast, there can be no standard-regarding reason against producing an outcome which is *not* governed by that standard. Thus, there can be no *promissory* reason against not making a promise, even a promise that I could keep. Hence, given a choice between bringing about an outcome to which a standard applies and that standard is satisfied, and bringing about an outcome that is outside the scope of the standard, I have no standard-regarding reasons to do the former rather than the latter.

The notion of a normative standard is also helpful in thinking about cases where our reasons are *prudential*, not moral, in nature. Suppose I am thinking of climbing Mount Everest. I know that, at that altitude, I will need a functioning oxygen mask – a need that I won't have if I don't climb the mountain. In *Figure 1.2*, Outcomes 1 and 2, in which I climb the mountain, are thus subject to a prudential standard – the standard of my need for a functioning oxygen mask – which does not apply to the outcome in which I remain in the flatlands.

I have an oxygen mask.

Outcome 1:
The standard of my need for an oxygen mask applies, and I satisfy this prudential standard.

I don't have an oxygen mask.

I climb Mount Everest.

Outcome 2:
The standard of my need for an oxygen mask applies, and I fail this prudential standard.

I remain in the flatlands.

Outcome 3:
The standard of my need for an oxygen mask does not apply.
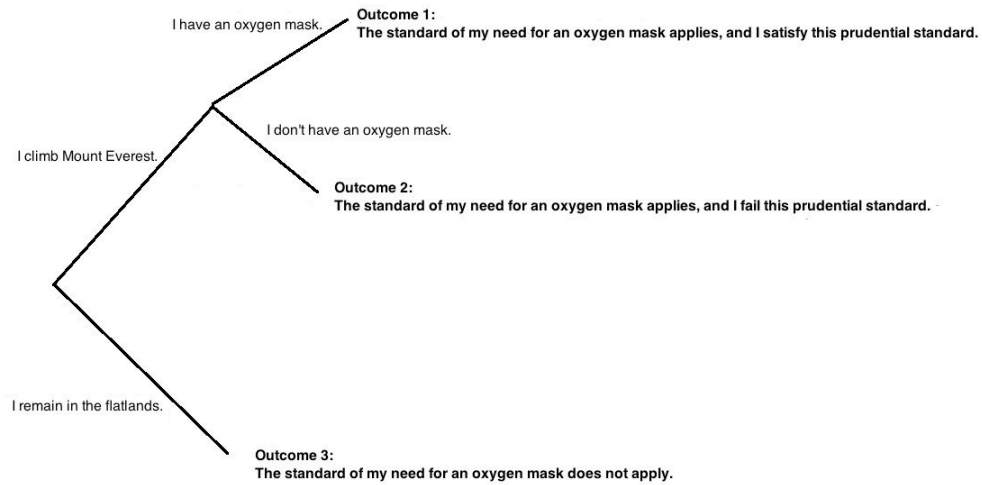
Figure 1.2

Here again, the fact that I foreseeably will be unable to satisfy a prudential standard gives me a reason to avoid outcomes to which that standard applies: The fact that I don't have a functioning oxygen mask gives me a reason not to attempt the climb. By contrast, the fact that I could satisfy a prudential standard does not, in itself, give me a reason to bring about an outcome to which this standard applies and is satisfied, rather than an outcome that falls outside the scope of this prudential standard: The mere fact that I have a functioning oxygen mask does not give me a reason to climb Mount Everest rather than to remain in the flatlands.

Consider, finally, how the notion of a normative standard can help us to explain the first half of the Procreation Asymmetry (*Figure 1.3*). I claim that any outcome in which I create a new person S is subject to a moral standard (that of S's well-being), which is grounded in the existence of S, a being with moral status. This moral standard is

36

satisfied if S has a life that is at least worth living, and failed if she has a life that is not

worth living. By contrast, the standard of S's well-being does not apply to an outcome in

which I do not create S.

S has a life that is worth living.   **Outcome 1:**
                                      The standard of S's wellbeing applies, and I satisfy this standard.

I create person S.   **Outcome 2:**
                                      The standard of S's wellbeing applies, and I fail this standard.
S has a life that is not worth living.

I do not create person S.

                                      Outcome 3:
                                      The standard of S's wellbeing does not apply.
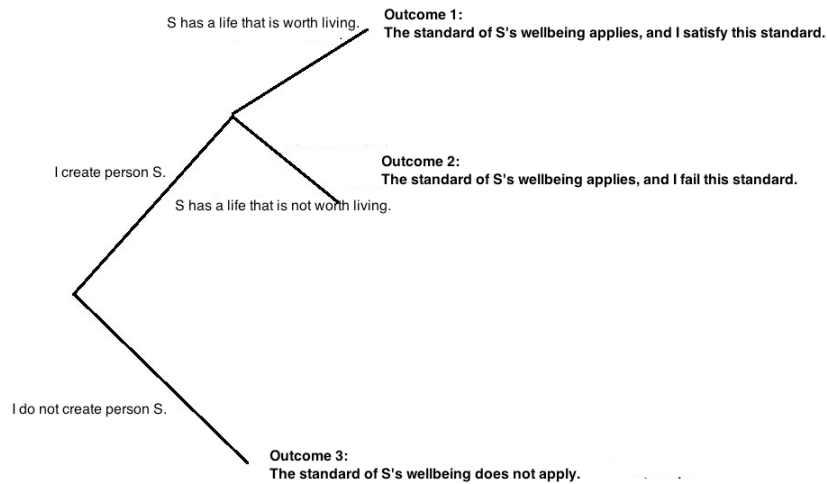
Figure 1.3

If I am unable to give S a life that is worth living, I have a standard-regarding reason not

to create S. This explains the first half of the Asymmetry. By contrast, the mere fact that I

could create a person S whose life would be worth living does not give me a moral

reason to do so. For I have no standard-regarding reason to bring about an outcome to

which the standard of S's well-being applies and that standard is satisfied, rather than

an outcome that falls outside of the scope of this normative standard. This explains the

second half of the Asymmetry.

37

Before I go on, let me pause to briefly comment on the dialectic of this and the previous section: Attentive readers will have noticed that the notion of normative standard is itself asymmetrical. There are standard-regarding reasons not to bring about an outcome to which a normative standard applies and that standard is failed, whereas there are *no* standard-regarding reasons to bring about an outcome to which a normative standard applies and that standard is complied with. For this reason, it would have been question-begging to appeal directly to the notion of normative standard in seeking to explain the Procreation Asymmetry, without first giving an *independent* argument to the effect that we have no moral reasons to create a person just because her life will be worth living. That is why the argument of the preceding section, in defense of the second conjunct of the Procreation Asymmetry, was indispensable. With that argument in place, I could *then* appeal to the notion of normative standard to provide a *unified* account of both conjuncts of the Procreation Asymmetry, without begging any questions.

7. Standard-regarding reasons as wide-scope conditional reasons

While the notion of a "normative standard" is novel, I believe that our standard-regarding reasons can be captured in terms of a more familiar concept – that of conditional reason.

Conditional reasons come in two basic forms, depending on whether the reason-operator takes wide or narrow scope:

A narrow-scope conditional reason: If I do p, I have reason to do q.

A wide-scope conditional reason: I have reason to (if I do p, do q).

I believe that the correct way to capture standard-regarding reasons is in terms of wide-scope conditional reasons. With regard to the case of procreation, I affirm the following wide-scope conditional reason:

> *The Threshold Requirement:* I have a moral reason to (if I create a new person, make it the case that this person's life is at least worth living).

Suppose that the Threshold Requirement had instead been presented as a narrow-scope requirement:

> *The Threshold Requirement (narrow-scope reading):* If I create a new person, I have a reason to make it the case that this person's life is at least worth living.

There are two connected objections to this narrow-scope reading of the Threshold Requirement. The first seems to me open to debate, but the second does strike me as dispositive:

(1) *The "Reason" Implies "Can" Problem:* There are some lives that, on account of a severe congenital malady such as Tay Sachs disease, are *irredeemably* miserable: there is nothing that anyone can do to effectively mitigate the suffering of the new person, nor can the new person be sufficiently compensated to make her life worth living overall. If I create a child whose life is irredeemably miserable, I will therefore be *unable* to make it the case that the child has a life that is at least worth living. But, according to the

narrow-scope version of the Threshold Requirement, this is precisely what I have reason to do, conditional on creating the child. The narrow-scope reading thus implies that I sometimes have moral reason to do what I am unable to do. This runs counter to the thought, accepted by some moral philosophers, that deontic operators like "ought" or "reason" imply "can".[35]

The wide-scope reading avoids the "reason" implies "can" problem. On the wide-scope reading, I have reason to avoid an outcome in which the following *combination* of propositions are both true, namely that (i) I create a child; and (ii) I do not make it the case that this child can have a life that is at least worth living. This is a reason that it is possible to comply with even in the case of an irredeemably miserable life. If it is foreseeable that if I make (i) true, (ii) will unavoidably be true as well, I can avoid making both (i) and (ii) true by not making (i) true in the first place.

(2) *The Problem of Unconditional Reasons:* The decisive problem for the narrow-scope reading is that it lacks the resources to explain how I could have an unconditional reason *not* to create a child whose life will be irredeemably miserable. According to the narrow-scope reading, *any* reasons that I have with regard to my future child's well-being – and not just any reasons to *confer* well-being on the child – are reasons that I acquire *only upon it being the case that I create that child*. But if any welfare-related reasons

---

[35] I myself am agnostic about this "reason"-implies-"can" requirement.

that I have with regard to a potential future person take effect only once it is the case that I create this person, I will necessarily lack any welfare-related reason to *avoid* creating an irredeemably miserable life in the first place. The narrow-scope reading thus lacks the resources to explain the first conjunct of the Asymmetry.

Does the wide-scope reading fare any better? At first blush, it, too, faces a serious obstacle: It is uncontroversial that narrow-scope conditional reasons allow for what deontic logicians call *factual detachment*: The narrow-scope conditional reason "If I do p, I have a reason to do q", together with the factual premise "I do p", allow me to deduce the *unconditional* statement "I have a reason to do q".

By contrast, when the reason-operator takes wide-scope, it is commonly thought that the detachment of unconditional reasons is illicit. In this respect, the "reason" operator is thought to behave like the alethic operator "necessarily". The following argument, involving factual detachment of the necessarily-operator, is clearly invalid:

> *Necessarily* (If there is water in the glass, something is in the glass).
>
> There is water in the glass.
>
> ⇒ *Necessarily* (Something is in the glass).

Indeed, it is this supposed resistance to factual detachment which recommends wide-scope conditional reasons to some philosophers in other contexts. According to a popular account of instrumental reason, for instance, the instrumental principle takes the form of a wide-scope conditional reason:

I have a reason to (if I have end E and doing M is the known necessary means to E, do M).[36]

Philosophers who embrace such wide-scope accounts of instrumental reason[37] do so precisely in order to avoid problems that would be created by factual detachment. For example, an account of instrumental reason would be unsatisfactory if it implied that I have an unconditional reason to take the necessary means to my evil ends, or that I have a reason to perform evil actions, provided they are the necessary means to my ends.

Nonetheless, I believe that two more limited detachment rules can be defended. They are inherently plausible and solve the problem at hand, but without endangering the viability of analyses of instrumental reason in terms of wide-scope conditional reasons.

Following Patricia Greenspan[38], I propose a new operator and two new detachment rules for wide-scope conditional reasons. According to _Rule 1_, the unconditional reason "I have a reason to do q" will be derivable from the wide-scope conditional reason "I have reason to (if I do p, do q)" and "*U*(I do p)". *U* is a new operator, which asserts that the truth of the proposition within its scope is _unalterable_ by the agent at the time of

[36] See, for instance, Simon Rippon, "In Defense of the Wide-Scope Instrumental Principle", _Journal of Ethics and Social Philosophy_ 5.2 (2011), pp. 1-21.

[37] For the record: I am not one of them.

[38] Patricia Greenspan, "Conditional Oughts and Hypothetical Imperatives," _Journal of Philosophy_ 72 (1975): 259-76.

action. This will typically be true in cases where I have already made the proposition true, and can no longer make it false. In such cases, where I am no longer able to satisfy the wide-scope conditional reason by making its antecedent false, Rule 1 implies that I have an unconditional reason to make the consequent true.

Rule 2 is the contrapositive of this. Given the wide-scope conditional reason "I have reason to (if I do p, do q)" and "U~(I do q)", I can detach an unconditional reason _not_ to do p.

These two new detachment rules for wide-scope conditional reason both strike me as highly plausible. What is more, they allow me to provide an explanation of both conjuncts of the Procreation Asymmetry within a unified framework, where any reasons to _confer_ well-being on a person are _conditional_ on the fact of her existence. Consider again the Threshold Requirement I presented above:

> _The Threshold Requirement:_ I have a moral reason to (if I create a new person, make it the case that this person's life is at least worth living).

According to Rule 2, if it is unavoidably true that, should I create a certain child, I will fail to make it the case that this child can have a life worth living (because the child's life would be irredeemably miserable), I have an _unconditional_ reason not to create this child. Thus, together with Rule 2, the wide-scope reading of the Threshold Requirement allows me to affirm the first conjunct of the Procreation Asymmetry. At the same time, there is _no_ unconditional reason under the Threshold Requirement to create a child, just because I will be able to give it a life worth living. This is in line with the second

conjunct of the Procreative Asymmetry, and allows me to meet the Objection from Symmetry. Rather, any reason to confer well-being on the child is *conditional* on that child's existence. Only once the proposition "I create a child" is unavoidably true — namely once I *have* created the child — does Rule 1 allow me to detach an unconditional moral reason to make it the case that the child has a life that is at least worth living.

My proposal also avoids the bootstrapping problem for wide-scope accounts of instrumental reason which I alluded to above. The problem, to recapitulate, is that I should not be able to infer an unconditional reason to take the necessary means to my ends from a wide-scope reason of the form

> I have a reason to (if I have end E and doing M is the known, necessary means to E, do M).

plus the mere fact that I have these ends; otherwise, I could in principle have a reason to do anything, however immoral, as long as this was a known, necessary means to some end of mine.

My proposal avoids this problem, since Rule 1 applies only to cases where it is *unavoidable* at the time of action that the antecedent of the wide-scope conditional reason is made true. But this will never be the case for situations governed by the instrumental principle: at any time when I can still take (or fail to take) the necessary means to my end (i.e. at any time when the instrumental principle gets a grip on me), I can also give up my end instead. Hence, it is never the case that, while I can still decide whether or not to do M, it is already unavoidable that the antecedent is true (i.e. that I have end E, to

which M is the necessary means). Hence, Rule 1 will never be applicable to the instrumental principle.

8. The Same-Person Case and the Maximization Requirement

Although the Threshold Requirement, by itself, suffices to generate both conjuncts of the Procreation Asymmetry, it does not capture all our well-being-related reasons in procreation. Note, for instance, that the Threshold Requirement cannot explain our confident intuition that the pregnant woman in Same-Person Case has a weighty moral reason to undergo the procedure that, at negligible cost to herself, will allow her child to have a very happy instead of a moderately happy life.

In order to see how my account can capture this intuition, we must refine our understanding of normative standards. Up to this point, I have discussed normative standards as if passing and failing a normative standard were *binaries*. But, as a matter of fact, most normative standards are likely to be non-binary, in the sense that there are different *degrees* of success and failure in satisfying the standard.

Consider again the case of promising. I can fully keep a promise, partially keep it to various degrees, or not keep it at all. On what might be termed a "rigorist" view about promising, only *fully* keeping a promise satisfies the promissory standard. It is prima facie wrong to make any promise that I foreseeably am unable to *fully* keep. Even the rigorist, however, will allow that *failure* to comply with the promissory standard may come in degrees: it will often be morally more serious to make a promise and not keep it

at all than to make a promise and partially keep it. A "permissive" view about promising goes one step further, holding that the promissory standard can be both failed and *satisfied* to various degrees. According to the permissive view, there may be a degree of promise-keeping D <100% at which the promissory standard counts as *satisfied*. In that case, there is no promissory reason, all else equal, against making a promise that I foreseeably am able to keep only to degree D. Consider a promise to pick you up at the station at noon. Clearly, there is some threshold of lateness such that, if I am late beyond that threshold, I do not count as having kept my promise to a degree that satisfies the promissory standard. If I know in advance that I unavoidably will be late beyond this threshold, I have a moral reason not to make the promise in the first place. (Of course, this reason could be outweighed by other considerations unrelated to the promissory standard). But likewise, we may hold that some delays are small enough that they do not prevent me from satisfying the promissory standard in this context – although I satisfy this standard to an even higher degree if I am *right* on time.

By reconsidering my mountaineering example from Section 6, we can further bolster the idea that the successful satisfaction of a standard, normative or prudential, often admits of degrees. Suppose I am contemplating climbing Mount Everest and have a choice between two oxygen masks: the kind of rudimentary oxygen mask available to Sir Edmund Hillary in 1953 or a state-of-the-art model. Suppose that both masks would *suffice* to minimally satisfy my need for oxygen. That is, in a scenario where *only* the rudimentary oxygen mask was available, I would not have a prudential reason against

climbing Mount Everest. Nonetheless, given a choice between both oxygen masks, I have a prudential reason to choose the state-of-the-art model, since it will *better* satisfy my need for oxygen. (Of course, the basic insight of Section 6 remains unaffected by this: even the availability of a state-of-the-art oxygen mask does not give me, in itself, a reason to climb Mount Everest).

It seems highly plausible that the *well-being* standard that applies to procreation can also be satisfied (and failed) to various degrees. While creating person X with a moderately happy life does satisfy the well-being standard (because X has a life that is worth living), making it the case that X enjoys a *very* happy life satisfies the well-being standard to a *higher* degree. Moreover, unlike in the case of promising, where there is such a thing as *fully* satisfying the promissory standard by keeping my promise 100%, the well-being standard is both upwardly and downwardly *unbounded.* For *any* life, however happy, we can imagine a yet happier life that would satisfy the well-being standard to a higher degree (and *mutatis mutandis* for miserable lives).

For cases in which the same normative standard can be satisfied to different degrees, I affirm the following principle:

> *Principle of Greater Satisfaction*: If I am to bring about one of two outcomes in which the same normative standard is satisfied (or failed) to different degrees, I have standard-regarding reason to choose the outcome in which the standard is satisfied to a higher rather than to a lower degree. ("Sameness" of normative standard is here understood to be a function, not just of the kind of normative standard in question, but of the identity of the individual who grounds the standard).

Thus, in a choice between making you a promise and keeping it to degree D < 100% and making you a promise and *fully* keeping it, all else equal, I have promissory reason to bring about the latter rather than the former outcome, even if neither outcome would fail the promissory standard. Note, however, that the reason given to us by the Principle of Greater Satisfaction is *merely contrastive*[39]: Holding fixed the fact that I will make you the promise, I have a reason to make the promise and fully keep it *rather than* to make the promise and only keep it to degree D. However, this does not imply that I also have an *unconditional* reason to (make the promise and fully keep the promise). Given a choice between (making the promise and fully keeping the promise), and *not making* the promise, the Principle of Greater Satisfaction is silent.

Applied to the case of procreation, the Principle of Greater Satisfaction implies the

> *Maximization Requirement*: If I have a choice between creating a new person X with a life worth living at well-being level W, and creating the same person with a life worth living at well-being level V, where W > V, then I have contrastive reason to bring about the former rather than the latter outcome, if I am to create X at all.

This explains our judgment in *Same-Person Case*. Note that the Maximization Requirement is compatible with the prioritarian intuition that the strength of the

---

[39] A contrastive reason, in general, is a reason to do p *rather than* q. According to Justin Snedgar ("Reason Claims and Contrastivism about Reasons", *Philosophical Studies* 166:2 (2013), pp. 231-242) *all* reasons for action are to be understood as contrastive reasons, i.e. as reasons to do one thing *rather than* another. According to Snedgar, "reason" expresses a relation with an argument place for a set of alternatives. Whether or not Snedgar is right about this, the converse is surely not the case: a contrastive reason to do p *rather than* q need not also be a reason to do p *simpliciter*.

contrastive reason diminishes, the higher the levels of well-being W and V in question. Thus, my contrastive reason to make it the case that my offspring's life is *extremely* happy rather than *very* happy may be weaker than my contrastive reason to make it the case that her life is *very* happy rather than *moderately* happy, even if, in absolute terms, the differential in well-being is equally large. It would thus become easier for our reasons under the Maximization Requirement to be outweighed by other considerations, the better off our offspring. And even when the balance of reasons still favors our reasons under the Maximization Requirement, failing to comply with these reasons would constitute a lesser moral wrong, the better off the recipient. All I am claiming, in putting forward the Maximization Requirement, is that, however well-off one's offspring, there is always some pro tanto reason to make them better off still, if this is possible.

9. Solving the Non-Identity Problem: The Selection Requirement

The Threshold Requirement by itself suffices to account for the Normative Asymmetry. Moreover, I believe that having satisfied both the Threshold Requirement and the Maximization Requirement, a progenitor could not be said to have *wronged* her offspring in any way by creating him. Nonetheless, in order to capture not just the Normative Asymmetry and our intuitions about the *Same Person Case*, but also to provide the intuitively correct verdict in Parfit's *Non-Identity Case*, we must go one step further.

The Threshold Requirement and the Maximization Requirement are both so-called *narrow person-affecting* principles, in the sense that violating either requirement would be bad, or worse, *for some person*. Unfortunately, it is notoriously difficult to account for our intuitive judgments about Parfit's non-identity cases in narrow person-affecting terms. Thus, in the example from Section 3, it seems intuitively wrong for the woman to create person X, with a moderately happy life, rather than the numerically distinct person Y, whose life would be very happy. However, acting in this way is not bad, or worse, for any person. Moreover, note that if there *were* a narrow person-affecting explanation why creating person X with a moderately happy life would be morally wrong in the *Non-Identity Case*, the same would presumably have to be true in a case in which the woman's only two options were either to create X with a moderately happy life, or else to create no-one. But intuitively, there would be nothing wrong with creating X in latter scenario. [40]

Hence, my challenge in this final section is to show how the bearer-dependent account can support a non-person affecting principle that solves the Non-Identity Problem, without lapsing back into totalism.

I will first present what strikes me as the correct principle:

---

[40] For an interesting attempt to provide a narrow person-affecting solution to the Non-Identity Problem, see Elizabeth Harman, "Can We Harm and Benefit in Creating?", *Philosophical Perspectives* (2004), pp. 89-113. For a critical discussion of this proposal, see my "Future Persons and Victimless Wrongdoing" (ms).

The Selection Requirement: If I have a choice between creating a new person Y with a life worth living at well-being level W, and creating a numerically distinct person X with a life worth living at well-being level V, where W > V, then I have contrastive reason to bring about the former rather than the latter outcome, if I am to create a new person at all.

The Selection Requirement captures our intuition that in *Non-Identity Case*: given a choice between creating two numerically distinct children, one of whom would foreseeably have a significantly better life than the other, the woman has reason to create the happier of the two children, even though acting contrary to this reason would not be bad or worse for anyone. Thus the selection-requirement is not a narrow person-affecting principle in the sense defined above.

Unlike totalism, however, the Selection Requirement solves the non-identity problem without postulating an *unconditional* reason to create the happier life, on account of the well-being it contains. Like in the case of the Maximization Requirement, the reason for bringing about the outcome in which the happier child exists isn't unconditional but merely *contrastive*. Holding fixed the fact that she will create a new life, the Selection Requirement tells the woman to create the very happy person Y *rather than* the moderately happy person X. By contrast, in a choice between creating a very happy life and creating no new life at all, the Selection Requirement is silent. The Selection Requirement thus captures the way in which *Non-Identity Case*, in which the woman must choose between conferring a moderate *existential* benefit on X or a large *existential* benefit on Y is intuitively different from a case in which an agent must trade off conferring a moderate *ordinary* benefit on W against conferring a large *ordinary*

benefit on Z. Here, too, the agent has a contrastive reason to confer the larger rather than the smaller benefit. But in this case, her reason is not *merely* contrastive. Rather, the agent has an unconditional pro tanto reason to confer *either* benefit, with the stronger pro tanto reason *outweighing* the weaker. Thus, it is true for either benefit that, if the agent's choice was between conferring *that* benefit or conferring no benefit at all, she would have decisive moral reason to confer the benefit.

My final hurdle is to show how the Selection Requirement is supported by the bearer-dependent view of well-being that I have championed in this essay. Here is how: The choice faced by the woman in *Non-Identity Case* is between two possible outcomes. In Outcome 1, person X exists, and the standard of X's well-being is satisfied to a *moderate* degree. In Outcome 2, Y exists, and the standard of Y's well-being is satisfied to a *high* degree. Now, the standard that applies, conditional on creating X, and the standard that applies, conditional on creating Y, are not the *same* standard. This is so, because sameness of standard is a function, not just of the *kind* of moral standard (here, the standard of someone's wellbeing), but of the *identity* of the person who grounds the standard.

However, I believe we should affirm the following general principle:

> *The Principle of Standard Selection*: If I have a choice between bringing about Outcome 1 to which standard X applies, or bringing about Outcome 2 to which standard Y applies, and (i) standard X and standard Y are standards of the same kind; (ii) standard Y is satisfied to a higher degree in Outcome 2 than standard X is satisfied in Outcome 1; and (iii) all else is equal, then I have contrastive reason to bring about Outcome 2 rather than Outcome 1.

52

The Principle of Standard Selection gives plausible answers, not just in the case of procreation and the non-identity problem, but in cases involving other kinds of standards as well. Suppose I can choose between making a promise to person A, which I will only be able to keep to degree D < 100%, or making the same kind of promise to a different person B, which I can keep *perfectly*, and that all else is equal. (Assume, as above, that D is sufficiently high such that, in a case where my choice is only between making the promise to A and not making a promise at all, I would have no promissory reason against making the promise, and would count as having satisfied the promissory standard if I keep my promise to degree D). Although the standard that would obtain, conditional on making the promise to A is not the *same* standard that would obtain, conditional on making the promise to B, in the sense that my reasons to keep my promise in the two scenarios are grounded in the claim-rights of numerically distinct promisees, I believe that I nonetheless have a reason to make the promise to B, rather than to A. The Principle of Standard Selection captures this.

If the Principle of Standard Selection is correct, it implies the Selection Requirement for the specific case of procreation. Moreover, although neither the Principle of Standard Selection nor the Selection Requirement are narrow person-affecting principles, they can be explained and justified in terms of the bearer-dependent view of wellbeing that I have proposed in this chapter. According to the bearer-dependent view, there are no moral reasons to be exercised over the non-existence of a potential person whose life would have been well worth living, since there is no person *for whose sake* we have

reason to be exercised. However, if we do decide to create a person, we have bearer-dependent reason to want her life to go well. Indeed, we have reason to want her life to go as well as possible in an *absolute* sense: If the new person's life, while still worth living, is of mediocre quality (for instance, due to some congenital malady), there is reason to regret this for that person's own sake. This is true, even if, as a matter of *practical feasibility*, this person's life goes as well as it could.

It follows that, by creating the very happy child, the woman in *Non-Identity Case* ensures that the world as it is comes *closer* to the world as she would ideally want it to be than would be the case if she creates the less happy child. It is precisely because, for each person, conditional on her existence, we have bearer-dependent reasons to want her life to go as well as possible in the absolute sense, that in deciding *whom* to create, we should aim to select that person whose life we expect to go absolutely best.

10. Conclusion

If my arguments in this chapter are sound, I have charted a third way between the standard dichotomy, so familiar from the literature on population ethics, of impersonal totalist (or wide person-affecting) views, which can explain the Non-Identity Problem but must deny the Asymmetry, and narrow person-affecting views, which face the opposite problem. The view about the reason-giving force of human well-being that I have defended in this paper is *not* impersonal, because our reasons to confer well-being on people are bearer-dependent, not state-regarding reasons. People's well-being

54

matters because people matter, not because there being people with well-being contributes to a better state of affairs. Yet, as the discussion in the preceding section has revealed, the moral reasons that are supported by this bearer-dependent view of well-being go *beyond* narrow person-affecting principles and allow for contrastive reasons in cases where the people affected by an alternative are numerically distinct. We can thus reject totalism and uphold the Asymmetry, while simultaneously solving the Non-Identity Problem.

Appendix to Chapter 1: Roberts' Variabilism

According to Melinda Roberts' view, which she dubs *variabilism*, "*all* persons matter morally but they all matter *variably*."[41] Instead of dividing *people* up in accordance with whether they matter morally or not, Roberts proposes that we instead divide up *losses* and *gains* to people in accordance with whether *they* matter morally or not. As Roberts employs the term, a person incurs a 'loss' whenever agents could have created more well-being for that person and instead create less. More precisely, "to say that a person *p* incurs a *loss* at a given world *w* as a result of a given act *a* is to say that there was still another world *w'* accessible to agents at the critical time such that their performance of an alternate act *a'* at *w'* is better for *p* than their performance of *a* at *w* is."[42] Unlike me, Roberts rejects the claim that a possible world w can only be *better* (worse) for a person than a world z if the person is actual at both world w and world z. She thus maintains that it is *better for* a person to be created with a life worth living than not to be created, even though the person exists in only one of these two possible worlds. Indeed, Roberts goes further than this, holding that a potential person need not even exist at some world w in order to count as suffering a loss (or receiving a gain) at that *world w*. Roberts can thus maintain that failing to bring into existence a potential person whose life would have been well worth living is a loss to this potential person, while averting the creation

[41] Melinda Roberts, "The Asymmetry: A Solution," *Theoria* 77 (2011), p. 335.

[42] Ibid., p. 337.

of a miserable life is a gain for the potential person, who is thus prevented from coming into existence.

The central idea of the variabilist approach, however, is that not all losses matter morally. Rather, variabilism asserts that the moral significance of any loss depends on where that loss is incurred in relation to the person who incurs it. That is: "the loss incurred at a world where the person who incurs that loss does or will exist has *full moral significance*, while a loss incurred by that same person at a world where that person never exists at all has *no moral significance whatsoever*."[43] Furthermore, according to Roberts, *gains* matter just in case they avoid a morally significant loss.

Roberts' variabilism succeeds at capturing both halves of the Asymmetry. It matters morally that we do not create a miserable life, since doing so would impose a morally significant loss on a person. (The loss is morally significant because the subject of the loss exists at the world at which she incurs the loss, namely the world at which she exists with a miserable life). By contrast, failing to create a happy life is of no moral significance, according to variabilism: Though it represents a loss to a potential person, this person does not, and never will, exist at the world at which the loss is incurred.

Besides our disagreement about the logic of the *better/worse for*-relation, which I set aside in the following, I have two principal objections to Roberts' variabilist proposal.

_____

[43] Ibid., p. 356.

First, while Roberts' variabilism captures the two halves of the Asymmetry, it does so in a manner that will strike supporters of a symmetry view as objectionably *ad hoc*. The variabilist claims that while all losses matter morally provided they are incurred by people at worlds where these people exist, the same is *not* the case for gains. Rather, "gains have moral significance, not when those gains are accrued at the world at which the person who accrues those gains *exists*, but [only] when the *losses* those gains *avoid* on behalf of that person are incurred at worlds where *the person who incurs those losses exists*."[44] It is this purported difference between the moral significance of gains and losses that allows variabilism to track the intuitive moral asymmetry between creating a miserable life and failing to create a happy life.

However, lest she beg the question against a proponent of the symmetry view, the variabilist owes us an *explanation* and *justification* for this purported difference between losses and gains. If we think that avoiding losses for existing people always matters morally, then what is our reason for thinking that conferring gains on existing people does *not* always matter?

Unfortunately, Roberts has little to offer on this score, beyond the fact that positing such a difference allows her to capture the Asymmetry in a way that avoids inconsistency. She writes:

_____

[44] Ibid., p. 365.

Why not make gains matter at worlds where the person who accrues those gains exists? The simple answer is that our aim has been to determine whether we can situate the two halves of the Asymmetry in a principle that we are not then compelled on consistency or other conceptual grounds to reject. To do things the other way around – to make gains matter at worlds where the person who accrues those gains exists – goes no distance at all in achieving that aim. It is rather a way of failing to account for the Asymmetry.[45]

This passage essentially concedes that variabilism has no support beyond the fact that it captures the Asymmetry. For those puzzled by how the Asymmetry itself could be true, given the Objection from Symmetry, it offers no deeper explanation; nor does it help to fortify the Asymmetry against its critics. Variabilism is thus of limited explanatory and dialectical power.

My second objection is that – like the presentist, necessitarian, and actualist proposals surveyed above – variabilism cannot account for most people's intuition in *Non-Identity Case*, namely that it matters morally that the woman create person Y, whose life will be very happy, rather than person X, whose life, while still well worth living, will go considerably less well. If variabilism is right, this intuition cannot be correct.

According to Roberts, gains to existing people matter only insofar as they prevent a loss to someone who would have existed at the world in which the loss occurs. But this isn't the case for the people in the non-identity case. They exist *only* in the world in

_____

[45] Ibid., p. 365.

which they gain by being created; in the world at which they lose by not being created, they do not exist. Therefore, the fact that person Y would gain more, if he is created, than person X, if he is created, is without moral significance, according to variabilism. For neither gain prevents a morally significant loss. Therefore, variabilism lacks the resources to account for the non-identity intuition.

Roberts herself does not consider this as a problem, since, surprisingly, she appears not to share the non-identity intuition:

> We can discern no (…) morally significant loss [in such cases]. On the other hand, neither do we clearly in such cases discern wrongdoing. (When it is maximizing for everyone else, is it really wrong to bring a genetically impaired but happy child into existence rather than a less impaired and happier child into existence?)[46]

For those of use who do have the intuition, and who would consider jettisoning it in order to account for the Asymmetry too high a price to pay, variabilism is not the way forward.

---

[46] Ibid., p. 362.

Chapter 2: The Evaluative Asymmetry and the Mere Addition Paradox

1. Introduction

Population ethics is often described as a field littered with paradoxes. We have already encountered one notoriously difficult problem in Chapter 1, namely how to reconcile the Normative Procreation Asymmetry with our intuitions about the Non-Identity Problem. As we saw, traditional approaches managed to capture our considered convictions about either the Non-Identity Problem or the Procreation Asymmetry, but only at the price of forfeiting those about the other. By contrast, the bearer-dependent view of wellbeing that I introduced in Chapter 1, according to which our reasons to confer wellbeing on a person are conditional on the fact of her existence, managed to successfully account for our intuitions in both cases.

The questions we explored in Chapter 1 concerned *normative* judgments about our moral reasons for action. By contrast, in this chapter, I shall explore the manifold complications that arise when we attempt to defend the *evaluative* claims that correspond to the normative claims established in Chapter 1 – that is, claims about the *goodness* of outcomes produced by our procreative decisions. After surveying a range of alternatives, none of which prove satisfactory, I argue that our best hope of making sense of the evaluative version of the Asymmetry intuition is by defending a theory

about the way in which truths about the relative goodness of outcomes are connected to our reasons for action: what I call the *biconditional buck-passing view of outcome betterness*.

This theory does not just allow us to make sense of the evaluative version of the Asymmetry intuition. It also allows me to provide a novel solution to Derek Parfit's famous Mere Addition Paradox.

2. From the Normative to the Evaluative Procreation Asymmetry: the Challenge

In Chapter 1 I argued in defense of

> *The Normative Procreation Asymmetry*:
> If a future person would foreseeably have a life that is not worth living, this in itself gives us a strong moral reason to refrain from bringing this person into existence.
> By contrast, there is no moral reason to create a person whose life would foreseeably be worth living, just because her life would be worth living.

Notably, I defended the second conjunct of the Asymmetry for *all* levels of positive wellbeing that a potential future person might enjoy. That is, I claimed that *however* good a potential future person's life could foreseeably be, this fact does not, in itself, give us a moral reason to bring this person into existence.

In addition, in discussing Parfit's Non-Identity Problem, I affirmed what I shall now call the

> *The Normative Same Number Claim*:
> If I have a choice between creating a new person with a life worth living at well-being level W, and creating a new person with a life worth living at well-being level V, where W > V, I have contrastive reason to bring

62

about the former rather than the latter outcome, if I am to create a new person at all.[47]

This proposition is true, I argued, whether or not the two potential lives in question are numerically identical or not.

The above claims are *normative*, i.e. they are claims about our reasons for action. In the present chapter, I shall explore the implications of embracing two corresponding *evaluative* claims, namely

> *The Evaluative Procreation Asymmetry*:
> It makes the world go worse, all else equal, to create a life that is foreseeably not worth living.
> By contrast, it does not make the world go better, all else equal, to create a life that is foreseeably worth living.

and

> *The Evaluative Same Number Claim*:
> If I have a choice between creating a new person with a life worth living at well-being level W, and creating a new person with a life worth living at well-being level V, where W > V, it makes the world go better, all else equal, to produce the former rather than the latter outcome.[48]

---

[47] The Normative Same Number Claim combines the Maximization Requirement and the Selection Requirement defended separately in Chapter 1.

[48] Throughout this chapter, I will phrase comparative evaluative claims in terms of which of two alternative courses of action "makes the world go better" or "produces the better outcome". I employ these locutions instead of the more common "produces the better *consequences*" in order to make clear that the action itself is part of the *evaluandum*, and not just the end-state that results from the action. However, for the sake of brevity or when paraphrasing other writers, I sometimes also speak simply in terms of which "outcome" is best. I will take it as read that in these instances, the evaluation of "outcomes" is meant to include the evaluation of associated actions.

Both these evaluative claims seem to me intuitively extremely plausible. The Evaluative

Same Number Claim is very widely shared amongst population ethicists. It is entailed,

for instance, by Derek Parfit's *Same Number Quality Claim*, according to which "[i]f in

either of two possible outcomes the same number of people would ever live, it will be

worse if those who live are worse off, or have a lower quality of life, than those who

would have lived."[49] The first conjunct of the Evaluative Procreation Asymmetry also

seems incontrovertible. How could it not make the world go worse, all else equal, to add

to it a life that is, on the whole, so miserable as to be not worth living? The second

conjunct of the Evaluative Procreation Asymmetry, too, has struck many philosophers

as highly plausible – even those who believe that they are ultimately forced to reject it

for theoretical reasons. Thus, John Broome, in *Weighing Lives* claims to be strongly

attracted to what he calls the "Intuition of Neutrality", according to which there is a

range of positive levels of well-being at which adding a person to an existing population

is "ethically neutral", in the sense that this person's existence makes the world go

neither better nor worse.[50] (This is in contrast to the view that Broome ultimately adopts,

---

[49] Derek Parfit, *Reasons and Persons* (Oxford University Press, 1984), p. 360.

[50] The neutral range can either be upwardly unbounded, in which case the Intuition of Neutrality entails the second conjunct of the Evaluative Procreation Asymmetry; or it can be upwardly bounded, such that lives whose level of wellbeing exceeds the upper bound of the neutral range make the world go better.

namely that there is only a *single*, albeit vague, level of lifetime well-being at which the existence of a person makes the world go neither better nor worse).[51]

But, besides the inherent appeal of these evaluative propositions, there is a further reason why I set out to defend them in this chapter. Given that I have argued in support of the Normative Procreation Asymmetry and the Normative Same Number Procreation Claim in Chapter 1, I believe that I am *committed* to the corresponding evaluative claims. In this, I differ from Michael Tooley, who attempts to uphold the Normative Procreation Asymmetry while denying the second conjunct of the Evaluative Asymmetry.[52] Tooley maintains that although it would make the world better, all else equal, to create a new happy life, there is *no* moral reason – not even a pro tanto reason – to do so.

This is not a tenable combination of claims, I believe. It is part of the conceptual constraints on our concept "good" that it cannot be normatively irrelevant in this way. The fact that one available course of action makes the world go better than the other must make at least a defeasible difference to what I have reason to do. I agree with Christine Korsgaard that the function of our concept "good" is "to mark out schematically the solutions to certain kinds of problems which we have to solve".[53]

---

[51] See John Broome, *Weighing Lives* (New York: Oxford University Press, 2004), chapters 10-12.

[52] See Michael Tooley, "Value, Obligation and the Asymmetry Question", *Bioethics*, Vol. 12, No. 2 (1998), pp. 93-110.

[53] Christine Korsgaard, "On Having a Good" (unpublished manuscript), pp. 22-3.

When inquiring which of two courses of action available to me makes the world go better, I believe that (at least part of) the problem to which claims about betterness provide the answer is "which course of action do I have more reason to choose?". Indeed, it is unclear what someone could *mean* who maintained that, although I have *no* reason – not even a defeasible pro tanto reason – to choose course of action A over course of action B, A would make the world go better than B, all things considered. In the absence of any normative implications, what could this claim about betterness possibly amount to?

I shall have more to say on the relationship between the goodness of outcomes and our reasons for action in Section 8. But, if my brief remarks above are plausible, they already support the following Bridge Principle between claims about what makes the world go better and claims about reasons for action:

> *Bridge Principle*:
> If, of two possible courses of action available to me, option A makes the world go better than option B all things considered, I have a defeasible pro tanto reason to choose option A rather than option B.

In putting forward this principle, I am not ruling out at this point that, even if A makes the world go better than B, my reason for choosing A rather than B can be outweighed, silenced, cancelled, etc. Nor am I claiming that if, of two courses of action available to me, I do not choose the one that makes the world go better, when this is *also* the option that I have most moral reason to perform, my action need be morally *wrong*. This presupposes a view about the relationship between moral reasons and wrongness – one

66

according to which it is always wrong not to do what I have stronger moral reason to do – that I reject.[54] Finally, in putting forward the Bridge Principle, I am not committing myself to a view about the *order of dependence* between facts about the comparative goodness of outcomes and my reasons for action. The Bridge Principle states only a material implication; it makes no claims about grounding or priority.

The Bridge Principle and the second conjunct of the Normative Procreation Asymmetry together entail the second conjunct of the Evaluative Procreation Asymmetry. If I have *no* reason to create a new life just because it would be worth living, then, by the contrapositive of the Bridge Principle, creating such a new life does not make the world go better, all else equal. This means that someone who, like myself, embraces both conjuncts of the Normative Procreation Asymmetry must be prepared to defend both conjuncts of the Evaluative Procreation Asymmetry as well. But, as we shall see from the next section onwards, this is far from a trivial undertaking. Indeed, I believe that embracing the Evaluative Procreation Asymmetry together with the Evaluative Same Number Claim will ultimately give us reason to reject a very widely held view about the goodness of outcomes or ways the world can go: what Larry Temkin has called the *Internal Aspects View of Outcome Goodness*.

---

[54] For instance, I believe that a course of action A can be *supererogatory* – in which case, I may have more moral reason to do A than B, yet it wouldn't be wrong of me not to do A.

3. The Unacceptable Tetrad

Larry Temkin characterizes the Internal Aspects View of Outcome Goodness as follows:

> *The Internal Aspects View of Outcome Goodness (IAVOG)*:
> "Roughly, for each outcome, O, how good that outcome is all things considered depends solely on how good it is with respect to each moral ideal that is relevant for assessing the goodness of outcomes, and on how much all of the relevant ideals matter vis-à-vis each other, where these depend solely on O's internal features. Moreover, for any two outcomes, O_1 and O_2, O_1 will be better than O_2 all things considered if and only if the extent to which O_1 is good all things considered, as determined solely on the basis of O_1's internal features, is greater than the extent to which O_2 is good all things considered, as determined solely on the basis of O_2's internal features. In addition, if O_1 is better than O_2 all things considered, the extent to which this is so will depend solely on the extent to which O_1 is good, all things considered, is greater than the extent to which O_2 is good, all things considered."[55]

If the IAVOG is correct, then how good an outcome O_1 is does not depend on which outcome it is being compared to, nor on what other outcomes are available. (Call the set of all outcomes that an agent can bring about in a given situation the "option set"). Rather, it is solely a function of the "internal", or intrinsic, properties of O_1. Moreover, whether, and by how much, O_1 is better or worse than another outcome O_2 again depends solely on the intrinsic properties of O_1 and O_2; it cannot be affected by other

---

[55] Larry Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Oxford University Press, 2012), p. 370.

outcomes in the option set. A theory of goodness that conforms to IAVOG will thus respect what economists call "basic contraction consistency" and "basic expansion consistency", often referred to together under the name "Independence of Irrelevant Alternatives" (though this label is not used consistently).[56] The version of the principle that I will work with in this chapter states:

> *The Independence of Irrelevant Alternatives Principle (IIAP)*:
> If outcome O_1 is better than outcome O_2 when the option set is S = {O_1, O_2, …, O_n}, then O_1 must still be better than O_2 when the option set is contracted to T, where T is a *subset* of S that contains O_1 and O_2. (*Contraction consistency*) Likewise, if O_1 is better than O_2 when the option set is T, then O_1 must be better than O_2 when the option set is expanded to S. (*Expansion consistency*)

If the comparison of O_1 and O_2 depends only on the intrinsic properties of these two outcomes, it cannot be affected by the presence or absence of further alternatives from the option set.

Utilitarian moral views, such as total or average utilitarianism, conform to the IAVOG. According to these views, there is a *single* moral ideal that is relevant to assessing the goodness of a given outcome: total or average utility, respectively. How well this ideal is realized is a function of intrinsic properties of that outcome: what

---

sentient creatures there are in that outcome and their levels of utility. It does not depend on properties of other outcomes. But value *monists* like utilitarians are by no means the only ones who can subscribe to IAVOG. Value *pluralists*, who hold that there are a number of different moral ideals – total utility, beauty, and friendship, say – that are relevant to assessing the goodness of an outcome, can support IAVOG as well. As long as the various moral ideals that together determine the goodness of an outcome are such that we can assign the outcome a goodness "score" which is solely a function of its intrinsic properties, a pluralist theory of outcome goodness will conform to the IAVOG. Indeed, I believe that moral theories that have a *teleological* structure, in the sense discussed in Chapter 1, will in general tend to conform to IAVOG. For the teleologist, the reasons to bring about a certain outcome, to make the world go a certain way, are *state-regarding* reasons. They are reasons that we have in virtue of the overall goodness of the outcome, which in turn is a function of all the good or valuable things that a state of affairs contains.

However, I will now show that, paired with a standard assumption about comparative evaluative relations, the IAVOG generates extremely implausible implications, or even outright incoherence, if we affirm both the second conjunct of the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim.

The assumption about comparative evaluative relations is this:

> *Completeness of the Standard Trichotomy of Comparative Value Relations ("Completeness")*: For any two outcomes O₁ and O₂ that we compare in terms of their all-things-considered goodness, exactly one of the following three comparative relations must hold: either O₁ is *better than*

O_2, or O_1 is *equally good as* O_2, or O_1 is *worse than* O_2. Call this the *Completeness* assumption for short.

If IAVOG and Completeness both hold, there is an ordinal ranking over all possible outcomes that is complete and transitive. The goodness of any outcome can be represented by a position on a single linear scale, which indicates whether it is better, equally good, or worse than some other outcome. Moreover, IAVOG and Completeness together entail the following principle:

> *Like Comparability of Equals*: If outcomes O_1 and O_2 are equally good, all things considered, then if O_1 is all-things-considered better (worse) than some third outcome O_3, O_2 is all-things-considered better (worse) than O_3.[57]

The tension between the Evaluative Asymmetry and the Evaluative Same Number Claim on the one hand, and IAVOG on the other, can be brought out in two steps. Let O_{VHL} stand for the outcome in which I create a new life that is, on the whole, very happy; let O_{MHL} stand for the outcome in which I create a new life that is, on the whole, moderately happy; let O_{NL} stand for the outcome in which I create no new life.

Step 1: According to the Evaluative Same Number Claim, O_{VHL} is better than O_{MHL}. Moreover, if IAVOG is true, O_{VHL} is better than O_{MHL} in virtue of the intrinsic properties of these two outcomes. The outcomes could be arranged like this on a linear scale

---

[57] This is a narrower version of Temkin's "Like Comparability for Equivalents" principle. See *Rethinking the Good*, p. 237.

representing their goodness:

O_MHL   O_VHL

Goodness of outcome
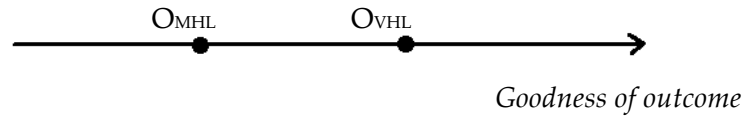
Figure 2.1

<u>Step 2</u>: According to the Evaluative Asymmetry, O_NL, the outcome in which I produce no new life, is not worse than O_MHL, the outcome in which I produce a moderately happy life. By Completeness, it follows that either O_NL is equally good as O_MHL or O_NL is better than O_MHL. In terms of *Figure 2.1*, this means that O_NL must not be located to the left of O_MHL. But likewise, the Evaluative Asymmetry implies that O_NL is not worse than O_VHL, the outcome in which I produce a *very* happy life. So, on the linear scale, O_NL must not be located to the left of O_VHL either, which means that it must be located to the right of O_MHL. This follows from the assumption of Completeness: If O_NL is *not worse* than O_VHL, then it must be either equally as good as or better than O_VHL. But then, by either the Like Comparability of Equals or the transitivity of "better than", O_NL must be better than O_MHL. And again, assuming IAVOG is correct, these things are true solely in virtue of the intrinsic properties of these three outcomes.

There are two problems that immediately become apparent. First of all, the conclusion that O_MHL is worse than O_NL is inherently quite implausible, at least in some circumstances. As I argued in Chapter 1, if a moderately happy life is the *best* life that I

72

can create, either now or in the future, then there is no moral reason not to create this new life. But this sits ill with the claim that, just in virtue of the intrinsic properties of these two outcomes, my creating a moderately happy life is a *worse* outcome than not creating a new life at all.

But, of course, the problem goes much further than that. A moment's reflection reveals that O_{NL} must be better than O_{VHL} as well. For we can imagine creating a yet happier life (an "ecstatic life"). And, naturally, O_{EL}, the outcome in which such an ecstatic life is created would be better than O_{VHL}. But then we can re-run the argument of Step 2 above. By the Evaluative Asymmetry, O_{NL} must not be worse than O_{EL}, which means that it must be better than O_{VHL}.

And why stop there? As I noted in Chapter 1, there is every reason to suppose that a person's lifetime wellbeing is potentially upwardly unbounded. For any life, however happy and fulfilled, we can imagine ways of making it go better yet for the person. And, plausibly, if the *personal* good of lives is in principle upwardly unbounded, so too is the *impersonal* goodness of creating such lives. That is, for any life, however happy, we can imagine a yet happier life such that the world goes better if the latter rather than the former life is created is better. Yet, if the Evaluative Asymmetry is correct, there is no life, however happy, such that creating such a life, in itself, makes the world go better than not creating this life.

Mathematically this would be possible, on the assumption that the impersonal goodness of creating happy lives increases asymptotically towards some limit, and that

the goodness of O_{NL} is located at or above that limit. But, for one thing, such an assumption seems very ad hoc. I can think of no good independent reason for making it. For another, even if this assumption spares us from outright incoherence, the resulting position is still implausible in the extreme. It would imply that creating even the most blissful life makes the world go worse than not creating any new life.[58]

Let us briefly recapitulate the argument of this section. We have seen that the combination of the following four propositions has unacceptable implications:

(A): The second conjunct of the Evaluative Procreation Asymmetry, according to which creating a happy life rather than creating no life does not make the world go better, all else equal.

(B): The Evaluative Same Number Claim, according to which creating a very happy life makes the world go better, all else equal, than creating a moderately happy life.

(C): The Internal Aspects View of Outcome Goodness, which holds, roughly, that the goodness of an outcome is solely a function of its intrinsic properties, and does not depend on contextual factors, such as which outcome it is being compared to or the other alternatives in the option set.

---

[58] Even the anti-natalist David Benatar would not embrace this extreme view. In a nutshell, Benatar's position in *Better Never to Have Been* is that not creating a new life is better than creating a new life that contains any personal bads (i.e. anything that *tends* towards making the life not worth living for the person). The presence of personal *goods* in a life, i.e. of things that tend towards making the life worth living, cannot compensate for the presence of personal bads, according to Benatar. Furthermore, Benatar assumes that *in practice* even the happiest lives will contain some personal bads. That is why he embraces anti-natalism, the view that it is prima facie wrong to create new lives. But, *in theory*, nothing stops us from imagining a blissful life that contains no personal bads, only personal goods. Creating such a life, even Benatar would concede, would not be worse than producing no new life.

(D): The Completeness assumption, which maintains that for any two outcomes which we compare in terms of their all-things-considered goodness, one must be better than the other, or they must be equally good.

Beginning in the following section, I will examine various options for avoiding the untenable implications of this tetrad, by giving up one or more of its members. Doing so will take us deeper into recent work in the field of population ethics.

4. Rejecting the Evaluative Asymmetry: Totalism, Averagism, and Critical Level Theories

Of all the possible ways of avoiding the tension between propositions (A)-(D), giving up proposition (B), the Evaluative Same Number Claim, seems to me the least promising. Indeed, with the possible exception of David Heyd, whose "generocentric" view I briefly discussed in Chapter 1, I know of no population ethicist who is tempted to reject this claim. For this reason, I shall not further pursue this option. The Evaluative Same Number Claim seems so obviously correct that denying it would function as a *reductio* of any view that lead to this conclusion.

By contrast, there is no shortage of population ethicists who reject (A), the second conjunct of the Evaluative Procreation Asymmetry. Most such philosophers maintain that there is only a *single* level of lifetime wellbeing such that creating a life with this level of wellbeing is equally good as not creating a new life. Call this level of wellbeing the *Impersonal Zero Level*. Creating a life whose wellbeing is above the Impersonal Zero

Level makes the world go better; creating a life whose wellbeing is below that level makes the world go worse.[59][60]

Jettisoning the Evaluative Procreation Asymmetry in favor of the view just sketched allows us to hold on to propositions (B), (C), and (D), while avoiding the problems described in the previous section. Assuming, as totalist utilitarians do, that the Impersonal Zero Level coincides with the *Personal* Zero Level, i.e. the level that separates levels of lifetime wellbeing at which a life is worth living for a person from levels of wellbeing at which it is not worth living, we obtain the following simple ranking of our three cases from the previous section: $O_{VHL} > O_{MHV} > O_{NL}$. (Here and in the following,

---

[59] As I noted above, John Broome's position in *Weighing Lives* is a more complicated version of this view. Broome maintains that there is only a single Impersonal Zero Level (what he calls the "Neutral Level"), but that this level is *vague*. We cannot state this level with precision, but only affirm that it lies within a certain range of wellbeing. Moreover, this range is potentially quite wide, according to Broome. For a large range of levels of wellbeing it is therefore the case that, if we add a person with this level of wellbeing to the population, we cannot *affirm* that the resulting distribution is either better or worse than the original population.

[60] One alternative worth mentioning is average utilitarianism. According to average utilitarianism, it will also be true, in any given situation, that there is a single level of lifetime wellbeing at which the addition of a new person makes the world go neither better nor worse, namely the level at which the addition of the person does not change the level of average wellbeing in the world. If the person's lifetime wellbeing is above this level, it makes the world go better; if it is below it, it makes the world go worse. Unlike the view sketched above, however, average utilitarianism implies that this level of wellbeing isn't fixed, but rather entirely dependent on contextual factors, namely the average level of wellbeing in the rest of the population. This feature of average utilitarianism is also at the root of well-known and convincing objections to this view: For instance, average utilitarianism implies that, for a population consisting of some number of people leading extremely miserable lives, the outcome is improved by adding another person whose life is only slightly less miserable. Likewise, in a world in which everyone's life is blissfully happy, it makes the world go worse to create a new person whose level of wellbeing is only slightly less than the average in the population. Both of these implications of average utilitarianism are exceedingly implausible, and give us reason to reject the view.

">"denotes "is all-things-considered better than"; "<" denotes "is all-things-considered worse than"; and "=" denotes "is equally good as").

However, abandoning (A) comes with costs. I have already commented above on the inherent plausibility of the Evaluative Asymmetry, as well as on the fact that it would be implausible to deny it while embracing the corresponding normative claim. In addition, however, rejecting the Asymmetry invites trouble of other kinds.

If adding new lives above the Impersonal Zero Level makes the world go better, all else equal, this should, in principle, be able to outweigh other changes that tend towards making the world go less well. Parfit's simplest argument for the Repugnant Conclusion exploits this to make trouble for totalism and cognate views. The Repugnant Conclusion is the following claim:

> *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.

This claim, Parfit notes, strikes us as intuitively unacceptable. How could a world in which everyone has a life that is barely worth living be better than one in which everyone who exists enjoys a very high quality of life? And yet, once we accept that containing more worthwhile lives makes the world go better, all else equal, this conclusion can seem hard to escape: In *Figure 2.2* below, let A represent an outcome in which 10 billion people live lives with a very high level of wellbeing. (The height of a box represents the wellbeing of people in a population, the width of the box represents

the number of people in the population). B is an outcome in which every person is somewhat less well off than in A; however, if the population in B is sufficiently more numerous than the population in A, bringing about B may make the world go better, all things considered, than bringing about A.
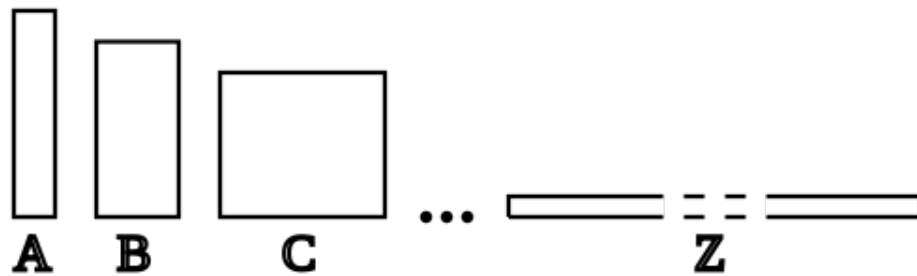
Figure 2.2

Likewise in a comparison between outcomes B and C, between C and D, etc., all the way to outcome Z, which consists of a huge population in which everyone's life is barely worth living. By transitivity, since B is better than A, C is better than B, and so on down to population Z, it follows that the world goes better in outcome Z than in outcome A. Call this the *Spectrum Argument* for the Repugnant Conclusion.

A proponent of the Evaluative Asymmetry, by contrast, rejects the very first step of the Spectrum Argument. Since, according to him, containing more happy lives does not make the world go better all else equal, there is no reason to accept that B is better than A (nor that C is better than B, etc.). The Spectrum Argument never gets off the ground.

Friends of totalism may attempt to draw the sting of the Spectrum Argument by modifying their view. So-called *critical level* theories attempt to avoid the Repugnant Conclusion by positing a split between the value that a life has for the person who lives it and its "contributory" value, i.e. the contribution it makes to the world's going better or worse. Whereas standard versions of totalist utilitarianism assume that any life that is worth living for a person (and thus has positive personal value) also has positive contributory value, critical level theorists deny this: a person's life only contributes positively to an outcome if the quality of the person's life exceeds some positive critical level. If the person's wellbeing is below this level, the contributory value of her life is negative. On the simplest version of the critical level view, the contributory value of a person's life is her wellbeing minus the critical level.

*Figure 2.3* illustrates how critical level views avoid the Repugnant Conclusion:
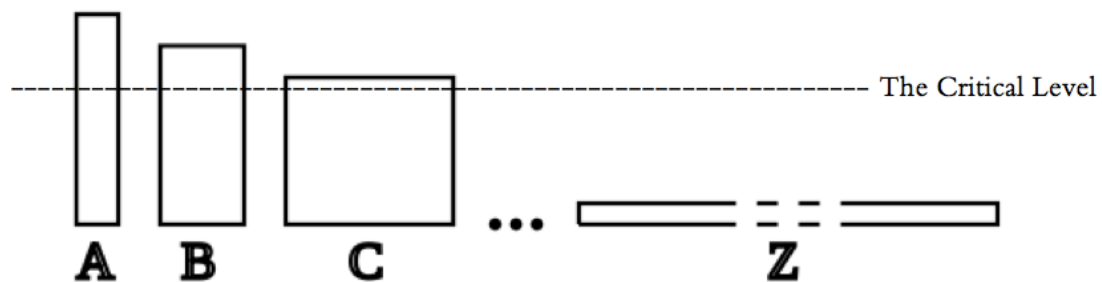
Figure 2.3

If the critical level is higher than the quality of the lives led by the people in outcome Z, the Repugnant Conclusion is blocked. In the case illustrated by *Figure 2.3*, the logic of the

Spectrum Argument can only take us as far as the conclusion that outcome C is better than outcome A. Since the people in outcome C lead lives that are considerably better than those of the people in outcome Z, this will strike many as less counterintuitive than the Repugnant Conclusion itself. (I myself confess to finding even this claim strongly unappealing). By contrast, since the people who live in Z have lives that are all below the critical level, Z constitutes an outcome that is unambiguously worse than A.

However, critical level views give rise to serious objections of their own. Critical level theorists avoid the Repugnant Conclusion by divorcing the personal and contributory value of human lives. But, by the lights of their own theory, this move seems ad hoc. We may ask: in virtue of what does the existence of happy human lives *above* the critical level have contributory value? The answer that most critical level theorists are inclined to give is: 'because these lives are *good for* those who live them.' But if that is the correct explanation for ascribing contributory value to lives *above* the critical level, shouldn't it apply *pari passu* to worthwhile lives below the critical level? For these lives, too, are good for those who live them.

Critical level views also give rise to counterintuive implications that are avoided by pure totalism. The best-known of these is Gustaf Arrehennius's

> *Sadistic Conclusion:* For any number of potential lives with very negative welfare, there are situations in which, according to the critical level

theory, it would make the world go better to add *these* lives rather than some larger number of lives with positive welfare.[61]

For instance, according to the critical level view, it could be better to add to the world a small number of miserable lives, rather than a large number of lives that are worth living but whose wellbeing falls below the critical level, because the negative contributory value of the latter addition exceeds that of the former. This is extremely counterintuive.

5. Rejecting Completeness: the Appeal to Different Number-Based Imprecision

A second option for dealing with the tension between propositions (A)-(D) is to challenge proposition (D), the Completeness assumption, which maintains that for any two outcomes which we compare in terms of their all-things-considered goodness, one must be better than the other, or they must be equally good. It has been suggested that alongside the traditional trichotomy of comparative evaluative relations, it is plausible to postulate at least one further relation:

> *Parity:* Two things are on a par in value if and only if neither is better than the other, yet they are also not precisely equally good.[62]

---

[61] See Gustaf Arrhenius, "An Impossibility Theorem for Welfarist Axiology", *Economics and Philosophy* 16 (2000), pp. 247–266.

[62] See, for instance, Ruth Chang, "The Possibility of Parity", *Ethics* 112 (2002), pp. 659-688.

Derek Parfit, who introduced the same notion under the name "imprecise equality" in *Reasons and Persons*, now maintains, in unpublished work, that there are in fact, *six* comparative value relations, since judgments of better- or worse than may also sometimes be imprecise.[63] In the following, I will employ "is on a par with" and "is imprecisely equal to" as synonyms.

We are prepared to encounter parity or imprecise equality in contexts where the two things being compared are sufficiently different in nature for us to judge that a *precise* comparison between their goodness is impossible. A popular illustration is artistic greatness across different domains of creation. Suppose you were asked: "Who was the greater artist: William Shakespeare or J.S. Bach?" You may be inclined to respond: "Neither of the two was a greater artist than the other." But must this imply that Shakespeare and Bach were *exactly equally* great artists? Surely not. The two domains of creation in which they were active – literature and music – are too different to permit such precise comparisons. An upshot of this is that if we imagined an artist, call him "Shakespeare Minus", who, by assumption, was a *somewhat* less great writer than Shakespeare, it does not follow that Shakespeare Minus was a less great artist than J.S. Bach. (We assume that amongst writers more precise comparisons of artistic greatness are possible than between writers and composers).

_____

[63] Derek Parfit, "Towards Theory X: Parts I and II" (unpublished manuscript).

On the other hand, the notion of parity must not to be confused with that of full-blown *incomparability.* If the artistic greatness of writers and composers were incomparable, then it couldn't be said of *any* writer that he was a less great artist than the composer J.S. Bach. This is plainly not the case. There are countless writers, say Arthur Conan Doyle, who are much less great artists than J.S. Bach.

Suppose we accept, *arguendo*, that two outcomes containing a *different* number of lives may often be on a par in goodness, since comparisons between populations of different sizes are *necessarily imprecise*. Call this the suggestion of *different number-based imprecision*. One philosopher who has advanced this proposal is Derek Parfit in "Towards Theory X: Part II".[64] If we accept different number-based imprecision, we can make the following move: The Evaluative Asymmetry commits us to the claim that (i)

---

[64] I should note that Parfit doesn't put forward different number-based imprecision in an attempt to defend the Evaluative Asymmetry. In fact, Parfit rejects the Evaluative Asymmetry. Instead, Parfit's view can be summed up in two claims:

> *V1:* When we add a person who has a worthwhile life to an existing population, this makes the outcome *in one respect* better. This is true in general, even for lives that are barely worth living.

> *Different number-based imprecision:* Adding a new person also increases imprecision.

According to Parfit, we must therefore distinguish between two types of cases where we add a new life worth living to an original population P. In the first case, the additional goodness from creating the new person is not enough to "overcome" different number-based imprecision. (The added value from creating the new person does not exceed what Parfit calls the "margin of imprecision"). Although the new population N_1 is *in one respect* better than the original population P, it is still, all things considered, imprecisely equal in value to the original population P. In the second case, the wellbeing of the person added is above what Parfit calls the *High Level*. In that case, the additional goodness from the new person is sufficiently large to overcome number-based imprecision. The new outcome N_2 is *all things considered* better than P, although it is still better by an imprecise amount. The view that I examine in this section differs from Parfit's in two respects: it rejects V1 and denies that there is a High Level.

having a moderately happy child does not make the world go better than having no

child – ¬(O_{MHL} > O_{NL}) – and likewise to the claim that (ii) having a very happy child does

not make the world go better than having no child – ¬(O_{VHL} > O_{NL}). The Evaluative Same

Number Claim commits us to the claim that (iii) having a very happy child makes the

world go better than having a moderately happy child – O_{VHL} > O_{MHL}. If we accept

different number-based imprecision, we could then cash out these three claims as

following:

> (i*) If I have a moderately happy child, the world goes *imprecisely equally*
> *as well* as it does if I have no child: O_{MHL} ≈ O_{NL}.

> (ii*) If I have a very happy child, the world goes *imprecisely equally as well*
> as it does if I have no child: O_{VHL} ≈ O_{NL}.

However, since the relation of imprecise equality is not transitive, affirming and (i*) and

(ii*) does not prevent us from affirming:

> (iii*) If I have a very happy child, the world goes *better* than if I have a
> moderately happy child: O_{VHL} > O_{MHL}.

The appeal to different number-based imprecision also helps us to avoid the

implications of the Spectrum Argument for the Repugnant Conclusion. For even if it

were the case that A is imprecisely equally good as B, B is imprecisely equally good as C,

and so on, all the way to outcome Z, it would not follow that A is imprecisely equally

good as Z. For the relation of imprecise equality is not transitive.

Note, finally, that rejecting Completeness and introducing imprecise equality as a

fourth comparative value relation does not force us to reject the Internal Aspects View of

84

Outcome Goodness. It may still be true that the goodness of an outcome is only a function of its intrinsic properties.[65]

Unfortunately, despite its attractions, I believe that the appeal to different number-based imprecision does not provide a fully satisfactory way of cashing out the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim. I shall mention three reasons why I think this is the case.

First, though I am convinced by the case for imprecise comparative relations in some contexts, I find the claim of *different number-based* imprecision insufficiently motivated. Relations of imprecise comparability typically hold between items that are qualitatively sufficiently different to vitiate precise comparisons with regard to some evaluative dimension. In the Shakespeare vs. Bach example, a judgment of imprecise equality seems justified by the fact that that the qualitative differences between music and literature are sufficiently great to render impossible *precise* comparisons of artistic greatness across these different domains. But, *prima facie*, it is not clear that a mere difference in *size* between populations makes for an important qualitative difference. In the absence of a clear rationale, the appeal to different number-based imprecision thus

---

[65] What we have to give up, by contrast, is the claim (entailed by the combination of IAVOG and Completeness) that the goodness of any outcome can be represented by a position on a *single* linear scale, which indicates whether it is better, equally good, or worse than some other outcome.

appears *ad hoc*.[66] (In Section 9, I supply a possible rationale for affirming judgments of imprecise equality in some contexts in population ethics. But, as we shall see, this is not *different number-based* imprecision).

Second, if we think that two outcomes containing different numbers of lives are, just for that reason, imprecisely comparable and often *on a par* in value, then this should be true, not just to cases where we add a new life worth living to an existing population, but also to cases where we create a new life that is not worth living. That is, the outcome that results from adding a miserable life to an existing population – O_{ML} – may be *on a par* with the outcome of not creating a new life, O_{NL}. This, however, contradicts most people's firm intuition that creating a new life that is so miserable as to be not worth

---

[66] This worry seems to me to arise in connection with Parfit's own appeal to different number-based imprecision in his draft "Towards Theory X: Part II". Parfit appears to affirm different number-based imprecision as a freestanding substantive commitment. (As far as I can tell, it doesn't follow from any other substantive claims he makes). This might appear particularly worrisome in a case like the following:

$$D = (5, 5, 5, \Omega)$$

$$D+ = (5, 5, 5, 5)$$

Let D and D+ represent two populations with three and four members, respectively, all of whose wellbeing is at level "5". ("Ω" denotes non-existence). Let 5 be a level of wellbeing at which a life is worth living, but below the High Level. Since the extra life in D+ is below the High Level, Parfit's view implies that D and D+ should be imprecisely equal in value. But why exactly should this be the case? The presence of the fourth person in D+ does not change average wellbeing, compared to D, but it does increase total wellbeing. According to V1, this makes D+ in one respect better than D. D+ is also just as equal as D. On all dimensions that Parfit identifies in his text as potentially determining the goodness of an outcome, D+ appears to do as well as or better than D. Given this, why should Parfit deem D+ to be only imprecisely equally good as D, and not better? I don't think Parfit can here appeal to a claim he makes earlier in the text, namely that "when two things are only imprecisely comparable, one of these things may be in one way better than the other, and in no way worse, without being better all things considered." For what I am questioning is, precisely, his rationale for deeming D and D+ to be only imprecisely comparable.

living makes the world go unambiguously worse, all else equal. Thus, while appeal to imprecise comparability can help us to accommodate the second conjunct of the Evaluative Procreation Asymmetry, it threatens to undermine the first.

The final problem with appealing to different number-based imprecision is more complicated. It arises in connection with Parfit's famous Mere Addition Paradox from Chapter 19 of *Reasons and Persons.*

6. The Mere Addition Paradox

Consider the three outcomes in *Figure 2.4*:

A   A+   B
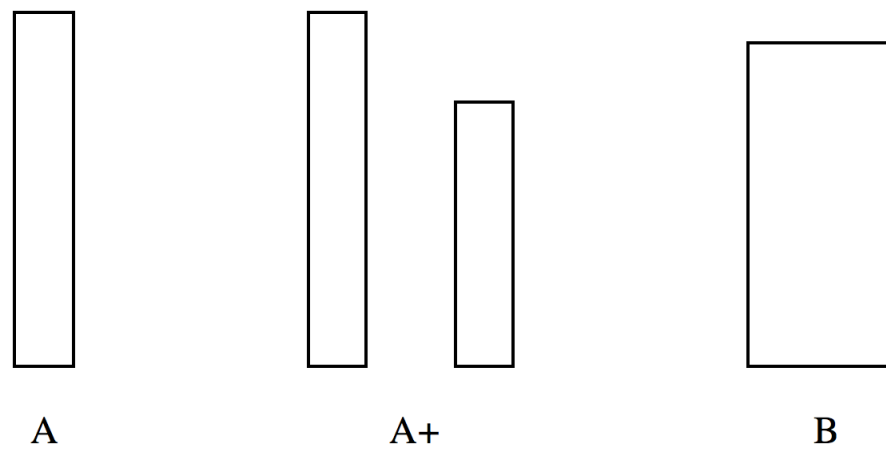
Figure 2.4

As in *Figures 2.2* and *2.3*, the width of each block shows the number of people living in an outcome and the height shows their quality of life. A contains 10 billion people, all with very worthwhile lives. A+ consists of two groups. The first group consists of the

same people as A, with the same high quality of life. A+ also contains a group of extra people, equal in number to those in A. These extra people are significantly worse off than the people in A, but still have lives that are well worth living. Outcome B contains the same two groups as A+. Everyone in outcome B is worse off than the people in A, but better off than the worse-off group in A+. Assume that the average quality of life in A+ (and hence also the total wellbeing in A+) is below that in B.

In *Reasons and Persons*, Parfit very convincingly argues that each of the following three pairwise comparisons is extremely plausible:

(1) A+ is worse than B.
(2) B is worse than A.
(3) A+ is not worse than A.

Let us briefly review the case for each of these three claims.

The intuition that A+ is a worse outcome than B is for many people the most robust. It is also supported by a number of distinct moral views: Total utilitarians, average utilitarians, telic egalitarians, and prioritarians all converge on the claim that B is a better outcome than A+.

The second claim, that B is worse than A, also strikes many people as strongly intuitive. *Everyone* who exists in B is worse off than everyone who exists in A (including the people who exist in both A and B). This verdict would be contradicted by total utilitarianism, which holds that any loss in the quality of lives in a population can be made up for by a sufficient gain in the quantity of a population. But as we have seen,

total utilitarianism itself is suspect, since it leads us to the Repugnant Conclusion via the Spectrum Argument.

Finally, consider the comparison between A and A+. Relative to A, A+ involves what Parfit calls "mere addition", meaning that (i) the existence of the extra people in A+ does not affect anyone else's well-being, (ii) they have lives that are worth living, and (iii) their existence does not involve injustice. Merely adding the extra people could be thought to make A+ worse than A, either by lowering average wellbeing or by introducing distributive inequality. Is it plausible that A+ is worse than A for either of these reasons? Parfit thinks not. The objection from average well-being is quickly dispatched with, since average utilitarianism has unacceptable implications in other cases, as we saw above. As to the objection from equality, Parfit responds by distinguishing two kinds of case: If we can avoid inequality by making some existing people better off, then avoiding inequality can make the outcome better. But if inequality can only be avoided if some people, who have lives worth living, do not exist, then avoiding inequality does not make the outcome better. Thus, Parfit claims, when inequality is produced by mere addition, it does not make the outcome worse: "We cannot plausibly claim that the extra people should never have existed *merely because (…) there are other people who are even better off.*"[67]

---

[67] *Reasons and Persons*, p. 425.

If we accept both the Internal Aspects View of Outcome Goodness and Completeness, however, these three intuitive judgments form an inconsistent triad. For instance, if A+ is not worse than A, and B is better than A+, this implies that B is *better*, not worse, than A. This contradicts (2). To avoid intransitivity, one or more of the judgments must be given up. Moreover, suppose it is (2) which we give up. This would again raise the threat of the Repugnant Conclusion. For the logic of the Mere Addition case can be iterated: Having established that B is better than A, we could then imagine a yet bigger population B+, which is produced from B through mere addition, and a population C, containing the same people as B+, but without inequality and with higher total wellbeing. If B > A, the same line of reasoning would suggest that C > B. This manoeuver can be repeated, until we arrive at outcome Z. By the transitivity of the "better than"-relation, we would be forced to accept the conclusion that Z is better than A, without making any of the totalist assumptions that drove the Spectrum Argument.

Rejecting Completeness and appealing to different number-based imprecision would allow us to avoid the specter of the Repugnant Conclusion. If different number-based imprecision holds, then (3), the claim that A+ is *not worse* than A does not imply that A+ must either be precisely equally good or better than A. Rather, A and A+ could be *on a par* in terms of goodness. The pairwise comparison of A+ and B is not affected by different number-based imprecision, since these are outcomes in which the same number of people exist. We can thus retain (1), the claim that A+ is worse than B. By contrast, (2), the claim that B is worse than A, must be given up: If A+ is imprecisely

equally as good as A, and B is better than A+, then B cannot be worse than A. Moreover, it is only by maintaining, in line with number-based imprecision, that B is on a par with A that the slide to the Repugnant Conclusion can be stopped: iterating the reasoning of the Mere Addition case, we may conclude that B is on a par with A (and hence not worse than A), that C is on a par with B (and hence not worse than B), etc., all the way to outcome Z. But since "on a par with" is not a transitive relation, we are not forced to the conclusion that outcome Z is on a par with, and hence not worse than, A.[68] Rather, we are free to maintain that outcome Z is worse than A.

A proponent of the Evaluative Asymmetry, however, should not rest content with this solution. As we just saw, appealing to different number-based imprecision helps avoid the Repugnant Conclusion, but only by rejecting the judgment that B is worse than A. This is a high price to pay, I think. It is not just that, intuitively, B does seem to many people *clearly* worse than A. There is also a simple argument from the Evaluative Asymmetry that leads to this conclusion. As illustrated by *Figure 2.5*, the differences between A and B can be decomposed into two distinct effects:

---

[68] This is not the Repugnant Conclusion strictly speaking, but a slightly weaker claim. The Repugnant Conclusion holds that outcome Z is *better* than A. By contrast, according to this weaker claim, Z is *not worse* than A. (Call this the *Weak Repugnant Conclusion*). However, though it is *logically* weaker than the Repugnant Conclusion itself, the Weak Repugnant Conclusion strikes me as just as unacceptable.
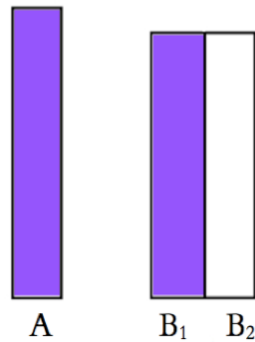
Figure 2.5

First, there is a change in the well-being of the people who exist in both A and B (highlighted in purple in *Figure 2.5*). This is the move from A to B_1. B_1 is incontrovertibly worse than A. Second, there is the addition of a group of new people, B_2, to B_1. However, if the Evaluative Asymmetry is right, adding new lives worth living does not make the world go better. Thus, B, the union of B_1 and B_2, is not better than B_1; and B_1, we said, is worse than A. Thus, all things considered, B is worse than A. The appeal to different number-based imprecision, by contrast, implies the negation of this claim.[69]

---

[69] This argument is similar to John Broome's "greediness objection" to number-based imprecision in "Should We Value Population?", *The Journal of Political Philosophy*, Vol. 13, No. 4 (2005), pp. 399-413, and elsewhere. In unpublished work I have argued that, given Broome's other commitments in population ethics, he is, in fact, ill placed to offer this argument. The greediness objection, it turns out, cuts as much against Broome's own position as against different number based-imprecision. Unfortunately, laying out my critique of Broome's argument would take us too far afield in the present context.

7. Taking Stock

We began by noting that the four following claims together have unacceptable implications:

(A): The second conjunct of the Evaluative Procreation Asymmetry, according to which creating a happy life rather than creating no life does not make the world go better, all else equal.

(B): The Evaluative Same Number Claim, according to which creating a very happy life makes the world go better, all else equal, than creating a moderately happy life.

(C): The Internal Aspects View of Outcome Goodness, which holds, roughly, that the goodness of an outcome is solely a function of its intrinsic properties, and does not depend on contextual factors, such as which outcome it is being compared to or the other alternatives in the option set.

(D): The Completeness assumption, which maintains that for any two outcomes which we compare in terms of their all-things-considered goodness, one must be better than the other, or they must be equally good.

Over the past three sections, I have explored the implications of rejecting (A), (B), and (D). None lead to a wholly acceptable resolution. Rejecting claim (B), the Evaluative Same Number Claim, is intrinsically very implausible. I therefore rejected this option out of hand.

Jettisoning claim (A), the second conjunct of the Evaluative Procreation Asymmetry also is an unsatisfactory way of relieving the tension between (A)-(D), given the aims of this chapter. After all, I am attempting to *uphold* both the Evaluative Procreation

Asymmetry and the Evaluative Same Number Claim. In addition, dropping the Evaluative Asymmetry in favor of a totalist moral view threatened to lead us either to the Repugnant Conclusion (via Parfit's Spectrum Argument) or Arrhenius's Sadistic Conclusion, both of which seem intuitively unacceptable.

So far, challenging (D), the Completeness Assumption, in favor of different number-based imprecision came closest to providing a satisfactory way of accounting for both the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim. But even this proposal had serious problems. For one thing, the assumption of *different number-based* imprecision seemed hard to justify, and therefore ad hoc. For another, different number-based imprecision had implications in Parfit's Mere Addition Paradox Case (namely that outcome B is on a par with outcome A), which are both counter to intuition, and reintroduce a tension with the Evaluative Asymmetry.

Having critiqued the main existing alternatives, it is time to present my own constructive proposal. My central gambit will be to question an assumption that all positions surveyed so far have held in common, namely that facts about the *goodness* of outcomes are either free-standing, or else *prior* to facts about *reasons* for action. Drawing on the work of T.M. Scanlon and A.C. Ewing, I will sketch and defend a view, what I call the "biconditional buck-passing view of outcome betterness", that reverses this relationship.

As we shall see, if the biconditional buck-passing view is correct, this allows me to reintroduce the notion of imprecise comparability, but on a sounder basis, and without

some of the implausible implications of *different number-based* imprecision. Armed with this notion, I am then in a position to show how the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim can be harmoniously reconciled.

The biconditional buck-passing view also provides a principled basis for challenging the Internal Aspects View of Outcome Goodness, in particular the Independence of Irrelevant Alternatives Principle, which it implies. This, in turn, puts me in a position to provide a solution to Parfit's Mere Addition Paradox. I will argue that there is a sense in which *all three* pairwise judgments in this case are correct. Yet, unlike Larry Temkin, who has also argued in favor of responding to the Mere Addition Paradox by rejecting the Internal Aspects View of Outcome Goodness, I do not think that my proposal threatens the transitivity of the "all-things-considered better than"-relation in any serious sense.

8. The Biconditional Buck-Passing View

In Chapter 2 of *What We Owe to Each Other*, T. M. Scanlon proposes what he calls his "buck-passing" account of goodness and value. According to Scanlon, the evaluative concepts *value* and *goodness* are analyzable in terms of the normative concept of *reason*. Contra "buck-stopping" theorists of goodness like G.E. Moore, Scanlon argues that for something to be good is not for it to have the simple – and, for Moore, unanalyzable – property of 'goodness'. Rather, it is to possess other properties that constitute reasons to respond to it in certain ways (for instance, to bring it about, to prefer it, to preserve it, to

respect or to cherish it, to admire it, to welcome it, etc.). As he writes:

> [B]eing good, or valuable, is not a property that itself provides a reason to respond to a thing in certain ways. Rather, to be good or valuable is to have other properties that constitute such reasons. Since the claim that some property constitutes a reason is a normative claim, this account also takes goodness and value to be non-natural properties, namely the purely formal, higher-order properties of having some lower-order properties that provide reasons of the relevant kind.[70]

Scanlon argues that his buck-passing view is supported by two considerations. The first is that, when we consider the reasons that we have to bring about, prefer, preserve, or admire things that are good or valuable, we realize that appealing to a special reason-giving property of goodness seems *redundant* once we have taken account of all of the thing's other properties. For example, our reasons to admire a good painting, to preserve it, and to prefer it to other less good works, seem exhaustively characterized by listing the various properties that *make* it a good painting: the vivaciousness of the colors, the subtlety of the facial expressions, the technical mastery of the artist, etc. Appealing, in addition, to a special reason-giving property of the painting's 'goodness' seems to add nothing. As Scanlon writes:

> It is not clear what further work could be done by special reason-providing properties of goodness and value, and even less clear how these properties could provide reasons for action.[71]

---

[70] T.M. Scanlon, *What We Owe to Each Other* (Cambridge, MA: Belknap Press, 1998), p. 97.

[71] Ibid., p. 97.

Secondly, the buck–stopper's view of goodness as a simple, reason–giving property sits ill with the sheer diversity of things that we call "good", and how little they appear to have in common:

> [M]any different things can be said to be good or to be valuable, and the grounds for these judgments vary widely. There does not seem to be a single, reason-giving property that is common to all these cases.[72]

Thus, Scanlon concludes, for some thing or outcome to be good is not for it to possess a simple property of goodness which provides us with reasons to undertake certain actions, or to have certain attitudes, with regard to it. Rather, it is simply the higher–order supervenient property of having lower–order properties which constitute the reasons in question.

In one of the first published discussions of Scanlon's buck-passing view, Jonathan Dancy criticized it for threatening to prematurely resolve the debate between consequentialism and deontology in favor of consequentialism. He writes:

> Deontologists have suggested in one way or another that there are duties, and so reasons, that are not value-involving. An action can be one's duty even though doing it has no value and its being done generates nothing of value. (…) The buck-passing view rules this out in advance. To have value is to have reason-giving features, we are told, and since this is an identity statement it goes both ways. So to have reason-giving features is to be of value. So the deontological view (…) is ruled out in advance of any significant debate.[73]

---

[72] Ibid., p. 97-8.

[73] Jonathan Dancy, "Should We Pass the Buck?", *Royal Institute of Philosophy Supplement* 47 (2000), p. 168.

Dancy's critique of the buck-passing view depends on reading the buck-passer as offering a *biconditional* account, according to which it is the case, not just that claims about value and goodness are ultimately analyzable in terms of reasons, but likewise that "to have reason-giving features is to be of value". As an interpretation of Scanlon's view, however, this biconditional interpretation appears inaccurate, and Dancy's criticism therefore wide of the mark. Scanlon never explicitly affirms the biconditional claim. Moreover, Roger Crisp reports that in conversation Scanlon asserted that while, on his buck-passing account, for something to be good or valuable just is for there to be reasons to respond to it in certain ways, it may be the case that there are true claims about reasons which do not entail claims about value.[74]

However, other philosophers, such as A. C. Ewing, who defended a view very similar to Scanlon's buck-passing account, explicitly endorsed the biconditional analysis.[75] According to Ewing's account in *The Meaning of Good*, a thing's being good or valuable means that it possesses properties such that it is the appropriate object of a pro-

---

[74] Roger Crisp, "Value, Reasons, and the Structure of Justification: How to Avoid Passing the Buck", *Analysis* 65 (2005), p. 83.

[75] I am indebted to Jonas Olson's "Buck-Passing and the Consequentialism/Deontology Distinction" in D. Egonsson, J. Josfesson, B. Petersson, & T. Rönnow-Rasmussen (eds.) *Hommage a Wlodek: Philosophical Essays in Honour of Wlodek Rabinowicz* (Lund University, 2007) for bringing to my attention the close parallels between Scanlon's view and Ewing's. My later remarks on whether adopting a *bi-conditional* buck-passing view prematurely resolves the debate between consequentialists and non-consequentialists are also indebted to Olson's discussion.

attitude. A 'pro-attitude', for Ewing, is "any favorable attitude to something (…) for instance, choice, desire, liking, pursuit, approval, admiration".[76] In a later text, Ewing writes that "[t]o say that some particular thing is intrinsically good is then to say that its nature in itself provides a reason for adopting a favorable attitude towards it".[77]

Ewing explicitly embraced the biconditional interpretation of his buck-passing thesis, because he thought it furnished him with a forceful response to a common consequentialist critique of deontological morality. According to this critique, deontological views have the puzzling implication that it is often right to deliberately bring about an outcome that makes the world go worse. Ewing clearly felt the sting of this challenge. He writes: "it is hard to believe that it could ever be a duty deliberately to produce less good when we could produce more."[78] Adopting a biconditional buck-passing view promises to provide an elegant solution to this problem. If the biconditional account is correct, Ewing argued, the outcome that we have most reason to bring about all things considered is, for that reason, *better* than other outcomes we could bring about. Thus, doing the right thing (what we have most reason to do, all things considered) and producing the best outcome could never come apart, since the latter analytically follows from the former. For instance, if deontological morality tells us that,

[76] A.C. Ewing, *The Definition of Good* (London: MacMillan, 1947), p. 149.

[77] A.C. Ewing, *Second Thoughts in Moral Philosophy* (London: Routledge and Kegan Paul, 1959), p. 63.

[78] The Definition of Good, p. 189.

all things considered we have most reason *not* to cut up the one innocent patient to save five others from dying, then abiding by these reasons is not only right, it also produces the best outcome. For what it means for an outcome to be best just is for it to be the outcome that we have most reason to bring about.[79] Ewing could thus capture the thought that it is never wrong to do what makes the makes the world go best. But the explanation suggested by his biconditional buck-passing view inverts the order of dependence assumed by most consequentialists: "[for the consequentialist] what we ought to do is derivable from the good, while the reverse is true if [the buck-passing] analysis is correct."[80]

I myself confess to being strongly attracted to this kind of biconditional buck-passing view. Let me add a consideration of my own in its support. At the outset of this chapter, I endorsed Christine Korsgaard's remark that the *function* of our concept "good" is "to mark out schematically the solutions to certain kinds of problems which we have to solve". In the course of presenting his buck-passing view, Scanlon expresses a similar thought:

> Judgments about what is good or valuable generally express practical conclusions about what would, at least under the right conditions, be reasons for acting or responding in a certain way.[81]

---

[79] Note that here, as everywhere else in this chapter, "outcome" is not equivalent to "end-state", but rather *includes* the action that brought about the end-state.

[80] Ibid., p. 188.

[81] *What We Owe to Each Other*, p. 98.

Let us apply this thought specifically to questions about the goodness of outcomes or ways the world can go. The paradigmatic practical question with regard to outcomes is: 'which outcome do I have reason to bring about, in a given situation?'.[82] Hence, if the function of our concept of "good" is to capture the answer to this practical question, then, in resolving the *normative* question which of two outcomes O_1 and O_2 I have more reason to bring about, I *ipso facto* answer the evaluative question about the comparative goodness of these two outcomes. Suppose that, in the knowledge of all relevant facts about O_1 and O_2 as well as all available alternatives, I have established that I have all-things-considered reason to bring about O_1 rather than O_2. How could it then *not* be the case that the world goes better, all things considered, if I bring about O_1? By assumption, in deciding which outcome I have more reason to bring about, I have duly taken into account all relevant reason-giving facts pertaining to the comparative preferability of O_1 and O_2, and have arrived at the conclusion that I have all-things-considered reason to choose O_1. On what *basis* could we then deny that the world goes better when I bring about O_1? Moreover, if the true normative theory holds that "deontological" considerations, such as rights, duties, claims, as well as the quality of my action itself,

[82] Of course, there are other questions too, such as the question which outcome to hope for, to welcome, fear, or be sad about, but answers to these questions tend, I believe, to largely travel together with answers to the former question.

bear on which outcome I have reason to bring about, then responding appropriately to these considerations (by choosing the outcome I have more reason to bring about, all things considered) is part of what makes the world go better.

I thus endorse the following

> *Biconditional buck-passing view of outcome betterness* ("the biconditional buck-passing view" for short):
> In a given choice situation, an available outcome O_1 is all-things-considered better than an available outcome O_2 just in case, and because, given full information about O_1 and O_2, I would have all-things-considered reason to bring about O_1 rather than O_2.[83]

This biconditional buck-passing claim is much narrower than Scanlon's or Ewing's account in a number of respects. Unlike these philosophers, I do not aim to provide an account of value or goodness quite generally (including, for instance, the value of things or persons). Rather, my biconditional buck-passing view merely aims to provide an account of evaluative claims about the comparative goodness of outcomes in a choice situation.

Let us return to Jonathan Dancy's criticism above. Is it true, as Dancy suggests, that a biconditional buck-passing view like mine automatically gives the game away to the consequentialist? Far from it, I think.

Dancy's criticism presupposes that any theory which posits a systematic connection

---

[83] Note that the biconditional buck-passing view supports the simple Bridge Principle of Section 2.

between doing what is morally right (i.e. what one has most moral reason to do) and making the world go best must, for that reason, be consequentialist in nature. But this strikes me as neither the only, nor indeed the most illuminating way of capturing the essence of the consequentialist outlook.

Indeed, as Jonas Olson observes, on perhaps the best-known way of distinguishing consequentialist and non-consequentialist ethical theories, namely John Rawls's, it would seem that, if anything, the biconditional buck-passing view biases things in favor of *non*-consequentialism.[84] For Rawls, the defining mark of consequentialist theories is that "the good is defined independently of the right, and then the right is defined as that which maximizes the good."[85] A non-consequentialist theory is one that denies this, i.e. a theory that "either does not specify the good independently from the right, or does not interpret the right as maximizing the good."[86] Since, according to my biconditional buck-passing view, facts about the goodness of outcomes are not independent, but rather a function of normative facts about reasons, the buck-passing view may be thought to hand the victory to non-consequentialism by default.

But I think there is a third way of drawing the distinction between consequentialism and non-consequentialism, which is superior to both Dancy's and Rawls's. The most

---

[84] See Olson, "Buck-Passing and the Consequentialism/Deontology Distinction".

[85] John Rawls, *A Theory of Justice* (rev. edn), (Cambridge, Mass.: Harvard University Press, 1999), p. 21.

[86] Ibid., p. 26.

fundamental difference between the consequentialist and the non-consequentialist perspective in moral philosophy concerns the question: "*How* do we have reason to respond to things that are valuable or good, according to this ethical theory?" Consequentialists are *monists* in this regard. The only kinds of reasons for responding to things that are valuable or good that the consequentialist recognizes are *teleological*: they are state-regarding reasons to maximally *promote* the occurrence of things that contribute to the value of the world.[87] Non-consequentialists, by contrast are *pluralists* about the *kinds* of ways in which we have reason to respond to what is of value. Thus, non-consequentialist theories prescribe not only promotion in response to that which is good or valuable, but also responses such as cherishing, honoring, protecting, respecting, admiring, etc.[88]

If this is a plausible way of drawing the distinction between consequentialist and non-consequentialist ethical views, then a buck-passing view (biconditional or not) can accommodate them both, and hence remain agnostic between them. As Scanlon himself notes:

> One could accept such [a buck-passing view of goodness and value] while still holding a purely teleological conception of value, since nothing in the argument just given rules out the possibility that the reasons associated with something's being valuable are all reasons to promote it,

---

[87] For a characterization of consequentialism in these terms, see Philip Pettit, "Consequentialism" in Peter Singer (ed.) *Blackwell's Companion to Ethics* (Oxford: Blackwell, 1992), pp. 230–240.

[88] These remarks are indebted to Scanlon's *What We Owe to Each Other*, chapter 2.

or perhaps to promote states of affairs in which it figures in various ways. My rejection of the latter view is based on the consideration of [other] examples (…) in which being valuable involves there being reasons to act or to respond in a wider variety of ways.[89]

But likewise, the biconditional buck-passing view does, of course, make room for ethical theories according to which our moral reasons for action are not purely teleological in nature. This includes, for instance, the bearer-dependent account of the morality of procreation that I offered in Chapter 1.

My defense of the biconditional buck-passing view in this section has, perforce, been all too brief for such a major claim. A more thorough discussion would examine, amongst other things, how a proponent of the view can respond to "wrong types of reasons"- objections, as well as make room for agent-relative reasons without incurring implausible commitments about the goodness of outcomes.

But I believe I have done enough by way of motivating this view to make it interesting and worthwhile to see how, if the biconditional buck-passing view is true, it provides us with a simple and even elegant way of reconciling the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim, as well as with the elements of a solution to Parfit's Mere Addition Paradox. It is to this task that I now turn.

---

[89] *What We Owe to Each Other*, p. 98.

9. From the Normative to the Evaluative Asymmetry: the Solution

If we accept the biconditional buck-passing view of outcome betterness sketched in Section 8, then a straightforward solution to reconciling the Evaluative Procreation Asymmetry and the Evaluative Same Number Claim comes within reach. I can now *derive* these claims from the corresponding normative claims, which I established in Chapter 1.

According to the first conjunct of the Normative Procreation Asymmetry, if a future person would foreseeably have a life that is not worth living, this in itself gives us a strong moral reason to refrain from bringing this person into existence. In other words, in a choice between bringing about O_{ML}, the outcome in which a miserable life is created, and O_{NL}, the outcome in which no life is created, I have strong moral reason to bring about the latter rather than the former outcome, all else equal. Together with my biconditional buck-passing view, the Normative Procreation Asymmetry entails the first conjunct of the Evaluative Procreation Asymmetry: Bringing about O_{NL} rather than O_{ML} makes the world go better, all else equal.

The second conjunct of the Normative Procreation Asymmetry implies that, in a *two-way choice* between creating a very happy new life and creating no new life, I have no moral reason, all else equal, to bring about the latter outcome rather than the former. (By "two-way choice", I mean that these are the only two outcomes in the option set). Together with the contrapositive of the biconditional buck-passing view, this implies

106

that the former outcome is not better than the latter. If Completeness holds, this would mean that O_{VHL} is either precisely equally as good as O_{NL} or worse than O_{NL}. But, as we saw in Section 5, we need not accept Completeness. There is the conceptual space to claim that O_{NL} and O_{VHL} might be 'on a par' in goodness.

And indeed, I think there are good reasons to make this claim. Let us say that if, of two outcomes O_1 and O_2, I neither have reason to bring about O_1 rather than O_2, nor reason to bring about O_2 rather than O_1, I am *rationally indifferent* between these two outcomes. However, rational indifference comes in two varieties: *fragile* and *robust*. Imagine a somewhat improved version of outcome O_2 – call it O_2+ – such that in a choice between O_2 and O_2+, I have reason to choose O_2+. If it is also the case that, given a two-way choice between O_1 and O_2+, I have reason to choose O_2+, then my rational indifference between O_1 and O_2 was fragile. By contrast, if, given a two-way choice between O_1 and O_2+ I remain rationally indifferent, then my rational indifference between O_1 and O_2 was robust. The natural evaluative counterpart of fragile indifference is *precise* equality of goodness. If two items are precisely equally good, then a slight improvement in one of the items will mean that they are no longer precisely equally good. By contrast, the evaluative counterpart of robust indifference is imprecise equality or parity. If two items are on a par then slightly improving one of the items does not imply that the improved item can no longer be on a par with the other item. Now, the views on the morality of procreation that I defended in Chapter 1 clearly imply that, all else equal, I am robustly rationally indifferent between O_{NL} and O_{VHL}. On the one hand I

argued that, *however* good a potential future person's life could foreseeably be, this fact does not, in itself, give me a reason to bring this person into existence. Thus, even *greatly* improving outcome O_{VHL} would not give me a reason to choose this improved version over O_{NL}. But, on the other hand I, also argued that, if the best life that I can create, either now or in the future, is a *moderately* happy life, then there is no reason not to create this new life. Thus, given a two-way choice between O_{NL} and a significantly worse version of O_{VHL}, I could remain rationally indifferent as well. For these reasons, I affirm the evaluative claim that O_{NL} and O_{VHL} are on a par in a two-way choice. By parity of reasoning, O_{NL} is also on a par with O_{MHL} in a two-way choice.

The Evaluative Same Number Claim can also be inferred from its normative counterpart. The Normative same number claim is a claim about a three-way choice: given that I can bring about O_{NL}, O_{MHL}, or O_{VHL}, I have contrastive reason to bring about O_{VHL} rather than O_{MHL}, if I choose one of these two options. Together with the biconditional buck-passing view, this implies that, given a three-way choice between O_{NL}, O_{MHL}, or O_{VHL}, the world goes better if I choose O_{VHL} rather than O_{MHL}.

What about the comparison between O_{NL} and O_{MHL}, and O_{NL} and O_{VHL} in this three-way choice? Must it be the case that, because O_{MHL} is a worse outcome than O_{VHL}, which in turn is not a better outcome than O_{NL} (by the Evaluative Asymmetry), O_{MHL} makes the world go worse than O_{NL}? That would seem a rather stark conclusion. Admittedly, creating a moderately happy life when I could instead have created a very happy life exposes me to justified criticism – whereas creating no life, when I could have created a

very happy life, does not. But to say that creating a moderately happy life under these circumstances makes the world go *worse* than creating no new life strikes me as too strong. Fortunately, I am not forced to bite this bullet. The appeal to imprecise equality allows me to maintain that O_{NL} is on a par in goodness with *both* O_{VHL} and O_{MHL}. But since imprecise inequality is not transitive, this does not stop me from affirming that O_{MHL} is worse than O_{VHL}, in line with the Evaluative Same Number Claim.

On the surface, this solution to reconciling the Evaluative Asymmetry and the Evaluative Same Number Claim may look just like the one that I considered in Section 5, when discussing the appeal to different number-based imprecision. But, in two crucial respects, it rests on very different foundations: Firstly, the *basis* for departing from Completeness and affirming imprecise comparability is entirely different. According to the claim of different number-based imprecision, evaluative comparisons between populations of different sizes are necessarily imprecise. Upon closer inspection, however, this claim lacked a clear justification. It also had implications (for the addition of lives that are not worth living, or the comparison between outcomes A and B in the Mere Addition Paradox) that, as a champion of the Evaluative Asymmetry, I am keen to avoid. By contrast, my present claims that O_{NL} is imprecisely equal in value to O_{MHL} and O_{VHL} rest on associated *normative* claims about our rational indifference between these options. The implications of my view for other cases are thus quite different from those of different number-based imprecision. (For instance, as we shall see in the next section, unlike different number-based imprecision, my view does not give us a reason to think

109

that options A and B in the Mere Addition Paradox are on a par in value).

Secondly, though nothing I have said in this section would in itself require us to jettison the Internal Aspects View of Outcome Goodness (– as I noted in Section 5, rejecting Completeness and embracing parity is still compatible with the IAVOG), we shall see in sections 10 and 11 that the biconditional buck-passing view, on which my defense of the Evaluative Procreation Asymmetry rests, does provide a principled basis for doing so. This will be crucial to my solution of the Mere Addition Paradox in the following section.

10. Solving the Mere Addition Paradox

The Mere Addition Paradox, I claim, trades on an ambiguity about the context of choice. There is a sense in which *all three* intuitive judgments about the Mere Addition Paradox are true, namely that

> (1) A+ is worse than B;
> (2) B is worse than A;
> (3) A+ is not worse than A.

Yet, understood in a different way, claim (3) is false. Let me explain.

The sense in which each of (1), (2) and (3) are correct is as pairwise comparisons in a *two-way* choice, i.e. in a choice situation where the option set contains only *these two* possible outcomes. Thus, it is true that

> (1*) A+ is worse than B, when the option set is Q = {A+, B};
> (2*) B is worse than A, when the option set is R = {A, B};
> (3*) A+ is not worse than A, when the option set is T = {A, A+}.

110

Each of these three claims can be given a straightforward justification on the bearer-dependent theory of wellbeing that I defended in Chapter 1, paired with the biconditional buck-passing view. In Section 6 we have already seen what this justification would look like in the case of (2*): There are bearer-dependent reasons to choose A over B (B is worse than A for the people who exist in both outcomes) and no bearer-dependent reasons to choose B over A: there is no moral reason to prefer an outcome in which the new people in B exist. Thus, all things considered, there is reason to bring about A rather than B in a two-way choice. This, plus the biconditional buck-passing view implies that B is worse than A.

Consider next (3*): The new people in A+ are significantly worse off than the people in outcome A (call them the "A-people"), as well as the A-people in outcome A+. Still, their lives are, by assumption, well worth living. And crucially, this is the best their lives could be going, *given that A and A+ are the only available options.* Under these circumstances, according to the bearer-dependent account, though we have no moral reason to create these new lives, we have no reason not to create them either. We can therefore be rationally indifferent between A and A+. This, plus the biconditional buck-passing view, implies that A+ is not worse than A.

Finally, consider (1*). A+ and B contain the same people. Total and average utility are higher in B. Moreover, the worst-off are better-off in B than in A+, and wellbeing is more equally distributed. All things considered, it is hard to deny that we have reason to

111

choose B rather than A+, and hence that A+ is worse than B, given the biconditional buck-passing view.

The air of paradox, I think, arises from the assumption that the pairwise judgments (1)-(3), which I have argued are true in *two-way* choices, i.e. as (1*)-(3*), carry over to a *three-way* choice, where all three outcomes are available. It would then be true that

> (1′) A+ is worse than B, when the option set is S = {A, A+, B};
> (2′) B is worse than A, when the option set is S = {A, A+, B};
> (3′) A+ is not worse than A, when the option set is S = {A, A+, B}.

This would spell big trouble, since (1′), (2′), and (3′) would together constitute an instance of what I call "intransitivity within an option set" as opposed to the "intransitivity across different option sets" that we witness in propositions (1*), (2*), and (3*).

From the perspective of practical reason, intransitivity *within* an option set would be much the worse problem, since it could imply contradictory instructions and paralyze choice. This is clearest in cases where the outcomes within an option set are ranked in a *cyclical* fashion, such that, for each outcome that we could bring about, there is a better outcome that we could bring about. Translating this back into a claim about reasons via the biconditional buck-passing view, it would be the case that there is no outcome that we could choose such that we wouldn't have reason to choose a different outcome: a classic rational dilemma. But even in the case above, which is not cyclical, intransitivity within an option would imply contradictory instructions: (1) suggests that I have more reason to bring about B than A+, whereas (2) and (3) together suggest the opposite. We

therefore have strong reason to deny that the outcomes within an option set can, in an intelligible sense, be intransitively ordered.

By contrast, intransitivity *across* option sets seems essentially innocuous from the point of view of practical reason. Since, by assumption, we are not talking about the same context of choice in (1*), (2*), and (3*), this combination of claims cannot imply contradictory instructions or lead to rational dilemmas.

Fortunately, as I shall now argue, it isn't the case that the pairwise judgments (1), (2), and (3), which are true in two-way choices, carry over to a three-way choice. For (3'), I will show, is false. (My views here diverge significantly from those of Larry Temkin, whose claims about the Mere Addition Paradox I briefly contrast with my own in an appendix to this chapter).

Here is how I think we should reason about in a three-way choice between A, A+, and B. There are two questions to be asked:

> Q1: *If* we bring about an outcome in which the new people exist, do we have reason to bring about A+ or B?
> Q2: Given the answer to Q1, do we have reason to bring about an outcome in which the new people exist rather than outcome A, in which only the A-people exist?

The answer to question Q1 is clear. In a choice between A+ and B, B seems the only defensible option, for all the reasons surveyed above. So the answer to Q1 is: *if* we bring about an outcome in which the new people exist, we have reason to bring about B rather than A. Translated into evaluative terms via the biconditional buck-passing view, this means that A+ is worse than B if the option set contains A, A+, and B. This is proposition

113

(1′) above.

Now consider Q2: Do we have reason to bring about an outcome in which the new people exist, or rather outcome A? We have just established that, *if* we bring about an outcome in which the new people exist, we would have to choose outcome B. Given this, Q2 reduces to a choice between outcomes A and B. Here, I can appeal to the argument I have already given for the two-way choice between A and B: There are bearer-dependent reasons to choose A over B (B is worse than A for the people who exist in both outcomes) and no bearer-dependent reasons to choose B over A. Thus, all things considered, there is reason to bring about A rather than B. This, plus the biconditional buck-passing view implies that B is worse than A, when the option set contains A, A+, and B. This is proposition (2′) above.

Given propositions (1′) and (2′), it would follow from the transitivity of the "better than"-relation that A+ is worse than A when the option set contains A, A+, and B. Thus, proposition (3′) is false. And indeed, this claim seems independently plausible: The lives of the new people in A+ go significantly worse than the lives of the A-people. Moreover, unlike in a two-way choice between A and A+, it can no longer be said to the new people that, though they are significantly less well off than the A-people, their lives are going as well as they could. Given that outcome B is also available in this scenario, this is no longer true. Unlike a case where the option set contains only A and A+, there now exists a way of removing the inequality between the A-people and the new people other than by preventing the new people from coming into existence. We could also remove the

114

inequality by making the new people better off, while reducing the wellbeing of the A-people by a lesser degree. Doing so would also increase overall wellbeing. For all these reasons, I believe that there would be a complaint of *injustice* against A+ in the three-way choice that didn't apply in a two-way choice between A and A+. By contrast, if we bring about outcome A, no-one has such a complaint. These considerations, I think, make it very plausible that while A+ is not worse than A in a two-way choice, A+ *is* worse than A in a three-way choice.

If these arguments are sound, I have shown how we can make sense of each of our pairwise intuitions in the Mere Addition Paradox, without incurring the cost of intransitivity within an option set. Interpreted as judgments about two-way choices, each pairwise judgment is plausible, and supported by the bearer-dependent theory of wellbeing that I have defended in Chapter 1. By contrast, I have shown that when we consider the Mere Addition Case as a three-way choice, we have sound moral reason to reject the claim that A+ is not worse than A. We then obtain a transitive ranking of the three options, according to which A > B > A+. This solves the Mere Addition Paradox.[90]

---

[90] At one point in Chapter 19 of *Reasons and Persons*, Parfit briefly considers a strategy for dealing with the Mere Addition Paradox that might be thought to resemble the solution I have proposed in this section. He writes:

> It may (…) be said: 'Suppose that these outcomes [in the Mere Addition Paradox] were the predictable effects of different possible acts. If we ask what we ought to do, we solve the Paradox. Assume that we could bring about either A, or A+, or B. It would be wrong to bring about A+. This would be wrong since there is a better outcome, B, that we could have brought about. But it would also be wrong to bring about B, since there is a better outcome: A.' (p. 429)

11. Against the Independence of Irrelevant Alternatives Principle

The solution to the Mere Addition Paradox that I have just presented implicitly rejects

the Internal Aspects View of Outcome Goodness. Let me make this explicit. As we saw

in Section 3, one of the implications of the IAVOG is that the following principle holds:

> *The Independence of Irrelevant Alternatives Principle (IIAP)*: If outcome O_1 is
> better than outcome O_2 when the option set is $S = \{O_1, O_2, \ldots, O_n\}$, then O_1
> must still be better than O_2 when the option set is contracted to T, where
> T is a *subset* of S that contains O_1 and O_2. (*Contraction consistency*)
> Likewise, if O_1 is better than O_2 when the option set is T, then O_1 must be

---

In response to this proposal, Parfit argues that attempting to resolve the Mere Addition Paradox by appeal to what we ought to do is not a solution to the paradox but "merely ignores it". He says:

> Most of our moral thinking may be about what we ought to do. But we also have views
> about the relative goodness and badness of different outcomes. As I have said, these are
> not views about moral goodness or badness, in the sense that applies to acts or to agents.
> If an earthquake kills thousands, this is not morally bad in this sense. But it is bad in a
> sense that has moral relevance. Our views about the relative goodness of different
> outcomes sometimes depend upon our views about what we ought to do. But such
> dependence often goes the other way. (…) [S]ome of our beliefs about what we ought to
> do depend upon our beliefs about the relative goodness of outcomes. Since these latter
> beliefs form the basis of some of our morality, we cannot refuse to consider an argument
> that is about these beliefs. (p. 429)

Do these remarks cut against my proposed solution? It might appear that they do, since my solution, by means of the biconditional buck-passing view, also makes appeal to normative facts about which of the outcomes in the Mere Addition Paradox we would have more reason to bring about, under certain conditions. But, in fact, I do not believe that Parfit's objection applies to my proposal. I have offered a theory of outcome goodness according to which facts about the comparative goodness of outcomes are downstream from, and true in virtue of facts about which outcomes we would have more reason to bring about, under certain conditions. Therefore, for me to give my view about which of the three possible outcomes in the Mere Addition Paradox we have most reason to bring about is *not* to ignore the question about the comparative goodness of these outcomes; it is to answer it, given my assumptions. Of course, one may wish to challenge one of the stepping-stones of my solution, namely the biconditional buck-passing view of outcome betterness. But that would be to question the content of my proposal, not its *relevance* to the problem at hand.

better than O₂ when the option set is expanded to S. (*Expansion consistency*)

By contrast, my claims about the Mere Addition Paradox violate the IIAP. I claim that A is better than A+ when the option set is S = {A, A+, B}, whereas A is *not* better than A+ when the option set is T = {A, A+}. This contradicts contraction consistency. Should I be troubled by this?[91]

Some philosophers have held the IIAP to be self-evidently correct. One of these is Derek Parfit himself, who writes in *Reasons and Persons*: "The relative goodness of two outcomes cannot depend on whether a third outcome, that will never happen, might have happened."[92] Indeed, the IIAP *would* be unassailable if we bought into other aspects of the Intrinsic Aspects View of Outcome Goodness, such as the thought that there is some fixed amount of impersonal goodness inherent in any outcome, which is solely a function of that outcome's intrinsic properties, and which is prior to, or at least

---

[91] It is sometimes suggested, for instance by Michael Neumann in "Choosing and Describing: Sen and the Irrelevance of Independence Alternatives" [sic], *Theory and Decision* 63 (2007), pp. 79-94, that putative counterexamples to the Independence of Irrelevant Alternatives Principle can be dealt with by individuating the relevant outcomes or alternatives in a more fine-grained manner. Whatever the merits of this proposal in dealing with other counterexamples to the IIAP may be, it cannot protect the IIAP against the threat posed by the Mere Addition Paradox. Given the conclusions I reached in the previous section, the only way in which we could uphold the IIAP in the Mere Addition Paradox would be by individuating outcomes in accordance with the option sets that they are a part of. That is, we would say that 'A+ -when-the-option-set-is- T = {A, A+}' is a *different outcome* from 'A+ -when-the-option-set-is- S = {A, A+, B}' and likewise for outcome A. This move would make it *trivially* true that the comparative goodness of two outcomes could not vary depending on the presence or absence of a third outcome in the option set. For any change in the composition of the option set would mean that we are talking about different outcomes. But of course such a move would entirely void the IPPA of its content.

[92] Reasons and Persons, p. 429.

independent of, our reasons for bringing about that outcome. In that case, the context of choice, meaning the presence or absence of third outcomes, could never effect the relative goodness of two outcomes.

However, the whole thrust of my argument in the second half of this chapter has been to reject this picture. As I argued in the previous section, there is good reason to think that some aspects of an outcome that affect its overall goodness, such as whether the outcome contains injustice or justified complaints, are not just a function of that outcome's intrinsic properties. Rather, they are *essentially context-dependent*, meaning that their presence or absence is a function of what other outcomes we could bring about in that situation. Furthermore, if my biconditional buck-passing view about outcome betterness is correct, facts about the relative goodness of two outcomes are often *downstream* from facts about which of the two outcomes we would have more reason to bring about, under certain conditions.

If the biconditional buck-passing view about outcome goodness is true, this would certainly put immense pressure on the IIAP. For there is no shortage of cases in which it is quite plausible that which of two outcomes O_1 and O_2 we have more *reason* to bring about depends on whether a third outcome is in the option set or not. I will give just one brief example.

Consider the following scenario. A runaway trolley is headed towards five people trapped on the main track. Let the following be possible outcomes associated with this scenario:

M: Let the trolley kill the five on the main track.

P: Throw Peter, who is standing on a bridge overlooking the tracks, into the path of the trolley, paralyzing Peter but saving the five.

R: Redirect the trolley away from the five and up on to the bridge towards Peter, killing Peter but saving the five.[93]

Suppose, first, that only M and P are in the option set. Many non-consequentialists think that it would be morally wrong to throw Peter, a mere bystander, onto the tracks in this case, even though doing so would "only" paralyze Peter and save five lives. Peter, it might be thought, has a right not to be used in this way, even in order to prevent a greater harm to five others. Given this belief, the outcome you have most moral reason to bring about in this case is M: letting the trolley kill the five.

Suppose now that our option set is expanded to include all three of M, P, and R. This may trigger what Frances Kamm calls the "Principle of Secondary Permissibility". Many people think that in a choice between M and R, it would be permissible to choose R: redirect the trolley away from the five and towards Peter, killing him. Suppose that this is what you *would* in fact do, were option P not available. Kamm argues that, under these circumstances, option P: throwing Peter onto the tracks, may become "secondarily permissible". For P is less harmful to Peter than R, the outcome you would otherwise

---

[93] This case is owed to Frances Kamm, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (Oxford University Press, 2007), p. 26.

permissibly bring about.[94] Since P is also the outcome that minimizes overall harm, this is arguably the outcome that you have most reason to bring about when the option set includes M, P, and R.

If this is right, it illustrates how the addition of a third outcome, R (which moreover isn't chosen in either version of our scenario), can affect which of outcomes M and P you have more reason to bring about. When the option set consists of only M and P, you have more reason to bring about M. By contrast, when the option set consists of M, P, and R, you have more reason to bring about P. If, in addition, the biconditional buck-passing view is correct, these normative claims entail the following evaluative claims: when the option set consists of only M and P, outcome M is better than outcome P; by contrast, when the option set consists of M, P, and R, outcome P is better than M.

Moreover, as I said, cases like the above, in which facts about which of two outcomes we have more reason to bring about intuitively depend on the availability or absence of third outcomes, are relatively plentiful in commonsense morality.[95] Therefore, conditional on also embracing the biconditional buck-passing view of outcome betterness, violations of the IIAP begins to look not like a bug, but like a feature of many plausible ethical views.

---

[94] For a fuller discussion, see *Intricate Ethics*, chapter 1.

[95] For two interesting recent discussions, see Alex Friedman, "Intransitive Ethics", *Journal of Moral Philosophy* 6 (2009), pp. 277-297 and Tim Willenken, "Deontic Cycling and the Structure of Commonsense Morality", *Ethics* 122.3 (2012), pp. 545-561.

12. Conclusion

In this chapter, I attempted to connect my defense of the Normative Procreation Asymmetry in Chapter 1 to corresponding evaluative claims about the goodness of the state of affairs produced by procreative decisions. After surveying a range of alternatives, none of which prove satisfactory, I proposed a view, the biconditional buck-passing view of outcome betterness, according to which facts about the comparative goodness of outcomes are a function of our reasons for bringing about one outcome rather than another. This allowed me to derive the Evaluative Procreation Asymmetry from the corresponding normative claims that I established in Chapter 1. The biconditional buck-passing view also provided me with a principled basis for challenging the Internal Aspects View of Outcome Goodness, and in particular the Independence of Irrelevant Alternatives Principle, which it implies. This, in turn, put me in a position to provide a novel solution to Derek Parfit's Mere Addition Paradox.

Appendix to Chapter 2: Temkin's Views on the Mere Addition Paradox

I have argued that we can solve the Mere Addition Paradox by jettisoning the Internal Aspects View of Outcome Goodness, which underpins the Independence of Irrelevant Alternatives Principle. But doing so, I claimed, does not, in itself, give rise to a serious worry about the transitivity of the "all-things-considered better than"-relation. We have reason, I argued, to distinguish between two different *types* of intransitivity: intransitivity *within* an option set and intransitivity *across* option sets. I suggested that our intuitive pairwise judgments about the Mere Addition Paradox – that B is worse than A, that A+ is worse than B, but that A+ is not worse than A – are correctly interpreted as true judgments about *two-way* choices, i.e. as judgments about cases where the two outcomes being compared are the only items in the option set. This set of pairwise judgments is intransitive, but only in the sense of intransitivity *across* option sets. Such intransitivity, I maintained in Section 10, is essentially innocuous from the point of view of practical reason, since an ordering that is merely intransitive across option sets cannot imply contradictory instructions or lead to rational dilemmas. Indeed, if the considerations against the Independence of Irrelevant Alternatives Principle in Section 11 are sound, many plausible accounts of the comparative betterness of outcomes will be intransitive across option sets. By contrast, I claimed that there is no intransitivity *within* an option set in the Mere Addition Paradox, which would have paralyzed choice and given rise to rational dilemmas. In a three-way choice, the outcomes are transitively ordered: A > B > A+. Since only intransitivity within an option

set is deeply problematic from the point of view of practical reason, I concluded that the Mere Addition Paradox does not raise a serious challenge to the thought that the "all-things-considered better than"-relation must necessarily be transitive. For, understood correctly, this desideratum only expresses a constraint on the ranking of outcomes *within* an option set, not across option sets.

On my account, there is intransitivity across, but not within, option sets in the Mere Addition Paradox, because one factor that affects the goodness of A+, namely whether A+ contains injustice, is *essentially context-dependent*. This means that whether or not A+ contains injustice is a function of *what other outcomes are in the option set*. Given the presence of B in the option set, the inequality in A+ constitutes an injustice. Whereas, if only A and A+ are in the option set, there is no injustice in outcome A+. This creates intransitivity across option sets, since A+ is worse than B when the option set is Q = {A+, B}, and B is worse than A, when the option set is R = {A, B}, but A+ is *not* worse than A when the option set is T = {A, A+}. However, in an option set that contains all three outcomes, the injustice in A+ matters both in comparisons with A+ *and with A*. That is, A+ is worse than B, on account of its injustice. *But it is also worse than A*, and for the same reason: A+ contains avoidable injustice, whereas A does not. Hence, on my analysis, there is no intransitivity within option sets in a three-way choice between A, A+, and B.

In his seminal discussions of the Mere Addition Paradox in "Intransitivity and the

Mere Addition Paradox"[96] and in *Rethinking the Good*[97], Larry Temkin makes a number of

remarks which reveal an importantly different take on this problem. For Temkin, the

Mere Addition Paradox *does* raise a deep challenge to the transitivity of the "all-things-

considered better than"-relation. Temkin does not claim that our intuitions about the

Mere Addition Paradox *compel* us to give up on the transitivity of "all-things-considered

better than". We may instead choose to give up one of our considered judgments about

the Mere Addition Paradox and preserve the transitivity of the better than-relation.[98]

What Temkin maintains is that *if* we do embrace certain deeply plausible ways of

thinking about the Mere Addition Paradox (amongst them Parfit's claims in *Reasons and

Persons* about when inequality makes an outcome worse and when it doesn't, which we

surveyed in Sections 6 and 10 above[99]), then this would force us to accept that the "all-

---

[96] Larry Temkin, "Intransitivity and the Mere Addition Paradox", *Philosophy & Public Affairs* 16.2 (1987), pp. 138-187.

[97] Larry Temkin, *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Oxford University Press, 2012).

[98] As Temkin points out in chapter 11 of *Rethinking the Good*, there are ways of thinking about the Mere Addition Paradox which allow us to retain the Internal Aspects View of Outcome Goodness, and which therefore raise no challenge to the transitivity of the "all-things-considered better than"-relation. But such views will imply that at least one of our intuitive pairwise judgments about the Mere Addition Paradox is false.

[99] Other moral views which, according to Temkin, would also lead to judgments about the Mere Addition that are deeply intransitive are certain forms of Maximin, as well as the Narrow Person-Affecting view. See *Rethinking the Good*, Chapter 12.

things-considered better than"-relation is indeed intransitive in a deep sense. Although Temkin himself does not explicitly draw a distinction between 'intransitivity across option sets' and 'intransitivity within an option set', it seems clear from his discussion that he believes the Mere Addition Paradox challenges the transitivity of "all-things-considered better than", not just across option sets, but *within* an option set as well. This is a much more radical conclusion than the one I drew. Fortunately I believe that it is not warranted; indeed, I will show that it is not borne out by the logic of the arguments which Temkin himself adduces.[100]

According to Temkin, what may underlie our intransitive pairwise intuitions in the Mere Addition Paradox, is the fact that certain values, such as equality or Maximin, may be what he calls "essentially pairwise comparative". For instance, Temkin thinks that on the view suggested by Parfit's claims about inequality in Chapter 18 of *Reasons and Persons*, it is true that "*[e]quality is [c]omparative*, not merely in the ordinary sense – that it involves judgments about how some fare relative to others – but in the sense that our judgment about a situation's inequality depends on the alternative it is being compared

---

[100] In *Rethinking the Good*, Temkin presents many other examples – from outside the domain of population ethics – that, he believes, also show that the transitivity of "all-things-considered better than" is in deep tension with some of our considered judgments. (See, in particular, the "spectrum of ordeals"-cases from Chapter 5 of this book). Although I find these cases fascinating and suggestive, I lack the space to discuss them here. All I aim to establish in this appendix is that Temkin is wrong to think that the *Mere Addition Paradox*, in particular, raises a deep worry about the transitivity of "all-things-considered better than".

to."[101] On this view about equality[102], Temkin writes, "the *relevant and significant* factors for comparing A and A+ regarding inequality differ from those for comparing A+ and B in a sense connected with inequality being *essentially pairwise comparative*."[103] Thus, "[t]here is no fact of the matter as to how bad the inequality in A+ *really* is considered just by itself. How bad it is depends upon the alternative compared to it. Compared to B, A+ is bad; compared to A, it isn't."[104]

The view suggested in these remarks is that how A+ is to be judged in terms of equality is essentially *a function of the alternative it is being compared to* – not, as I suggested, of the option set as a whole. If this were indeed the case, then how A+ compares to A in terms of inequality would *not* differ depending on whether or not B was also in the option set or not. Rather, in *both* a two-way and a three-way choice, A+ would not be worse than A in terms of equality. In that case, however, our intransitive judgments in two-way choices *would* carry over to a three-way choice between A, A+, and B, thus raising the specter of intransitivity *within* an option set. "All-things-

---

[101] "Intransitivity and the Mere Addition Paradox", p. 147.

[102] Note that this is not Temkin's own view about equality. On Temkin's own view, equality is *not* essentially pairwise comparative. Indeed, the view about equality that Temkin defends in *Inequality* (New York: Oxford University Press, 1993) is compatible with the IAVOG.

[103] "Intransitivity and the Mere Addition Paradox", p. 149.

[104] Ibid., p. 150.

considered better than" would indeed be intransitive in a *deep* sense, if our intuitive judgments about the Mere Addition Paradox were sound.

Fortunately, I think that Temkin is mistaken that, on the view implied by Parfit's argument, facts about how A+ is to be judged in terms of equality are solely a function of the alternative it is being compared to, as opposed to the option set as a whole. The logic of Parfit's argument does not, in fact, support this position. As I stressed in section 10, our basis for thinking that A+ is not worse than A in a two-way choice is that the inequality between the A-people and the new people could only be eliminated by preventing the new people from coming into existence. But crucially, this rationale no longer applies in a three-way choice, which includes option B. Here, there *is* a way of avoiding the inequality in A+ that makes the new people better off. Hence, not availing ourselves of this option and producing outcome A+ is injustice; and this, in turn, makes A+ worse than *both* outcomes A and B in the three-way choice.

If this is right then the comparative goodness of A and A+ is *not*, as Temkin maintains, essentially pairwise comparative. Rather, it is, as I said, *essentially context-dependent*. This, as we saw, supports the view that the correct judgments in the Mere Addition Paradox are intransitive across option sets in two-way choices. But in a three-way choice between A, A+, and B, we obtain a transitive ranking. Temkin's belief that the Mere Addition Paradox poses a *deep* challenge to the transitivity of the "all-things-

127

considered better than"-relation thus rests on drawing the wrong lesson from Parfit's argument.[105]

---

[105] Similar remarks apply to Temkin's discussion of "Maximin" and the Narrow Person-Affecting View. Consider the former. Arguing that "maximin" is essentially comparative, Temkin writes: "while usually it is morally regrettable if the worst-off fare worse in one situation than another, this is not always so, and, in particular, is not if it results from mere addition." ("Intransitivity and the Mere Addition Paradox", p. 151). Therefore, A+ is not worse than A in terms of Maximin, though it is worse than B in terms of Maximin. Again, however, these remarks seem to overlook the moral importance of the option set. If our concern is that the worst-off not fare worse than they otherwise *might*, this does not support the thought that Maximin could never judge a situation bad if it is produced by mere addition. What matters are the *available alternatives*. So, a move from A to A+, even though it is just 'mere addition' could be criticized on grounds of Maximin, provided outcome B was also available.

Chapter 3: Objections and Replies

1. Introduction

In this third and final chapter, I consider and rebut two of the most common kinds of objection to the Procreation Asymmetry. According to the first set of objections, affirming the Procreation Asymmetry, which holds that

> If a future person would foreseeably have a life that is not worth living, this in itself gives us a strong moral reason to refrain from bringing this person into existence.
> By contrast, there is no moral reason to create a person whose life would foreseeably be worth living, just *because* her life would be worth living.

commits us to

> *Anti-Natalism*: It is prima facie wrong to have children.

According to the second objection, affirming the Procreation Asymmetry prevents us from making sense of the very widespread intuition that we have weighty moral reasons to be concerned with the long-term survival of humanity and to work against its premature extinction.

I will discuss each of these objections in turn.

2. Intra-Personal Objections to Procreation

Arguments for Anti-Natalism can be subdivided into two broad categories, what I call intra-personal and whole-life objections to procreation. If affirming the Procreation

Asymmetry forced us to embrace either type of argument, or deprived us of the means to rebut it effectively, this would be a serious strike against the Asymmetry, since Anti-Natalism strikes most people as implausible. I begin by examining intra-personal objections to procreation.

2.1 The Simple Intra-Personal Argument

Whether or not a life is worth living for a person is a function of the various *personal goods* and *bads* it contains. A personal good is something that tends towards, and can contribute to, making a life worth living for a person; a personal 'bad' is something that tends towards, and can contribute to, making it not worth living for that person. What exactly one considers a personal good or bad depends on one's theory of wellbeing. Since my defense of the Asymmetry intuition in Chapters 1 and 2 has aimed to remain agnostic on this score, I will not here offer a substantive account of personal goods and bads.

According to what I call the *Simple Intra-Personal Argument*, affirming the Asymmetry, in combination with the empirical assumption that any life we could create will contain some personal bads, commits us straightforwardly to Anti-Natalism. The argument runs as follows:

> *The Simple Intra-Personal Argument*
> (1) Because there is a reason not to create a life that is not worth living, the fact that the life of a potential future person will foreseeably contain some personal bads is a prima facie moral consideration against creating that person, since personal bads tend towards making a life not worth living.

(2)　　Because there is no reason to create a life worth living just because it will be worth living, the fact that the life of a potential future person will foreseeably contain some personal goods, which tend towards making a life worth living, does not affect the balance of moral reasons for or against creating that life.

(3)　　Any life we can create will contain both personal goods and personal bads.

(4)　　Thus, any life we can create will contain some things (personal bads) which are prima facie moral considerations *against* creating that life and other things (personal goods) which do not affect the balance of reasons for or against creating that life.

(5)　　Therefore, for any life we can create, the balance of moral reasons will be *against* creating that life, all else equal.

Therefore, it is prima facie wrong to create any life by having children.

I argue, however, that the Simple Intra-Personal argument does not threaten the Procreation Asymmetry. The argument is fatally flawed, because premise (2) of the argument commits a fallacy of division. From the claim that there are no moral reasons to create a life that is overall worth living, it mistakenly infers that the presence of personal goods in a life, which tend towards making it a life worth living, cannot affect the balance of reasons for or against creating that life. But this is incorrect.

To see this more clearly, consider first the following analogy:

Tennis Pro: Should a professional tennis player play a public match against an amateur? Consider the following stylized facts:
(a) If the professional loses the match, this is extremely embarrassing for him. Losing the match would constitute a reason not to play.
(b) By contrast, there is no glory in winning the match. Everyone expects the professional to win. Therefore, winning the match would not, in itself, constitute a reason to play.
(c) It is foreseeable that, over the course of the match, the professional will lose some points and win some points.

From these stylized facts, the following argument might be constructed, on analogy with

the Simple Intra-Personal Argument:

> (1) Because there is a reason not to play a match that he loses, the fact that the professional will foreseeably lose some points is a prima facie consideration against playing the match, since losing points conduces towards losing the match.
>
> (2) Because there is no reason to play a match just because he wins, the fact that the professional will foreseeably win some points, which conduces towards winning the match, does not affect the balance of reasons for or against playing the match.
>
> (3) In any match that the professional plays against the amateur, it is foreseeable that he will both win and lose some points.
>
> (4) Thus, any match he plays against the amateur will contain some things (lost points) which are prima facie considerations *against* playing the match and other things (won points) which do not affect the balance of reasons for or against playing the match.
>
> (5) Therefore, for any match that the professional can play against the amateur, the balance of reasons will be *against* playing.

This argument is clearly fallacious. So long as there is no shame in losing individual

points but only the entire match, the mere fact that the professional is bound to lose *some*

points over the course of the match does not suffice to give him an all-things-considered

reason against playing. If the professional is sure to win enough points to win the match,

the fact that he will lose some points has no significance to what he has reason to do, all

things considered. While, by assumption, the fact that he will win the match does not, in

itself, give the professional a reason to play, there is also no reason for him *not* to play

the match just because it will contain some lost points. Hence, if the professional has

other reasons for playing the match – for instance, to raise money for charity – then this

is what he has reason to do, all things considered.

Something very similar can be said with regard to procreation: The Asymmetry maintains that we do not have moral reasons to create a life *just because* it would be worth living. Thus, the fact that a life would contain personal goods does not give us a reason *in favor* of creating the life. However, counting *in favor* of procreation is not the only way in which personal goods can affect the balance of reasons for or against creating a life. Personal goods matter, I claim, by *canceling* or *neutralizing* the moral force of personal bads (which give us prima facie moral reasons *against* creating that life). That is, if a potential person's life will contain sufficient personal goods to make her life worth living for her (by compensating her for the presence of whatever personal bads her life contains), this makes it the case that there is no moral reason, all things considered, not to create this person. Therefore, if someone has other reasons for wanting to have children, this is what she has reason to do, all things considered. (I consider some potential reasons in Section 3 below).

In "Asymmetries in the Morality of Causing People to Exist", McMahan considers the Simple Intra-Personal Argument and discusses precisely this way of responding to it. However, he expresses doubts. It seems to him "strikingly ad hoc" to maintain that in procreative contexts personal goods can cancel the force of personal bads but don't themselves constitute reasons in favor of procreating, whereas in deliberating about actions that will affect already existing persons (what McMahan calls "individual-affecting choices"), personal goods have both a canceling *and* a reason-giving function:

> Why should goods have both reason-giving and canceling functions in individual-affecting choices but only a canceling function in procreative

133

choices? Why should goods have one kind of impersonal value – canceling value – but not the other – reason-giving value? My impression is that no one has tried to answer these questions in part because no one has asked them. But having asked them, I confess that, at the moment at least, I do not know how to answer them.[106]

Given my discussion in Chapters 1 and 2, this is an objection to which I am well placed to respond. The above remarks about personal goods in procreative and individual-affecting choices would only be ad hoc if they were offered as *free-standing* claims, or as claims that are *explanatorily upstream* from the Asymmetry itself. But they are not, in the context of my discussion. Rather, they are *derivative* of my earlier defense of the Asymmetry. It is *because* of the fact that we have no moral reason to create a new life just because it would be worth living that it is also true that the presence of personal goods in this life (which conduce towards making it worth living) do not give us a reason to create it. If this is the correct order of explanation, then the above claims about the cancelling and reason-giving function of personal goods in procreative and individual-affecting choices piggy-back on my account of the Asymmetry in Chapter 1. And, of course, the whole thrust of the bearer-dependent account of well-being that I elaborated in that chapter is to explain the difference in our moral reason for conferring ordinary

[106] Jeff McMahan, "Asymmetries in the Morality of Causing People to Exist" in David Wasserman & Melinda Roberts (eds.), *Harming Future Persons* (Dodrecht: Springer, 2009), p. 54.

benefits on already existing people on the one hand, and benefiting people by bringing them into existence, on the other.

2.2 Shiffrin's Argument from Unconsented Harm

My response to the Simple Intra-Personal Argument relied on the thought that sufficient personal goods can compensate a person for the presence of some personal bads in her life. But in "Wrongful Life, Procreative Responsibility, and the Significance of Harm", Seana Shiffrin points out that "compensation", in this context, can mean two different things:

> (i) Personal goods (or benefits) could compensate me for personal bads (or harms) in the sense that I am all things considered *better off* for having received both these benefits and harms.
> (ii) Personal goods (or benefits) could compensate me for personal bads (or harms) in the sense that it is *morally unproblematic* to impose on me these harms without my consent, so long as I also receive these benefits.

Shiffrin argues that compensation in the first sense does not imply compensation in the second sense, and that this makes procreation morally problematic. Unfortunately, doing full justice to Shiffrin's subtle discussion would take us too far afield. Shiffrin's own argument is unconnected to the Asymmetry intuition. Its relevance to the dialectic of this chapter is that, if sound, it would block the response to the Simple Intra-Personal Argument that I gave in the previous section. I shall therefore limit myself to briefly laying out her argument and explaining why I find two of its main premises unpersuasive.

135

Shiffrin's argument, in a nutshell, runs as follows:

Shiffrin's Argument from Unconsented Harm: [107]

      (1)   Any human life contains some non-trivial harms (e.g. pain and illness).

      (2)   It is morally problematic, and often impermissible, to impose an unconsented nontrivial harm on a person in the course of procuring for her *pure benefits*, i.e. "benefits that are just goods and which are not also removals from or preventions of harm."[108]

      (3)   Although a person can be benefited, all things considered, by being brought into existence, a possible future person is not harmed by not being brought into existence.

      (4)   Hence, to create a child with a life worth living is to impose on her unconsented non-trivial harms in the course of procuring for her pure benefits. [from (1), (2), and (3)].

Therefore, it is at least morally problematic, to have children, even when the child is benefited all things considered by being brought into existence [from (2) and (4)].

One key claim of Shiffrin's argument is premise (2). The intended contrast to imposing unconsented harms in the course of procuring *pure benefits* is imposing unconsented harms on a person in the course of *removing or averting an even more serious harm* to her. The latter class of actions are often clearly permissible, according to Shiffrin. She gives the example of breaking an unconscious driver's arm, when this is the only way to rescue her from her car, which is about to explode. (Call this case *Rescue*).

---

[107] This reconstruction of Shiffrin's argument is indebted to David DeGrazia, *Creation Ethics: Reproduction, Genetics, and Quality of Life* (New York: Oxford University Press, 2012), chapter 5.

[108] Seana Shiffrin, Wrongful Life, Procreative Responsibility, and the Significance of Harm," *Legal Theory* 5 (1999), p. 124.

Contrast *Rescue* with a case that involves unconsented harming in the course of

conferring a pure benefit. Shiffrin's own example is colorful, and worth citing in full:

> *Gold Manna*: Imagine a well-off character (Wealthy) who lives on an island. He is anxious for a project (whether because of boredom, self-interest, benevolence, or some combination of these). He decides to bestow some of his wealth upon his neighbors from an adjacent island. His neighbors are comfortably off, with more than an ample stock of resources. Still, they would be (purely) benefitted by an influx of monetary wealth. Unfortunately, due to historical tensions between the islands' governments, Wealthy and his agents are not permitted to visit the neighboring island. They are also precluded (either by law or by physical circumstances) from communicating with the island's people. To implement his project, then, he crafts a hundred cubes of gold bullion, each worth $5 million. (The windy islands lack paper currency.) He flies his plane over the island and drops the cubes near passers-by. He takes care to avoid hitting people, but he knows there is an element of risk in his activity and that someone may get hurt. Everyone is a little stunned when this million-dollar manna lands at their feet. Most are delighted. One person (Unlucky), though, is hit by the falling cube. The impact breaks his arm. Had the cube missed him, it would have landed at someone else's feet. Unlucky may have his arm repaired for much less than $5 million and benefits from the extra cash. He admits that all-things-considered, he is better off for receiving the $5 million, despite the injury.[109]

Although Unlucky is all-things-considered benefited by Wealthy's actions, Shiffrin finds

the imposition of unconsented harm on Unlucky morally deeply disturbing. She

suspects that this is because the harm Unlucky undergoes is not, as it is for the rescued

[109] Ibid., pp. 127-8.

person, in the service of preventing his suffering a greater harm; rather, it is to bestow a pure benefit.

However, we may wonder whether Shiffrin's explanation truly captures the morally relevant difference between *Rescue* and *Gold Manna*. One salient feature of *Rescue*, which is absent from *Gold Manna*, is that in *Rescue* the trapped driver is *in principle* incapable of giving her consent to the rescuer's actions. Due to her loss of consciousness, she has, in a sense, temporarily ceased to be an agent and, at that very moment, has no attitude or quality of will towards her own rescue. This is not the case in *Gold Manna*. At the time of the gold drop, the islanders are fully capable agents. There is nothing stopping them from making up their own minds about whether they wish the risky drop to proceed or not. (Indeed, elsewhere in the text Shiffrin suggests that, if asked, Unlucky would probably have refused his consent – though, in retrospect, he is glad that the drop occurred). It is only the political tensions between the two islands that prevent Wealthy from consulting the islanders and obtaining their consent.

It seems to me quite plausible that harming a person without their consent, even when this benefits the person all-things-considered, is especially morally serious in cases where that person is *not* in principle incapable of giving her consent. It is in such cases that we risk *overriding* the person's agency – turning her, against her will, into a merely passive object of our own agency. This risk is absent in cases like *Rescue* where, by assumption, there is nothing to be overridden at the time of our action.

138

To examine whether this difference between *Rescue* and *Gold Manna* – and not that between conferring a pure benefit and averting a more serious harm – may be driving our intuitions about these two cases, consider the following case:

> *Long-distance Amputation*: You suffer from a medical condition that will cause you to lose your right leg if left untreated. For a complicated medical reason, the only available treatment involves amputating your left hand. Suppose that, objectively speaking, you would be better off losing your hand than your leg. The only doctor capable of carrying out the procedure lives in a distant country, and cannot communicate with you. However, he has a device that permits him to carry out the operation remotely. So, without your consent, he amputates your hand, thereby saving you from losing your leg.

Like *Rescue* and unlike *Gold Manna*, this case involves an unconsented harm in order to prevent a greater harm, not to confer a pure benefit. Yet the doctor's actions in *Long-distance Amputation* are clearly morally problematic. At the same time, like *Gold Manna* and unlike *Rescue*, this is a case where consent is not in principle impossible, and thus the beneficiary's will is potentially being overridden. This suggests that it may be the latter factor, not the former, which in fact accounts for the intuitive moral difference between *Rescue* and *Gold Manna.*

This matters to the success of Shiffrin's Argument from Unconsented Harm. If the relevant moral difference between *Rescue* and *Gold Manna* is that between preventing a greater harm and conferring a pure benefit, then, given premise (3) of her argument, procreation is more like *Gold Manna* than *Rescue*, and hence morally suspect. By contrast, if the morally relevant difference is the one that I have highlighted, procreation is closer to *Rescue* than to *Gold Manna*, and therefore morally innocuous in most cases. After all, if

there is one act to which we cannot *in principle* give our consent, it is the act of our own creation.

Suppose, however, that we grant Shiffrin premise (2) of her argument. Does the argument then go through? I believe not. Premise (4), the claim that in procreating we *impose* unconsented non-trivial harms on our offspring, also strikes me as unsound in most cases.

The claim that being brought into existence can impose harms on children would be less controversial if what Shiffrin meant by "harms imposed by procreation" were limited to things that are the direct causal consequence of procreation itself – congenital conditions such as spina bifida, for instance. The scope of Shiffrin's intended claim, however, is much broader. For Shiffrin, *all* harms that befall a person later in life count as potentially imposed on her by her creation.

Following David DeGrazia, I want to suggest that Shiffrin elides an important distinction: that between *imposing* harm on a child and *exposing* a child to potential harm by creating it[110]. To expose someone to harm is, roughly, to create the conditions whereby she may later be harmed – be it by the intervening agency of other persons, by her own actions, or by natural events. By contrast, for an action A to count as imposing

---

[110] DeGrazia acknowledges that he owes the distinction between imposing harm and exposing to harm to Frances Kamm. Kamm also discusses this distinction in her *Bioethical Prescriptions* (2013), pp. 319-320, footnote 5.

harm on some person, the harm must be a relatively direct causal consequence of A itself.

I agree with De Grazia that this difference between imposing harm and exposing to harm is morally significant. We routinely do things that would count as exposing children to harm (without their consent), yet which seem morally innocuous. For instance,

> Parents who enroll their children in school expose them to any manner of possible harms: humiliation in class, despair at being 'dumped' by friends, depression secondary to poor performance or social difficulties, emotional and physical bullying, minor harms during recess or gym class, sleep deprivation, illness contracted from other students, etc.[111]

Yet, given the opportunity for substantial benefit that going to school represents, our behavior seems faultless in most cases – perhaps even morally required.

Seen in this light, the fact that any human life will foreseeably contain its share of disappointments, of pain and sadness and loss, does not suffice to show that all parents impose serious harms on their children by creating them, nor that procreation is, in general morally suspect. If the benefits of being alive are large enough to make the child's life well worth living overall, bringing her into existence and thereby exposing her to the vicissitudes of life is morally innocuous.

---

[111]*Creation Ethics*, p. 153.

3. Whole-Life Objections to Procreation

The intra-personal objections to procreation that I examined in the previous section tried to derive an anti-natalist conclusion by focusing on goods or bads *within* any human life.

The second set of anti-natalist arguments take a different form. These arguments grant that creating an individual life that is *on balance* well worth living would not be morally problematic, *if we could be sure that this is what will come about*. The problem is that, given the current state of medical knowledge, any act of procreation involves some epistemic risk, however small, of producing a life that is not worth living. Furthermore, if enough people decide to procreate, then, due to the Law of Large Numbers, it is foreseeable that some lives not worth living will be created (along with many worthwhile lives). The question is whether we can justify our decision to procreate, given the downside risk of creating someone with a life not worth living. This is what I call a "whole-life" objection to procreation, since the argument takes as its units *entire lives*, some of which are not worth living, as opposed to the intrapersonal version of the argument which focuses on goods or bads *within* a life.

Though, as we shall see, whole-life objections to procreation do not rest on the Asymmetry itself, embracing the Asymmetry makes it more challenging to respond to these objections. If the Asymmetry is true, we might reason, then two of the following things could happen when I choose to procreate: Either, I am unlucky and create a life that is not worth living. This is an outcome I have reason to avoid. Or, I am lucky and create a life worth living. But even this is an outcome that I have no moral reason to

bring about, for its own sake. So, in the best case, my act produces an outcome that I have no moral reason to bring about, and in the worst case, it produces an outcome that I have strong moral reason to avoid. How could procreation be justifiable under these conditions?

The first step in responding to this whole-life objection is to realize that thinking of our reasons for procreation in purely moral terms is a drastic oversimplification. The motivations of most parents in procreating are not 'moral' in nature. We do not bring children into this world in order to add to the stock of happy lives that are ever lived, nor even to benefit our offspring in particular. Rather, most people who decide to have children do so because they feel that without them their lives would be lacking in crucial respects – from the joy and pride that children often bring their parents, to the manifold and complex ways in which our existence can be enriched and rendered more meaningful by founding a family. That these sentiments are, in an extended sense, 'self-regarding', does not render them any less profound or worthy of taking seriously in moral reflection. James Lenman puts it well:

> The desire to have children is a selfish sort of sentiment, to be sure, but in a peculiar and complicated way. Partly it is a matter of wanting there to be a constituency for that range of our moral and altruistic instincts that we bring to bear on our immediate successors. If there is no such constituency, our lives are impoverished in central and vital ways. The desire for – as the song has it – somebody to love is that peculiarly

143

sociable form of selfishness that is fundamental to human moral community.[112]

It is therefore wrong to assume that if there is no moral reason to create a new life just because that life would be worth living, there could be no valid reason *tout court* to procreate. Rather, the challenge posed by the whole-life objection is to explain how it could be permissible for parents to act on their desire to procreate, given the non-trivial risk of producing a child with a life that is not worth living.

Before we look more closely at this second type of anti-natalist argument, I want to introduce a distinction between two types of conditions that might make a person's life not worth living.

3.1 N-conditions and C-conditions

Suppose we have some theory about the *distinctive necessary properties* of each particular person, i.e. the properties that *this person* could not have lacked, and which only this person could have possessed. Consider, for instance, the *Origin View* considered by Parfit in *Reasons and Persons*:

> *The Origin View*: Each person has this distinctive necessary property: that of having grown from the particular pair of cells from which this person actually grew.[113]

---

[112] James Lenman, "On Becoming Extinct", *Pacific Philosophical Quarterly*, Vol. 83 (2002), p. 261.

[113] Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984), p. 352. As Parfit points out, this property "cannot be fully distinctive. Any pair of identical twins grew from such a pair of cells. And any

I propose to divide genetic conditions that make a life not worth living into two types: those that co-vary with a person's distinctive necessary properties (whatever they are) and those that do not. (I bracket the issue of lives that are made not worth living due to *environmental* or *social*, as opposed to *genetic* factors).

Conditions that do *not* co-vary with a person's distinctive necessary properties I shall call *N-conditions*. These are genetic conditions that make a person's life not worth living, but which *that* person might not have had. By contrast, *C-conditions* are conditions which, given the current state of medical knowledge, do co-vary with a person's distinctive necessary properties. In other words, it is not *practically* possible for this person to have existed *without* his C-condition. Another person could have existed instead, without the C-condition, but this person would have been a numerically distinct individual. Which actual genetic conditions will count as N-conditions and which as C-conditions will depend on empirical factors as well as on the account of the necessary distinctive properties of persons that we endorse.

Next, suppose, for the sake of argument, that there are no C-conditions, only N-conditions, but that every procreative act involves some small risk of creating a person

---

fertilized ovum might have later split, and produced twins." (ibid). The Origin view would have to be amended to meet these problems. But since it serves only as an illustration for a more general argument, I need not do so here.

with an N-condition. Would procreation be morally problematic under these assumptions?

Not necessarily, I believe. As I argued in Chapter 1, just as it harms a person (in the non-comparative sense) to be created with a miserable life, being created with a life worth living benefits a person. (What I denied is that the latter fact, in itself, gives us any moral reason to procreate). Now, if for any potential future person his *ex ante* probability of being born with an N-condition was sufficiently low, and his life without an N-condition would be sufficiently good for him, I believe that the risky act of procreation could be retrospectively justified to whoever ends up being created: Being created had a high likelihood of benefiting him and only a small chance of harming him. This is true even if the new person is unlucky and ends up being afflicted with an N-condition. Thus, even though we have no moral *reason* to create a new person just because his life would foreseeably be happy, the fact that, in this scenario, any potential new person has a good probability of having a happy life makes it morally *permissible* for parents to procreate.[114]

---

[114] For a much more detailed discussion of this type of "ex ante" justification of risk imposition, see my "Contractualism and Social Risk: How to Count the Numbers Without Aggregating" (ms).

3.2 C-conditions and the New Problem of Non-Identity

The second way in which a person might have a life not worth living is due to a C-condition, which co-varies with the person's numerical identity. There is no practical possibility in which *this* person exists and does not have the condition that makes his life miserable. There could have existed a person without the C-condition, but this would have been a numerically different individual. For a potential future person at risk of being born with a C-condition, the act of procreation thus represents what I shall call a *pure downside risk*, i.e. a risky action which, for *her*, has no potential upside, only a potential downside (although it does have a potential upside for someone else). "Heads I win, tails you lose".

The existence of C-conditions poses a much more serious ethical problem than that of N-conditions. Suppose, to take a stylized example, that there are 999 "happy sperm" for every 1 "wretched sperm". Someone might say, in defense of procreation, that the chance that you end up having a child with a life worth living is very good, namely 999/1000. However, crucially, this does not mean that, if we have a miserable child, we could say to that child: "there was a 999/1000 chance that *you* would have a life worth living". This is false, since, *ex hypothesi*, a child born from a wretched sperm could *only* have had a life not worth living, as a matter of practical possibility. Thus, procreation couldn't be defended to someone who is born with a C-condition as a gamble that was very likely to work out in his favor. It is true that we were very likely to produce *a*

person whose life was well worth living. But that person would have been non-identical with the wretched child. I call this, a little grandly, the *New Problem of Non-Identity*.

Of course, the differences between N-conditions and C-conditions matter only if we believe, as moral contractualists and other non-consequentialists do, that the justifiability of our actions to each potential individual – in terms of how our action might affect *them* – is ethically significant. If, by contrast, we think that what matters is only the expected overall outcome of our procreative act, as consequentialists assume, the difference between N-conditions and C-conditions is not morally relevant; what matters is the likelihood that a given procreative act will result in *a* person (de dicto) whose life is not worth living. Since I myself belong to the former camp, I will attempt to solve the New Problem of Non-Identity from a broadly contractualist perspective.

According to

> *Scanlon's Contractualist Formula*: An act is right or permissible just when such acts are licensed by a principle that no-one could reasonably reject.[115]

Scanlon also affirms the following

> *Greater Burden Principle*: "It would be unreasonable (…) to reject a principle because it imposed a burden on you when every alternative principle would impose much greater burdens on others."[116]

---

[115] This formulation of the contractualist formula is a composite of the many subtly different versions of his formula that Scanlon present throughout *What We Owe to Each Other*.

[116] T.M. Scanlon, "Contractualism and Utilitarianism", in A. Sen and B. Williams (eds.) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982), p. 111.

Since merely potential people are not burdened by not being brought into existence (there is no-one *to* be burdened), I believe that, if there is someone who could claim to be burdened by an anti-natalist moral principle prohibiting procreation, it is potential parents. By contrast, those who are burdened by an alternative moral principle permitting procreation are those unfortunate children who are born with C-conditions. How ought we to weigh these potential complaints against one another?

The New Problem of Non-Identity arises from the fact that procreation is a risky activity. In procreating, parents run the risk of imposing on their offspring life with a C-condition. Elsewhere[117], I have argued that in thinking about the morality of risk-imposition from a contractualist standpoint, it is often appropriate to discount a person's complaint against suffering a burden by their *likelihood* of suffering that burden. Might this thought allow the contractualist to avoid the New Problem of Non-Identity?

At first blush, it might seem that it should not. Imposing the risk of suffering a C-condition on a potential future person appears importantly different from ordinary cases of risk-imposition involving already existing people. In ordinary cases of risk-imposition concerning already existing people, the following is true:

> The probability that person X exists and is not burdened by the risky action A = 1 – the probability that X exists and is burdened by A.

[117] In "Contractualism and Social Risk".

Therefore, if the probability that X is burdened by A is low, the probability that X exists and isn't burdened is high.

But this is precisely what is not the case when dealing with C-conditions. Here, although it is true that Pr(X exists and X is burdened by the C-condition) is low, it is also the case that

Pr(X exists and X is not burdened by the C-condition) = 0.

The only life that an individual with a C-condition *can* have is a life in which he is burdened. In this connection, the fact that it was unlikely that this individual would exist and be burdened may appear to take on an entirely different significance than in ordinary cases of risk-imposition. After all, how could it make our action more justifiable to *this* unfortunate individual X that there was a high likelihood that, instead of him existing with a wretched life, some *other person* could have existed with a life worth living? That seems like cold comfort.

Fortunately, I believe that there is a way of solving the New Problem of Non-Identity within the contractualist framework. My argument proceeds in three steps. First, I believe that it is a mistake to view the New Problem of Non-Identity as fundamentally different from that of imposing pure downside risks on already existing people. The fact that in the New Problem of Non-Identity, unlike that of ordinary risk imposition, it is true that

Pr(X exists and X is not burdened by the C-condition) = 0

does not, on deeper reflection, have the moral significance that it at first appears to have. In both the New Problem of Non-Identity and in ordinary cases of pure downside risks the following is true: the only possible worlds in which those exposed to these risks are *affected* by the risky action A are those in which they are *harmed* by A. The fact that in ordinary cases of pure downside risks the victim (call her S) had a chance of *not* being burdened by the risky action A does nothing to *compensate* her for the risk imposed on her by A. It is not as if the fact that there are some possible worlds in which S is *not* burdened by A is a *benefit* provided to her *by action A*. Hence, the fact that if S is not negatively affected by A she will still *exist* in a case of *ordinary* downside risk, whereas she would not exist if the downside risk was a C-condition, seems neither here nor there in assessing the justifiability of A to S. If these arguments are right, the New Problem of Non-Identity is not fundamentally different from that of imposing pure downside risks on already existing people.

Second, despite taking seriously the "separateness of persons", which suggests that imposing harms or risks on one person cannot be straightforwardly compensated for by sufficiently large benefits to others, contractualists need not reject as impermissible all actions that impose pure downside risks on some. Scanlon himself suggests as much in *What We Owe to Each Other*:

> (…) there are many things that we do or depend on that involve risk of serious harm to others. Suppose, then, that we are considering a principle that allows projects to proceed, even though they involve risk of serious harm to some, provided that a certain level of care has been taken to

reduce these risks. It is obvious what the generic reason would be for rejecting such a principle from the standpoint of someone who is seriously injured despite the precautions that have been taken. On the other side, however, those who would benefit, directly or indirectly, may have good generic reason to object to a more stringent requirement. In meeting the level of care demanded by the principle, they might argue, thy have done enough to protect others from harm. Refusing to allow activities that meet this level of care would, they could claim, impose unacceptable constraint on their lives.[118]

Third, as I argued in my paper "Contractualism and Social Risk", actions that impose pure down-side risks, just like other instances of non-gratuitous risk-imposition, should be morally evaluated from an *ex ante* perspective. This means that *both the harms and benefits should be discounted by their improbability* in determining the strength of people's reasons for rejecting a principle licensing us to perform such actions.

Taken together, these three points allow us to see how a principle licensing procreation can satisfy Scanlon's Greater Burden Principle, despite the small risk of producing a child with a C-condition. Once we discount the reasons that a potential future person who bears the risk of suffering a C-condition has for rejecting a principle licensing procreation by the tiny likelihood of this person being created and suffering from his condition, and compare them to the certain and significant burden that outlawing procreation would place on parents, it is plausible that a principle licensing procreation passes the test of justifiability to each person.

---

[118] T.M. Scanlon, *What We Owe To Each Other* (Cambridge, MA: Belknap Press, 1998), pp. 236-7.

This concludes the first half of my discussion. If I am right, embracing the Procreation Asymmetry does not force us to embrace anti-natalism. Both intra-personal and whole-life objections to procreation can be successfully met without giving up the Asymmetry intuition.

I now turn to the second main objection, namely the thought that affirming the Procreation Asymmetry prevents us from making sense of the very widespread intuition that we have weighty moral reasons to be concerned with the long-term survival of humanity and to work against its premature extinction.

4. The Survival of Humanity

It is probably inevitable that humanity will someday go out of existence. The end could be sudden and cataclysmic – the result of an asteroid strike or global nuclear war. Or extinction might creep up on us gradually, through catastrophic climate change, resource depletion, or falling birthrates. What seems hard to doubt is that, one way or another, humanity's career will eventually come to an end. In the extremely unlikely event that mankind has managed to survive until then, the end of the Universe will finally do us in.

Most people accept that humanity's demise is only a matter of time. Yet, at the same time, many of us believe that it would be very bad, indeed one of the worst things that could possibly happen, if, for preventable reasons, the end came much sooner rather

than later. Insofar as this is within our power, we assign significant importance to ensuring that humanity survives for as long as possible.

These sentiments can be explained as arising, in part, from self-interest. Most obviously, there is the small but non-trivial possibility that *we* will be part of that last generation which is violently extinguished. No-one wants to belong to the generation that allowed humanity to go extinct – if only because this might coincide with our own premature demise.

We also have less narrowly egoistic reasons for wanting humanity to continue *beyond* our own deaths – for wanting there to be, in the words of Samuel Scheffler, an "afterlife".[119] Many of the activities and practices that give meaning to our own lives, Scheffler argues, are intimately connected with, and sustained by, our belief in the continued survival of humanity. In the absence of an afterlife, goal-oriented activities such as cancer research or social reform would lose much of their sense, since a large part of the intended benefit of these activities would vanish together with mankind itself. Likewise, activities designed to sustain certain values over time, such as traditions or cultural practices, would be threatened with futility, since the end of humanity would preclude the success of such efforts.

---

[119] Samuel Scheffler, *Death and the Afterlife* (New York: Oxford University Press, 2013).

Yet, although these two factors play an important role in our thinking about humanity's survival, they do not tell the full story, I believe. Even a Schefflerian interest in the afterlife, which transcends a mere concern for our *personal* survival, can be expected to drop off rather quickly a few generations after our deaths, as the things which Scheffler identifies as giving us reason to care about posterity – a "personalized" relation with the future or the survival of certain cherished values or traditions – fade away. By contrast, some of our present efforts to control existential risks to humanity (such as early detection programs for large earth-bound asteroids) involve extinction-events that would, in all likelihood, occur many centuries, if not thousands of years, from now. If indeed there are good reasons to undertake these efforts, they cannot primarily be grounded either in our collective interest in our own survival or in a Schefflerian concern with the afterlife.

Jonathan Glover reaches a similar conclusion by means of a thought-experiment. He has us imagine that we could take a drug that causes everyone currently alive to become infertile – thereby ensuring that the youngest generation alive will be the last – but at the same time makes us so blissfully happy as to more than compensate us for this fact. Even in this scenario where, by assumption, any *self-regarding* reasons to prefer the survival of humanity are outweighed, Glover thinks it would be wrong for us to let

humanity become extinct.[120] Like many others, Glover has the firm intuition that mankind has weighty *moral* reasons to ensure that it survives for as long as possible.

The final objection to the Procreation Asymmetry that I shall consider contends that affirming the Asymmetry leaves us unable to make sense of this powerful intuition.

It is true that by far the most common explanation found in the literature for why it matters that humanity survive for as long as possible does indeed depend on explicitly rejecting the Procreation Asymmetry. This is what I call the *Argument from Additional Lives*, which is favored by many consequentialist philosophers.[121] The argument proceeds as follows:

> *The Argument from Additional Lives*
> (1)   The world will go better, all else equal, if additional happy lives are lived than if they are not lived.
> (2)   The world will go worse, all else equal, if additional miserable lives are lived than if they are not lived.
> (3)   If humanity survives for as long as possible rather than going extinct sooner, the goodness of the additional happy lives that will be lived will outweigh the badness of the additional miserable lives that will also be lived.
> (4)   Hence, the world goes better, all things considered, if humanity survives for as long as possible rather than going extinct sooner. [from (1), (2), and (3)]
> (5)   We have moral reason to do what will make the world go better.

---

[120] Jonathan Glover, *Causing Death and Saving Lives* (Harmondsworth: Penguin Books, 1977), p. 69.

[121] Versions of this argument are found, inter alia, in Jonathan Glover's *Causing Death and Saving Lives*, in Gregory Kavka, "The Futurity Problem" in R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations* (Temple University Press, 1978), in Derek Parfit's *Reasons and Persons*, and John Broome, *Climate Matters: Ethics for a Warming World* (New York: Norton, 2012).

Therefore, we have moral reason to help humanity to survive for as long as possible rather than letting it go extinct sooner. [from (4) and (5)].

The factual assumption in premise (3) – that, if we prevent the premature extinction of humanity, this will maximize the net balance of happy over miserable lives that are ever lived – is non-trivial and might be contested. However, given likely advances in medicine, agriculture, and technology over the coming centuries, and in spite of serious challenges such as global warming, the gradual exhaustion of non-renewable resources, etc., it is a reasonable assumption to make. Let us grant it for the sake of argument.

The key normative premise of the Argument from Additional Lives is proposition (1), the assumption that the world will go better just in virtue of additional happy lives being lived. This, of course, contradicts the evaluative version of the Procreation Asymmetry, which I defended in Chapter 2. The question then is: supposing we hold fast to the Procreation Asymmetry and jettison the Argument from Additional Lives, is there still a convincing justification for our belief that we have weighty moral reasons to prevent the premature extinction of humanity?

In his provocative article "On Becoming Extinct", James Lenman argues that there is not. Given that it is inevitable that humanity will go extinct some day, Lenman contends that morally speaking it does not matter whether this occurs sooner or later.

Lenman, like myself, rejects the Argument from Additional lives. His reasoning is the following: If the central premise of the Argument from Additional lives is correct, namely that the world goes better, all else equal, if more happy lives are lived, this

157

would imply that we have moral reason to want, not just that humanity become *diachronically* more numerous, by surviving for as long as possible, but also that it be *synchronically* more numerous, by containing as many people as possible in any given generation (provided enough of them live happy lives). Indeed, since proponents of the Argument from Additional Lives hold that the value of a happy life does not depend on *when* it is lived[122], our moral reasons for wanting either kind of numerousness would be exactly equivalent[123]. But Lenman, like myself in Chapter 2, embraces the Evaluative Procreation Asymmetry, and therefore rejects the contention that the world would go significantly better if, at every point in time, humans (including those with happy lives) were synchronically more numerous. He therefore rejects the Argument from Additional Lives. Considering the analogy of an animal species, he argues:

> When we consider synchronically the size of the white rhino population, it is not clear that it matters how large that population is. If what matters is the instantiation of the universal – white rhino or whatever – that is already, as it were, taken care of. Of course, if there were fewer white

---

[122] See, for instance, John Broome, *Weighing Lives* (Oxford: Oxford University Press, 2004), pp. 126-128 and Glover, *Causing Death and Saving Lives*, pp. 69-71.

[123] Gregory Kavka is one philosopher who explicitly embraces this equivalence. He writes:

> Imagine God deciding between creating a universe with one planet occupied by *n* happy people, and a universe with two planets, *each* occupied by *n* people, each just as happy as those in the first universe. Does the fact that there are twice as many happy people in the latter universe constitute a reason for God preferring to create it? (...) Notice that our problem concerning future generations is quite analogous to this one, the difference being that the extra equally happy people are located in later generations rather than on another planet. Now I confess to being one of those who strongly feels it would be better for God to create the greater number of equally happy planets (or generations). ('The Futurity Problem', p. 196).

rhinos, it might be said that individuals of that species that might have existed will fail to exist and perhaps those individuals have intrinsic value. (…) But it is hard to make much sense of the thought that this is a bad thing – either for those individuals themselves or otherwise.[124]

If we cannot appeal to the value of additional happy lives being lived, are there other, specifically human traits, such as our sensitivity to moral reasons, sophisticated language use, etc. that could support the thought that we have weighty moral reasons to want humanity to go on for as long as possible? Lenman thinks not. Again, he argues, it is helpful to first consider the connected question: do these special traits of human beings give us reason to want humanity to be *synchronically* more numerous? If, as Lenman believes, the answer to this question is 'no', then what basis it there for thinking that these characteristics of human beings give us reason to want humanity to survive as long as possible, thereby becoming *diachronically* more numerous? For Lenman there is no good answer to this question. By contrast, in the following section I will present an argument, which, I believe, successfully responds to Lenman's challenge.

But let me first dispatch with an unsuccessful objection to Lenman's position. Some scholars have advanced other moral reasons to oppose the premature extinction of humanity which, unlike the Argument from Additional Lives, do not also imply a reason to make humanity synchronically more numerous. In particular, Gregory Kavka,

---

[124] "On Becoming Extinct", p. 256.

Jonathan Bennett, and Michael Tooley each endorse variants of the following proposal:

Human history, like individual human lives, has a certain narrative structure[125]. When we consider the prospect of humanity's extinction, we should think of it as an evil in the same way as we (including Lenman) think of the premature death of a person as an evil. Given that human history could potentially continue for millions of years to come, which would no doubt witness momentous developments in the cultural, scientific, political, and spiritual life of our species, an event that led to our premature extinction within the next few centuries or millennia would amount to tragically cutting short the human adventure in its infancy.[126]

Lenman's response to this proposal is convincing:

> (…) it is implausible to suppose that human history – or that of any species – has a natural narrative structure in the same way as a human life. (…) If we reject grand philosophical pictures that endow human history with some essential pattern, all that can be meant by metaphorical talk of our species' childhood is those times that are *relatively* early in its career whenever they may turn out to be. The individual human tragedy of dying young has no obvious analogue in the career of our species as a whole.[127]

---

[125] In the same article "On Becoming Extinct", Lenman himself argues that individual human lives have a narrative shape and that it is better, other things being equal, to prolong them to the full time span typical of our species than to allow them to end early.

[126] See Jonathan Bennett, "On Maximizing Happiness" in R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations* (Temple University Press, 1978), p. 65; Gregory Kavka, 'The Futurity Problem', pp. 196-198; Michael Tooley, *Abortion and Infanticide* (Oxford: Oxford University Press, 1985) pp. 257-58.

[127] "On Becoming Extinct", p. 259.

This seems exactly right. The flawed analogy between the early extinction of humanity and the premature death of a person rests on an overly anthropomorphized picture of humanity. But humanity is not a super-individual. It does not have a personal identity or interests, or projects the realization of which requires its continued survival. Humanity cannot enjoy life, and its only projects are those of its individual members. It therefore cannot be said to lose anything by being discontinued.

The upshot of Lenman's discussion is that, if we embrace the Procreation Asymmetry and reject the Argument from Additional Lives, there really is no good reason to think that it matters, *sub specie aeternitatis*, whether humanity's career comes to an end sooner or later. Of course, it rightly matters to *us* that humanity not go extinct within our lifetime, or within a few generations of our death. But given that humanity is bound to come to an end eventually, it is, alas, inevitable that *some* generation of human beings will suffer this tragic fate. It would be hubristic to suppose that it would be worse, impersonally speaking, if this misfortune befell us or our immediate descendants rather than a generation of humans further in the future.

5. The Argument from the Final Value of Humanity

If Lenman's position was the end of the story, this would be a highly counterintuive result of the Asymmetry – its *reductio ad absurdum* even.

Fortunately, we are not forced to choose between our intuitions about the morality of individual procreation and the survival of humanity. I will conclude this chapter by

sketching an argument, the Argument from the Final Value of Humanity, which can make sense of our conviction that it matters morally that humanity survives for as long as possible, without appealing to the thought that the world goes better, all else equal, the more happy lives are ever lived.

Let us say that something has *final* value if it is valuable for its own sake, independently of the instrumental value it may have for anyone. In many cases, something with final value will be *intrinsically* valuable, i.e. valuable in virtue of its internal or non-relational properties; but it need not be so.[128]

It is commonplace to claim of a wide range of things that they have final value in this sense: wonders of nature, great works of art, animal and plant species, languages, cultures, etc. The suggestion that *humanity* too, with its unique capacities for complex language use and rational thought, its sensitivity to moral reasons, its ability to produce and appreciate art, music and scientific knowledge, its sense of history, and so on, should be deemed to possess final value, therefore seems extremely plausible to me.[129] The Argument from the Final Value of Humanity proceeds by showing that there is a

---

[128] For more on the four-way distinction between intrinsic vs. extrinsic and instrumental vs. final value, see Christine Korsgaard, "Two Distinctions in Goodness" in her *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996).

[129] For a detailed defense of the claim that the human (and other) species have final value, see Holmes Rolston, "Duties to Endangered Species", *Bioscience* 35.11 (1985), pp. 718-726.

conceptual link between responding *appropriately* to the final value of humanity and being disposed to ensure its survival.

The thought is, indeed, a natural one. We undertake great efforts to ensure the survival of many other things we consider finally valuable. Millions are spent annually on the preservation of rare plant and animal species (even those of no identifiable scientific or aesthetic value to us), and there are countless organizations worldwide working to preserve dwindling cultures and languages. In undertaking these efforts, we need not believe that being a member of this particular culture, or speaking this particular language, is *better for anyone*, nor indeed, that it is better for anyone, all things considered, that there exist members of this culture or speakers of this language – though these things might also be true. Rather, it is held, the culture or language deserves to be preserved because they are *valuable in their own right*. The analogous claim that, given humanity's final value, we have moral reason to help it to survive therefore strikes me as extremely plausible.

Indeed, many philosophers have argued that caring for a thing's survival is, at least in normal circumstances, part of what it *means* to value that thing finally. Thus, Samuel Scheffler writes: "What would it mean to value things but, in general, to see no reason of any kind to sustain them or retain them or preserve them or extend them into the

future?".[130][131] G.A. Cohen, in an unpublished version of his essay "The Truth in Conservatism", makes the following argument: "A thing that has [final] value is worthy of being revered or cherished. We do not regard something as being worthy of being revered or cherished if we have no reason to regret its destruction, as such." This, Cohen thinks, is precisely what distinguishes valuing something finally from valuing it merely for its instrumental properties, which often goes hand in hand with treating the thing as fungible:

> One can say, quite properly, upon acquiring a valuable thing, 'I shall value this until something better comes along', but one cannot in the same way say 'I shall cherish this until something better comes along': that could happen to be a correct prediction, but it could not express a decision to cherish.

Here we have, then, a way of explaining why we ought to care about humanity's survival that is quite distinct from the Argument from Additional Lives. According to the Argument from the Final Value of Humanity, each successive generation collectively has a defeasible moral reason to work for the survival of humanity, since this is how we appropriately respond to its final value. (Such a moral reason must be defeasible, I

---

[130] Samuel Scheffler, "Immigration and the Significance of Culture", *Philosophy & Public Affairs* 35, No. 2 (2007), p. 106.

[131] So much at least seems true of things – like persons, objects, species, languages, etc. – which, given adequate care, *can* survive over extended periods of time. As Chris Korsgaard pointed out in conversation, the same concern for a thing's survival isn't necessarily part of responding appropriately to the final value of essentially *ephemeral* things, such as events or sensory experiences (e.g. a concert). The relevant contrast in the appropriate attitudes may be roughly that between "cherishing" (for things that can survive with adequate care) and "savoring" (for essentially ephemeral things).

believe, since we can imagine circumstances in which conditions on Earth – or on other planets – become so bleak as to make the lives of future persons not worth living. In such a situation, any moral reason to perpetuate humanity in virtue of its final value would, I think, be outweighed or silenced).

Notice that unlike the Argument from Additional Lives, the Argument from the Final Value of Humanity does not assign the same importance to humanity's becoming diachronically and synchronically more numerous, all else equal. What matters according to this argument is that humanity *survives*, i.e. that it persists into the future. The moral grounds it advances for ensuring the survival of mankind do not also constitute grounds for creating additional human lives in the *present*.[132]

We can further our understanding of the Argument from the Final Value of Humanity by responding to the objections that James Lenman raised against this line of reasoning in "On Becoming Extinct". As we saw, Lenman considers a view according to which the special traits of humanity, or of other species like the white rhinoceros, give us moral reasons to want to avoid their premature extinction. As he writes, it is "a natural thought (...) that human beings have [final] value, impersonally regarded. And that

---

[132] Note also that a reason to ensure the survival of *humanity* goes beyond merely ensuring that human beings exist in the future. To see why, imagine a world in which each generation of humans dies and vanishes without trace before the next one is born (imagine that each generation lays eggs before its death, but disappears before their offspring has hatched). Thus, each new generation lives without knowledge of previous generations of humans. The human *species* survives in this scenario, but a lot of what we mean by "humanity", and a lot of what seems to contribute to its final value – relationships between parents and children, cultural traditions, our sense of history, etc. – is lost.

therefore it is a good thing that human beings should continue to exist for as long as possible."[133] Lenman, however, ends up rejecting this view, since he believes that if the traits that constitute the final value of humanity, or of the white rhinoceros, are reasons for wanting the species to survive for *longer*, then surely they must also be reasons for wanting these species to be synchronically more numerous. Since Lenman finds latter claim implausible, this supposed implication functions as a *reductio* of the Argument from Final Value: "If it is unclear how it would make things better to stretch out, synchronically, in a single generation, the numbers of white rhinos, it is unclear why it should make things better to stretch them out diachronically by having more generations."[134]

I believe that Lenman's objection can be met. Like proponents of the Argument from Additional Lives, Lenman seems to assume that the only appropriate response to something's being finally valuable is to *promote* it, by maximizing its incidence in the world in both the diachronic *and* the synchronic sense. But this is a mistake. Cherishing, protecting, respecting, savoring, etc., are all modes of responding to something's final value which do not commit one to wanting more instances of it. Consider the question of *cultural* survival: It doubtless makes sense to maintain that aboriginal culture in Australia is valuable in its own right and should be helped to survive, by ensuring that

---

[133] "On Becoming Extinct", p. 255.

[134] Ibid., p. 256.

there continue to be members of aboriginal culture in the future. But clearly this does not commit one to claiming that the world would be better if aboriginal culture had *as many* members as possible. When what has final value is a form of life or a species, what we ought to care about is – to borrow Lenman's terminology – the ongoing instantiation of the universal, not the number of instantiations.

Note the importance of the word "ongoing" in the previous sentence. If I had said "what we ought to care about is the instantiation of the universal" this might have suggested the following argument: "If some type of thing x has final value then, for any time t, the world goes better if there exist tokens of x at t." This is implausible, as Lenman correctly points out: "We may think it a wonderful thing that the world contains many examples of jazz music, but how much should we regret its absence from, say, the world in the sixteenth century?"[135] Likewise, we don't bemoan the fact that humans didn't exist in the age of the dinosaurs. Just as deeming humanity to possess final value isn't a reason to want there to be as many humans as possible at present, humanity's final value isn't a reason to want it to be maximally extended across time.

However, it is a mistake to treat this as a *reductio* of our concern for the *future survival* of humanity, as Lenman does. This conclusion would only follow if our reasons for wanting there to be humans in the future were at the same time reasons for wanting

---

[135] Ibid., p. 257.

humanity to have existed in the pre-historic past. And this, again, would only be true on a conception of final value as something to be maximally promoted across space and time.[136] On the understanding of final value that animates the Argument from Final Value, by contrast, our reasons for wanting humans to exist in the future are not temporally neutral. Rather, they stem from our concern for the *survival* of those things we value finally.

6. Conclusion

I have defended the Procreation Asymmetry against several challenges, which all claim that the Asymmetry has counterintuitive implications to which we might not be willing to commit. Against these objections, I have demonstrated that: (1) because creating a person does not necessarily impose harm upon her, because parents have valid self-regarding reasons for wanting to procreate, and because the small risk of creating a miserable life by procreating is justifiable to each person, we can accept the Asymmetry without committing to anti-natalism. (2) If we agree that humanity has final value, we have a weighty reason to protect its survival for as long as possible - but no moral

---

[136] Note, incidentally, that a similar temporal asymmetry manifests itself in cases concerning our *own* existence at different points in time. It does not appear incoherent to be indifferent about not having existed five years *before* one was actually born, yet to be deathly afraid of no longer existing five years from now. The object of our prudential self-interest is *survival*, not maximal instantiation over time. For a penetrating discussion of temporal asymmetries of self-concern, see Frances Kamm, *Morality, Mortality Volume I: Death and Whom to Save From It* (New York: Oxford University Press, 1993), chapters 2 and 3.

reason to promote its growth to ever greater numbers. We can thus accept the

Asymmetry while retaining our concern for the future survival of humanity.

Bibliography

Anand, P., "The Philosophy of Intransitive Preference", *The Economic Journal* 103.417 (1993), pp. 337-346.

Anderson, E., *Value in Ethics and Economics* (Cambridge, MA: Harvard University Press, 1993).

Arrhenius, G., "Can the Person Affecting Restriction Solve the Problems in Population Ethics?", in Roberts & Wasserman (eds.) *Harming Future Persons* (London: Ashgate, 2009), pp. 289-314.

---, "The Person Affecting Restriction, Comparativism, and the Moral Status of Potential People", *Ethical Perspectives* 10.3 (2005), pp. 185-195.

---, "An Impossibility Theorem for Welfarist Axiology", *Economics and Philosophy* 16 (2000), pp. 247–266.

Ashford, E., "The Demandingness of Scanlon's Contractualism", *Ethics* 113 (2003), pp. 273-302.

Bader, R., "The Asymmetry" (unpublished manuscript).

---, "Aggregating v. Balancing" (unpublished manuscript).

---, "Neutrality and Conditional Goodness" (unpublished manuscript).

Beckstead, N., *On the Overwhelming Importance of Shaping the Far Future* (unpublished manuscript).

Benatar, D., *Better Never to Have Been: The Harm of Coming Into Existence* (Oxford: Oxford University Press, 2006).

Bennett, J., "On Maximizing Happiness" in R. I. Sikora and B. Barry (eds.), *Obligations to Future Generations* (Philadelphia: Temple University Press, 1978), pp. 61–73.

Bradley, B., "Asymmetries in Benefiting, Harming, and Creating", *The Journal of Ethics* 17 (2013), pp. 37-49.

---, "The Value of Endangered Species", *The Journal of Value Inquiry* 35 (2001), pp. 43-58.

Broome, J., *Climate Matters: Ethics for a Warming World* (New York: Norton, 2012).

---, "Should We Value Population?", *The Journal of Political Philosophy* 13.4 (2005), pp. 399-413.

---, *Weighing Lives* (Oxford: Oxford University Press, 2004).

---, "The Value of a Person" in his *Ethics out of Economics* (Cambridge: Cambridge University Press, 1999), pp. 228-242.

Bykvist, K., "The Benefits of Coming into Existence", *Philosophical Studies* 135.3 (2007), pp. 335-362.

---, "The Good, the Bad, and the Ethically Neutral", *Economics and Philosophy* 23 (2007), pp. 97-105.

Chang, R., "The Possibility of Parity", *Ethics* 112 (July 2002), pp. 659-688.

---, "Introduction" in R. Chang (ed.) *Incommensurability, Incomparability, and Practical Reason* (Cambridge, MA: Harvard University Press, 1997), pp. 1-34.

Cohen, G.A., "The Truth in Conservatism: Rescuing Conservatism from the Conservatives" (unpublished manuscript).

Darwall, S., *Welfare and Rational Care* (Princeton, NJ: Princeton University Press, 2002).

DeGrazia, D., *Creation Ethics: Reproduction, Genetics, and Quality of Life* (New York: Oxford University Press, 2012).

Driver, J., "Promises, Obligations and Abilities", *Philosophical Studies* 44 (1983), pp. 221-3.

Frick, J., "Future Persons and Victimless Wrongdoing" (unpublished manuscript).

---, "Contractualism and Social Risk: How to Count the Numbers Without Aggregating" (unpublished manuscript).

Friedman, A., "Intransitive Ethics", *Journal of Moral Philosophy* 6 (2009), pp. 277-297.

Greenspan, P., "Conditional Oughts and Hypothetical Imperatives," *Journal of Philosophy* 72 (1975), pp. 259-76.

Glover, J., *Causing Death and Saving Lives* (Harmondsworth: Penguin Books, 1977).

Hare, C., "Voices From Another World: Must We Respect the Interests of People Who Do Not, and Will Never, Exist?", *Ethics* 117 (2007), pp. 498-523.

Harman, E., "Critical Study of David Benatar's *Better Never To Have Been: The Harm of Coming into Existence*", *Noûs* 43.4 (2009), pp. 776–785.

---, "Can We Harm and Benefit in Creating?", *Philosophical Perspectives* 18 (2004), pp. 89-113.

Heyd, D., "Procreation and Value: Can Ethics Deal With Futurity Problems?",
　　Philosophia 18 (1988), pp. 151-170.

Holtug, N., "On the Value of Coming into Existence", *Journal of Ethics* 5.4 (2001), pp. 361-
　　384.

Hurka, T., "Value and Population Size", *Ethics* 93.3 (1983), pp. 496-507.

Kamm, F.M., *Bioethical Prescriptions: To Create, End, Choose, and Improve Lives* (Oxford and
　　New York: Oxford University Press, 2013).

---, *Intricate Ethics: Rights, Responsibilities, and Permissible Harm* (Oxford University Press,
　　2007).

---, *Morality, Mortality Volume I: Death and Whom to Save From It* (New York: Oxford
　　University Press, 1993).

---, "Supererogation and Obligation", *The Journal of Philosophy* 82.3 (1985), pp. 118-138.

Kavka, G., "The Futurity Problem" in R. I. Sikora and B. Barry (eds.), *Obligations to
　　Future Generations* (Philadelphia: Temple University Press, 1978), pp. 186--203.

Korsgaard, C., "On Having a Good" (unpublished manuscript).

---, "Interacting With Animals: A Kantian Account" in T. Beauchamp and R. G. Frey
　　(eds.) *The Oxford Handbook on Animal Ethics* (Oxford University Press, 2011), pp.
　　91-118.

---, "Two Distinctions in Goodness", in *Creating the Kingdom of Ends* (Cambridge
　　University Press, 1996), pp. 249-274.

Kraut, R., *Against Absolute Goodness* (Oxford University Press, 2011).

Lenman, J., "On Becoming Extinct", *Pacific Philosophical Quarterly* 83 (2002), pp. 253-269.

Leslie, J., "Why Not Let Life Become Extinct?", *Philosophy* 58.225 (Jul., 1983), pp. 329-338.

McMahan, J., "Causing People to Exist and Saving People's Lives" 17.1-2 *The Journal of
　　Ethics* (2013), pp. 5-35.

---, "Asymmetries in the Morality of Causing People to Exist" in M. Roberts and D.
　　Wasserman (eds.) *Harming Future Persons* (Dordrecht: Springer, 2009), pp. 49-68.

---, "Review: Problems of Population Theory", *Ethics* 92 (1981), pp. 96-127.

Moore, G.E., *Principia Ethica* (Oxford University Press, 1972).

Mulgan, T., "Dissolving the Mere Addition Paradox", *American Philosophical Quarterly* 37.4 (2000), pp. 359-372.

Narveson, J., "Moral Problems of Population", *The Monist* 57 (1973), pp. 62-86.

---, "Utilitarianism and New Generations", *Mind* 76 (1967), pp. 62–72.

Neumann, M., "Choosing and Describing: Sen and the Irrelevance of Independence Alternatives" [sic], *Theory and Decision* 63 (2007), pp. 79-94.

Nozick, R., *Philosophical Explanations* (Cambridge, MA: Belknap Press, 1981).

Olson, J., "Buck-Passing and the Consequentialism/Deontology Distinction" in D. Egonsson, J. Josfesson, B. Petersson, & T. Rönnow-Rasmussen (eds.) *Hommage a Wlodek: Philosophical Essays in Honour of Wlodek Rabinowicz* (Lund University, 2007).

Parfit, D., "Towards Theory X: Part One" (unpublished manuscript).

---, "Towards Theory X: Part Two" (unpublished manuscript).

---, *On What Matters* (Vol. 1&2) (Oxford University Press, 2013).

---, *Reasons and Persons* (Oxford University Press, 1984).

---, "Future Generations: Further Problems", *Philosophy and Public Affairs* 11 (1982), pp. 113-172.

Pettit, P., "Consequentialism" in Peter Singer (ed.) *Blackwell's Companion to Ethics* (Oxford: Blackwell, 1992), pp. 230-240.

Regan, D., "Why Am I my Brother's Keeper?" in Wallace, Pettit, Scheffler and Smith (eds.), *Reason and Value: Themes from the Moral Philosophy of Joseph Raz*, pp. 202-230.

Rippon, S., "In Defense of the Wide-Scope Instrumental Principle", *Journal of Ethics and Social Philosophy* 5.2 (2011), pp. 1-21.

Roberts, M., "A Variabilist Solution to the Asymmetry: Why Bringing the Miserable Child into Existence is Wrong but Leaving the Happy Child out of Existence Isn't" (unpublished manuscript).

---, "The Asymmetry: A Solution," *Theoria* 77 (2011), pp. 333-367.

Rolston, H., "Duties to Endangered Species", *Bioscience*, 35.11, The Biological Diversity Crisis (Dec., 1985), pp. 718-726.

Ross, W.D., *The Right and the Good* (Oxford: Clarendon Press, 1930).

Scanlon, T.M., "Ideas of the Good in Moral and Political Philosophy" (unpublished manuscript).

---, *Being Realistic about Reasons* (Oxford University Press, 2014).

---, *What We Owe to Each Other* (Cambridge, MA: Belknap Press, 1998).

---, "Contractualism and Utilitarianism", in A. Sen and B. Williams (eds.) *Utilitarianism and Beyond* (Cambridge: Cambridge University Press, 1982), pp. 103-110.

Scheffler, S., *Death and the Afterlife* (Oxford University Press, 2013).

---, "Immigration and the Significance of Culture", *Philosophy & Public Affairs* 35.2 (2007), pp. 93-125.

Sen, A., "Internal Consistency of Choice", *Econometrica* 6.3 (1993), pp. 495-521.

Shiffrin, S., "Wrongful Life, Procreative Responsibility, and the Significance of Harm," *Legal Theory* 5 (1999), pp. 117-48.

Smilansky, S., "Is There a Moral Obligation to Have Children?", *Journal of Applied Philosophy* 12.1 (1995), pp. 41-53.

Snedgar, J., "Reason Claims and Contrastivism about Reasons", *Philosophical Studies* 166:2 (2013), pp. 231-242.

Temkin, L.S., *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning* (Oxford University Press, 2012).

---, "Intransitivity and the Mere Addition Paradox", *Philosophy & Public Affairs* 16.2 (1987), pp. 138-187.

Tooley, M., "Value, Obligation and the Asymmetry Question", *Bioethics* 12.2 (1998), pp. 93-110.

---, *Abortion and Infanticide* (Oxford: Oxford University Press, 1985).

Weber, M., "Is Equality Essentially Comparative?", *Ethical Theory and Moral Practice* 10.2 (2007), pp. 209-226.

Willenken, T., "Deontic Cycling and the Structure of Commonsense Morality", *Ethics* 122.3 (2012), pp. 545-561.

Williams, B., "The Makropulos Case: Reflections on the Tedium of Immortality" in his *Problems of the Self* (Cambridge University Press, 1973), pp. 82-100.