

# Lives Worth Starting and the Non-Identity Problem

---

By

Isabella Ana-Maria Trifan

Submitted to:

Central European University

Department of Philosophy

In partial fulfilment of the requirements for the degree of Master of Arts

Supervisor: Professor Andres Moles

Budapest, Hungary

2015

## Abstract

In this paper I propose a person-affecting solution to the notorious non-identity problem. The non-identity problem arises in cases where a prospective parent's options are to either conceive a child that she knows will suffer from a severe disability, but would still have a life worth living, or conceive a different, healthy child at a different time. Most people believe the morally right choice is to bring the healthy child into existence rather than the disabled one, but it seems that this conclusion cannot be supported by concern for the disabled child's well-being, as it seems that for her the only two possibilities are a life with a disability, or no life at all. Following David Benatar, I argue that there is a distinction between a life worth starting and a life worth continuing. A bad situation that would make life not worth starting may not be bad enough to make life worth ending. If this is true, there need not be a paradox in saying that non-existence in the sense of never starting a life can be preferable to a life worth living understood as continuing. Further, I argue for the right not to be brought into existence with a life not worth starting, even though it may be worth continuing once it has started. I appeal to Ronald Dworkin's view of life as a challenge, and his understanding of a good life as a skilful performance in the face of this challenge. The opportunities for well-being, or the limits for how good our performance can be, are set by the parameters of our life, namely our specific circumstances. Some of these parameters are normative, which means that no matter how well one responds to their circumstances (i.e. they have a life worth continuing), their life still goes badly because they should not have been facing those circumstances in the first place (their life was not worth starting). One has the right not to be caused to exist with a life not worth starting, where this means a life that is set within inadequate parameters that impermissibly limit one's opportunities for well-being.

## Contents

Chapter 1: Introduction: The Non-Identity Paradox .....	1
Chapter 2: David Benatar and The Harm of Coming Into Existence.....	10
2.1. Can Someone Ever Be Harmed by Being Brought Into Existence?.....	11
2.2. Lives Worth Living: Worth Starting or Worth Continuing? .....	12
2.3. Starting Versus Continuing: A General Distinction? .....	15
2.4. Continuing to Live and the Badness of Death .....	18
Chapter 3: The Badness of Death.....	20
3.1. Can Death Ever be Bad for the One Who Dies? .....	20
3.2. What Is the Badness of Death? .....	27
3.2.1. Deprivation of future goods .....	28
3.2.2. Degree of personhood .....	32
3.2.3. Loss of investment .....	35
Chapter 4: What Is Wrong about Starting A Life that Is Not Worth Starting? .....	38
List of References.....	51

# Chapter 1

## Introduction: The Non-Identity Paradox

One of the primary concerns of moral theory has always been the various ways in which our actions affect other people. What are the consequences of our actions for other people? What is permissible to do given the foreseeable effects our actions will have on others? What do we owe each other? Do we have special duties to particular people? In addressing these questions we appeal to familiar moral principles that are philosophically sound and yield answers which are compatible with some of our deep-seated intuitions about what is right or wrong. Some of our actions affect not only existing people and the quality of their lives, but also the number of people that will come to exist, their identity, and the quality of their lives respectively. In cases, by no means few, where our actions affect the number and the identity of people that will exist in the future we might find that moral principles that virtually everyone accepts, and that have plausible implications in general, lead to conclusions that most people would find unacceptable for the morality of what we owe future people.

The non-identity problem is an umbrella-term that captures a set of problems arising from applying familiar moral principles to decisions about what we owe future people. It has been coined by Derek Parfit and has been under debate for more than 40 years now, ever since the publication of his “On Doing the Best for Our Children”<sup>1</sup> in 1976 and his later treatment of the paradox in his very influential *Reasons and Persons*.<sup>2</sup> The problem has also been formulated around the same time by Thomas Schwartz and Robert M. Adams.<sup>3</sup> The paradox arises when thinking about the obligations we have towards people who do not yet

---

<sup>1</sup>Derek Parfit, “On Doing the Best for Our Children,” in *Ethics and Population*, ed. Michael Bayles (Cambridge, MA: Schenkman, 1976).

<sup>2</sup>Derek Parfit, *Reasons and Persons* (Oxford: Oxford University Press, 1984).

<sup>3</sup>Thomas Schwartz, “Obligations To Posterity,” in *Obligations to Future Generations*, ed. Richard Sikora and Brian Barry, (Philadelphia: Temple University Press, 1978), 3–13; Robert M. Adams, “Existence, Self-Interest, and the Problem of Evil,” *Nous* 13 (1979), 53–65. The same problem is discussed by Gregory Kavka and he dubs it “the paradox of future individuals”. See Gregory S. Kavka, “The Paradox of Future Individuals,” *Philosophy and Public Affairs* 2 (1982), 93–112.

exist but whom we might cause to exist in a particular unavoidable flawed condition. This may happen, for instance, when prospective parents carry a serious genetic disorder that would be passed down to the child and that would cause her a considerable amount of suffering or place her at a significant social disadvantage. Most people would say that parents should not bring such a child into existence if they have full information about these consequences and can avoid bringing them about. However, assuming that the child's condition is not so bad so as to make her life not worth living, the choice parents face seems to be between offering this particular child a life that, while bad, is nevertheless worth living, or no life at all. These alternatives suggest that if the parents choose to have the child they would not be harming her in any way as she would have a life worth living, even though her condition would be bad. Many would still insist that this potential unhappy child should not be caused to exist, but this conclusion can be allowed to stand only if one can show how it is that non-existence can be preferable to a life worth living.

The difficulties raised by the non-identity problem are so formidable for our commonsense morality, that some authors, such as David Boonin and David Heyd,<sup>4</sup> have recommended that we accept the counterintuitive conclusion that causing a severely disabled child to exist even when one has full information and can avoid it is not morally wrong. To better understand what is at stake in the non-identity problem it is useful to consider two cases.<sup>5</sup>

Betty: Betty takes her newborn baby for a checkup. The doctor says that there is some good news and some bad news. The bad news is that, as things now stand, the baby is going to develop a handicap. The doctor explains that the handicap will be significant, meaning something that uncontroversially diminishes one's quality of life in a non-trivial way (e.g., more like blindness than like color-blindness). It will be non-terrible, meaning that although life with this handicap is considerably worse than life without it, it is nonetheless clear that it does not come close to making life worse than no life at all

<sup>4</sup> David Boonin, *The Non-Identity Problem and the Ethics of Future People* (Oxford: Oxford University Press, 2014); David Boonin, "How to Solve the Non-Identity Problem," *Public Affairs Quarterly* 22 (2008), 127-57; David Heyd, *Genethics: Moral Issues in the Creation of People* (Berkeley: University of California Press, 1992), 30-33.

<sup>5</sup> Both cases are from Boonin, "How to Solve," 127-28.

(assuming that such a thing is possible). And it will be irreversible, meaning that once the handicap develops, there will be nothing that anyone can do to treat it. The good news, though, is that Betty can prevent all of this from occurring simply by giving the baby a tiny pill once a day for the next two months. The pill is easy to administer, has no side-effects, and will be paid for by Betty's insurance company. Fully understanding all of this, Betty decides that having to give the baby a pill once a day for two months is too inconvenient and so chooses to throw the pills away. As a result, she ends up with an incurably blind child rather than a sighted child.

Wilma: Wilma is not yet pregnant, but is planning to try to have a baby and so goes to the doctor for a pre-conception checkup. The doctor says that there is some good news and some bad news. The bad news is that if Wilma conceives, as things now stand, she will conceive a child with a significant, non-terrible, irreversible handicap. The good news, though, is that Wilma can prevent this from occurring simply by taking a tiny pill once a day for the next two months before conceiving. The pill is easy to take, has no side-effects, and will be paid for by her insurance company. Fully understanding all of this, Wilma decides that having to take a pill once a day for two months before conceiving is too inconvenient and so chooses to throw the pills away and conceive at once. As a result, she ends up with an incurably blind child rather than a sighted child.

Virtually everyone would agree that both Betty and Wilma act immorally when they throw the pills away. The two cases seem to be similar in that their actions result in their having blind children rather than sighted children. But the crucial difference between the two cases is that what Betty does has a negative effect on an already existing person, whereas what Wilma does determines the identity of a future person. Betty's action has an identifiable victim, whereas Wilma's action seems to be victimless. And here is the crux of the non-identity problem: not only does Wilma's action not have a victim, but the action that is supposedly harming her child is in fact the same action that causes her child to exist in the first place. Assuming that our identity depends on the time when we were conceived,<sup>6</sup> had Wilma waited for two months before conceiving she would have given birth to a different child who would have been sighted. Her blind child, let us call her Pebbles as Boonin does, could only have been born if Wilma had conceived at once, which she in fact did. Thus Pebbles would not have been benefited by Wilma's taking the medication and waiting for two months, as the child born instead would have been a different one. So for Pebbles the only

---

<sup>6</sup> Parfit formulates this assumption as The Time-Dependence Thesis: "If any particular person had not been conceived within a month of the time when he was in fact conceived, he would in fact never have existed." The main idea is that our identity depends on the sperm and ovum that were combined for us to be conceived. See Parfit, *Reason and Persons*, 352.

two options would have been to either be brought into existence without sight, or not at all. This suggests that Wilma is not harming anyone when she throws away the pills. Betty is making her already existing child worse off than she would otherwise have been if Betty had given her the pills. The same child would have been sighted rather than blind. But Wilma's action does not make Pebbles worse off than she would otherwise have been because she would not have existed at all had Wilma acted differently. This suggests that Wilma is not doing anything wrong when she throws away the pills. And this is the conclusion that most people would find unacceptable.

David Boonin very clearly lays out the argument that the non-identity problem gives rise to, and it is this argument that has to be rebutted if we are to reject the conclusion that Wilma did nothing morally wrong when throwing away the pills.

- P1: Wilma's act of conceiving now rather than taking a pill once a day for two months before conceiving does not make Pebbles worse off than she would otherwise have been.
- P2: If P's act harms Q, then P's act makes Q worse off than Q would have been had P not done the act.
- C1: Wilma's act of conceiving now rather than taking a pill once every two months before conceiving does not harm Pebbles.
- P3: Wilma's act of conceiving now rather than taking a pill once every two months before conceiving does not harm anyone else.
- C2: Wilma's act of conceiving Pebbles does not harm anyone.
- P4: If P's act does not harm Q, then P's act does not wrong Q.
- C3: Wilma's act of conceiving Pebbles does not wrong anyone.
- P5: If P's act does not wrong anyone, then P's act is not wrong.
- C4: Wilma's act of conceiving Pebbles is not morally wrong.<sup>7</sup>

It is important to note that this argument rejects the conclusion that Wilma does something wrong for person-affecting reasons. Person-affecting accounts hold that an act is morally wrong by virtue of its harming or wronging someone. In other words, an act is wrong by virtue of someone's being a "victim" of it in some sense. This view is rivaled by impersonal accounts of morality, which maintain that an act can be morally wrong even if no identifiable individual is affected by it. Impersonal views seem to be better equipped for

---

<sup>7</sup> Boonin, "How to Solve," 129-31.

avoiding the challenges that the non-identity problem brings up. One such view is the one represented by Joel Feinberg and John Harris, later labelled by Jeff McMahan as “The Impersonal Comparative Approach.”<sup>8</sup> They hold that what is wrong about bringing severely disabled people into the world when one could have done otherwise is not that the person being born has been wronged, but simply that the agents responsible deliberately chose to “increase unnecessarily the amount of harm or suffering in the world.”<sup>9</sup> Derek Parfit voices the same idea when he writes: “If in either of two possible outcomes the same number of people would ever live, it would be worse if those who live are worse off, or have a lower quality of life, than those who would have lived.”<sup>10</sup> Thus in judging Wilma’s act to be wrong, one can base their judgment not on there being a victim of her act, but on comparing two possible states of affairs of the world, one in which a disabled child is born, the other in which a normally-abled child is born, and conclude that it is wrong, other things being equal, to intentionally bring about the worse outcome of the two. If one takes such a view, one is immune to the argument presented by Boonin as that argument only shows that Wilma’s act cannot be wrong for person-affecting reasons.

While taking an impersonal stance on morality is a plausible endeavour in itself and would avoid the non-identity problem, this type of view is not without its problems and counterintuitive implications.<sup>11</sup> Moreover, some of us might still feel that what is wrong about Wilma’s act has to do with the child being born, as when it comes to procreative decisions we

---

<sup>8</sup> Jeff McMahan, “Wrongful Life: Paradoxes in the Morality of Causing People to Exist,” in *Rational Commitment and Social Justice*, ed. Jules L. Coleman and Christopher Morris (New York: Cambridge University Press, 1998), 229.

<sup>9</sup> John Harris, *Wonderwoman and Superman: The Ethics of Human Biotechnology* (Oxford: Oxford University Press, 1992), 90.

<sup>10</sup> Parfit, *Reasons and Persons*, 360.

<sup>11</sup> See for example Parfit’s *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 suggests that if we base our morality about future people on utilitarian calculations with the aim to bring about the possible world that contains the most good, we may be forced to say that a world with a large number of people with a low quality of life is preferable to a world with less people that enjoy a better quality of life. See Parfit, *Reasons and Persons*, 387.



ultimately care about how our actions impact the quality of life of the people who actually come to exist. I believe a person-affecting solution to the non-identity problem would be more intuitive and would address the problem in the right manner, as the main concern here is with the morality of the relation between moral agents and other future moral agents. There is this special connection that we worry about: we cause other people to exist and we are responsible for the circumstances in which they are born, but if these circumstances are bad in one or many ways, we are ultimately not the ones forced to deal with them. The lives of our children become theirs to lead, and they alone are forced to live in whatever circumstances they find themselves as a result of our decision to cause them to exist. I believe prospective parents who, often times, postpone conceiving because they believe they will be able to offer their children better lives if they do are rightly motivated by thinking of the situation that the actual child will face – will it be a good one? could we have given her a better life? is it unfair for us to put someone into a bad situation by causing her to exist? I believe these are the right sort of concerns that should inform our procreative decisions. Thus, if the non-identity problem can be solved in person-affecting terms it will be a more intuitive kind of answer than the impersonal accounts can offer.

The purpose of this thesis is to explore a person-affecting solution to the non-identity problem that, to my knowledge, has not been convincingly pursued. The proposal belongs to David Benatar, as presented in his controversial book *Better Never to Have Been: The Harm of Coming Into Existence*.<sup>12</sup> He suggests that the apparent paradox that the non-identity issue forces us into (i.e. that we want to say that an act like Wilma's is wrong although it seems that there is no one being wronged) is due to a misunderstanding of what the phrase "life worth living" means. The bite of the non-identity problem depends on our assuming that Pebbles is not worse off by being brought into existence because her life, even if it is bad, is still good

---

<sup>12</sup> David Benatar, *Better Never to Have Been: The Harm of Coming into Existence* (Oxford: Clarendon Press, 2006).

enough to be worth living. Were the quality of her life so poor that it would not have been worth living (say, if she would have predictably been born blind and, moreover, with a condition that caused her considerable pain every day of her life), arguably we could say that she has been made worse off in the sense that non-existence would have been preferable to a life not worth living. But it is when we assume that her life is worth living that we are forced to admit that she could not have been better off by not coming into existence.

Benatar's suggestion is that there is a relevant difference between understanding a life worth living to mean "a life worth starting," or "a life worth continuing."<sup>13</sup> The point is that different standards apply when we judge whether a life is worth starting in the case of yet non-existing people as compared to whether an already existing life is worth continuing. "The judgment that an impairment is so bad that it makes life not worth continuing is usually made at a much higher threshold than the judgment that an impairment is sufficiently bad to make life not worth beginning."<sup>14</sup>

If correct, the distinction seems to dissolve the necessary logical difficulty in saying that non-existence can be preferable to a life worth living, since it takes this judgment to be about lives worth continuing. So even though it could not be the case that it is preferable not to begin a life that would be worth starting, "[t]here is nothing paradoxical about the claim that it is preferable not to begin a life that would be worth continuing."<sup>15</sup>

Ultimately, Benatar's project is not only to argue that coming into existence can sometimes be a harm (in cases like Pebbles'), but that it is always a harm. He believes that we would all be better off not having come into existence. As a matter of fact, any actual life contains some bad. Even the luckiest of people experience the loss of a loved person, some disappointment, ill-health in old age and death. No matter how much good one's life contains, Benatar believes that never existing would have been better for them because some evil will

---

<sup>13</sup> Benatar, *Better Never to Have Been*, 22.

<sup>14</sup> *Idem*, 23.

<sup>15</sup> *Idem*, 24.

inevitably befall them, and this is enough to make never coming into existence preferable. He bases his argument on an asymmetry according to which the absence of pain is good even if there is no one there to benefit from it, whereas the absence of pleasure is not bad unless there is someone there to be deprived of it.<sup>16</sup>

I do not follow Benatar in arguing that coming into existence is always a harm. In fact, this thesis does not engage at all with the question of when exactly coming into existence is a harm. Instead, the purpose of this paper is merely to establish that it *can* be a harm to come into existence, and that Benatar's distinction is promising for solving the non-identity problem in person-affecting terms. To this aim, I start chapter two by addressing the question of whether coming into existence can ever be a harm, and I conclude that it is a harm when the life one would have would not be worth living. In the next section I introduce Benatar's idea that a life worth starting is not the same as a life worth continuing, in support of the view that there is no necessary logical paradox in claiming that one would be better off never *starting* a life that would nevertheless be worth living, understood as worth *continuing*. However, I argue that Benatar does not do enough to support the cogency of this distinction. In the final section of the second chapter I introduce a different reason for which we have different standards in deciding whether a life is not worth starting as opposed to whether it is not worth continuing. I suggest that the element that raises the threshold for when a life is not worth continuing is the badness of death. Death only happens to actual people, and the way it affects them explains why they need to be in an especially bad condition for it to be rational to judge their life worth ending. For the badness of death to play the role I believe it has in judgments about when life is worth ending, but not when it is worth beginning, I first need to show how death can ever be bad for the person who dies, which I take up in the first section of chapter three. Then, I sketch an account of what makes death bad for people. Finally, I suggest how

---

<sup>16</sup> *Idem*, 30.

Benatar's distinction can help solve the non-identity problem. I believe we can reject premise four in Boonin's argument, namely that if P's act does not harm Q, it does not wrong Q. I do not believe that one is strictly speaking harmed by being brought into existence when their life is worth continuing, but I suggest that one can be wronged by having their life started when it was not worth starting, and this is sufficient to solve the non-identity problem in person-affecting terms.

## Chapter 2

### David Benatar and The Harm of Coming Into Existence

Before setting out to show that coming into existence can be a harm, some preliminary clarifications are needed. I will understand the definition of harm to be the following.

Individual P1 is harmed by P2 iff: by doing (or allowing) act a, P2 brings it about that P1 is worse off in terms of well-being than P1 would have been in the absence of a. <sup>17</sup>

Showing that one is being harmed by being brought into existence is one way in which to overcome the non-identity paradox, and this is what Benatar tries to do. But many, including myself, do not ultimately believe that a successful argument can be made to the effect that one would be better off never existing than with a life worth continuing (even if not worth starting). Instead, I hold that there are grounds to think one is nevertheless wronged by being brought into existence with a life not worth starting, even if it would be worth continuing. It is generally accepted that one can be wronged without being made worse off, for example when one is lied to even though there is no consequence of the lie, i.e. one is not worse off as a result of the lie. One has the right not to be lied to and one is wronged when that right is violated, even if she is not being made worse off than had the lie not occurred.<sup>18</sup> Similarly, I will argue in the last chapter that there are grounds for a right not be brought into existence when one would have a life not worth starting. Thus if one is conceived when her life would not be worth starting she is wronged even if her life would be worth continuing after it has started.

---

<sup>17</sup> Thomas Søbirk Petersen, “Being worse off – but in comparison with what?: On the baseline problem of harm and the harm principle,” *Res Publica* 20 (2014), 208.

<sup>18</sup> Here I subscribe to the common view that harming someone involves making them worse off than they would have been otherwise. There are authors, however, who think that one can be harmed even without being made worse off. See, e.g., Joseph Raz, *The Morality of Freedom* (New York: Oxford University Press, 1986), 416 where he argues that we harm people by failing to provide them with what they are due, even if this omission does not leave them worse off than they would have been had we fulfilled our duties towards them. So on this kind of view, whenever you are wronged you are also harmed regardless of whether you are actually worse off or not as a result of that wrong. But I leave such views aside for now.

In the next sections of this chapter I will evaluate and expand on Benatar's proposal for solving the non-identity problem. Since he is concerned with showing that coming into existence is always a harm, I will evaluate whether he manages to show (i) that coming into existence is ever a harm, and I will argue that he does - in cases where life would not be worth living -, and (2) that coming into existence can be a harm even when life would be worth living, i.e. worth continuing once it has started, and I will suggest that he is unsuccessful in this endeavour. I will therefore not go on to discuss the details of his account further. Instead, I will support the distinction he introduces between a life worth starting and a life worth continuing in order to argue, in the last chapter, that although one cannot be *harmed* by being brought into existence when their life would be worth continuing, they are nevertheless wronged by being caused to exist when their lives would not be worth starting.

### **2.1. Can Someone Ever Be Harmed by Being Brought Into Existence?**

To address the non-identity problem in cases such as Pebbles', Benatar needs to show that she has been harmed by being brought into existence. And in order to say that there are some cases where people are harmed by being caused to exist it must first be established how it can ever be the case that coming into existence is a harm, as this is by no means obvious.

Benatar explains the standard argument that is put forward in support of the position that coming into existence can never be a harm, and it is in fact a strong conclusion of the non-identity problem, namely that coming into existence cannot be a harm even when one's life is not worth living:

1. For something to harm somebody, it must make that person worse off.
2. The "worse off" relation is a relation between two states.
3. Thus, for somebody to be worse off in some state (such as existence), the alternative state, with which it is compared, must be the one in which he is less badly (or better) off.
4. But non-existence is not a state in which anybody can be, and thus cannot be compared with existence.
5. Thus coming into existence cannot be *worse* than never coming into existence.

6. Therefore, coming into existence cannot be a harm.<sup>19</sup>

This argument contends that for someone to be harmed she needs to be worse off in the situation she is in as a result of the harmful event as compared to the state she could have been in had the event not occurred. In the case of non-existence, however, there is no proper counterfactual state in which the person could have been in, since by definition she would not have existed at all.

Along with Benatar, I follow Joel Feinberg in maintaining that this comparison is intelligible nonetheless.<sup>20</sup> One does not need to be able to say that the person would exist in the counterfactual situation in order to say that she would be better off or worse off in it. It is enough to say that the state of affairs in which one finds herself is bad enough that non-existence (no state at all) would be preferable. Our attitudes towards suicide and euthanasia are proof that such a claim makes sense. Many of us think that suicide and euthanasia can be rational in some cases and irrational in others. As Feinberg suggests, when people claim that their life is so bad that they would be better off dead, they do not mean that they would exist in a different state after they die, in which they would be happier. They simply mean that their current state is so unbearable, that not being in any kind of state would be better. And many of us take this position to be not only intelligible, but rational as well. If we can accept that one is better off dead than living a horrible life that is not worth living, we should be able to accept that not coming into existence at all can be better than living a life not worth living, and therefore we should accept that being brought into existence with such a life is a harm.

## **2.2. Lives Worth Living: Worth Starting or Worth Continuing?**

Assuming the comparison between existence and non-existence is intelligible and it is accepted that non-existence is preferable to a life not worth living, the non-identity paradox

---

<sup>19</sup> Benatar, *Better Never to Have Been*, 20-21. Original emphasis.

<sup>20</sup> Joel Feinberg, "Wrongful Life and the Counterfactual Element in Harming," *Freedom and Fulfilment* (Princeton: Princeton University Press, 1992), 3-36.

arises in the case of lives worth living. No matter how bad one's life is, provided that it is above the threshold of a life worth living, it seems that having been brought into existence could not have been a harm. A life of some (minimal) value is always preferable to non-existence. And yet many of us still want to say that it is in some way wrong to bring a child into the world with very bad life prospects (even if not so bad as to make her life not worth living). Benatar himself insists that it is not only a wrong, but a harm to her.

It is at this point that he introduces the distinction which this thesis focuses on. He suggests that the non-identity paradox rests on the ambiguity of the phrase "a life worth living." We can take it to mean "a life worth starting" or "a life worth continuing". This ambiguity runs in parallel with that of the two possible meanings of "non-existence". Non-existence can mean never coming into existence, or ceasing to exist. Ceasing to exist only happens to actual people, and, as we shall see later, involves certain losses that never coming into existence does not. Similarly, judging whether a life is not worth starting is different from judging whether a life is not worth continuing. The second type of judgment only applies to actual people (as Benatar puts it, these are "present-life"<sup>21</sup> judgments). The first type of judgment applies to potential, future people ("future-life" judgments).

Crucially, the standards that we use in these two types of evaluation are not the same, Benatar contends. The same bad condition (say, missing a leg) might not be enough to deem someone's life not worth continuing (i.e. worth ending), whereas it might be enough to deem a potential person's life not worth starting. The threshold for judging a life worth ending is higher than that for judging a life not worth starting: one's condition needs to be much worse to make life worth ending than it needs to be to make a life not worth starting.

The relevance of this distinction for our judgments about future people is that it seems to avoid the necessary logical paradox of saying that non-existence can be preferable to a life

---

<sup>21</sup> Benatar, *Better Never to Have Been*, 22.



worth living. If a life is worth living in the sense that it is worth *starting* it is true that non-existence (not beginning a life) cannot be preferable to it. But if a life is not worth starting, even though it might be worth continuing, it *can* potentially be said that non-existence (not beginning a life) is preferable to it. Thus, the paradox need not arise, depending on how we pair up the two meanings of “a life worth living” with the two meanings of “non-existence”. Not starting a life cannot be preferable to a life worth starting, and ending a life cannot be preferable to a life worth continuing. But if a life is worth continuing it does not follow that it was worth starting. It is in this sense that non-existence, i.e. never starting a life, can be preferable to a life worth living, i.e. continuing.

For Benatar’s distinction to solve the non-identity problem in cases such as Wilma’s, three things need to be established: (i) that the distinction is cogent, (ii) that the future-life evaluation is the relevant one for determining the morality of bringing new people into the world, and (iii) what it is that makes life not worth starting. Question (i) will be taken up in the next subchapter, where I argue that Benatar’s justification for there being two different standards for judging a life to be worth starting as opposed to worth continuing is not convincing. I then proceed to offer my own reason for defending his distinction. Question (ii) is crucial for ultimately solving the non-identity issue in person-affecting terms because this requires showing that having my life started when it was not worth starting is a harm to me (or at least a wrong). Even if we accept the future-life versus present-life distinction it is not clear why having my life started when it would not be worth starting, but would be worth continuing, constitutes a harm or a wrong. This will be taken up in the last chapter of the thesis. As for (iii), I do not intend to provide a substantive answer to the question of what it is that makes life not worth starting although it does not make life worth ending. My purpose here is merely to establish how it *can be* the case and that this shows promise for moving the

non-identity debate forward. From here on I will simply assume that an appropriate list of criteria for (iii) can be found.

### **2.3. Starting Versus Continuing: A General Distinction?**

I believe that Benatar's attempt to establish the cogency of his distinction fails. Indeed it has been directly challenged by people such as David DeGrazia. DeGrazia asks "Why should it matter whether one is already in existence when the question is raised as to whether one's life is worth living? What is the basis for there being two different standards?"<sup>22</sup>

In support of the difference between the standards that we use in present-life as opposed to future-life cases Benatar seems to offer two kinds of reasons. First, he seems to suggest that different reasons applying to starting a project as opposed to continuing it is a basic distinction which is true generally. He offers a comparison with a more trivial case, namely going to the cinema. A film can be sufficiently bad that it is not worth going to the cinema to watch it, but, if you are already there, the same film is not considered to be bad enough that it would be worth leaving the cinema before it finishes. Apart from the fact that many might disprove this example (many would rather leave the cinema rather than sit through a bad film), it seems that when it (and other examples from everyday life with the same structure) works it is due to loss aversion. I suggest that in most or all such cases from everyday life, while it may be true that people tend to judge that the same state of affairs or course of action is not worth abandoning even though it would not have been worth starting in the first place, this evaluation is irrational because it is based on a quirk of our psychology when we contemplate losses.

---

<sup>22</sup> David DeGrazia, "Is it wrong to impose the harms of human life? A reply to Benatar," *Theoretical Medicine and Bioethics* 31 (2010), 320.

This feature of human psychology has been amply theorized and empirically documented. Amos Tversky and Daniel Kahneman call it loss aversion.<sup>23</sup> Richard Thaler calls it “the endowment effect”.<sup>24</sup> William Samuelson and Richard Zeckhauser talk about status quo bias.<sup>25</sup> They all describe an irrational asymmetry in people’s attitudes towards loss as opposed to gain. One tends to prefer keeping what she already has rather than risking losing it, even when an equal chance for gain is presented to her. As Tversky and Kahneman explain, “[t]he common reluctance to accept a fair bet on the toss of a coin suggests that the displeasure of losing a sum of money exceeds the pleasure of winning the same amount.”<sup>26</sup>

It seems that while loss aversion can explain instances where people are reluctant to leave the cinema during a bad film even though they wish they had never come to see it in the first place (and had not invested time and money in it), I suggest that loss aversion is not what is at work in present-life/future-life judgments. Loss aversion is an irrational psychological feature of human beings, but there is a different, rational reason for which we take a life worth starting not to be the same as a life worth continuing. A closely related explanation to loss aversion is to do with opportunity costs. We might think that the difference between present-life and future-life cases is that in the former we have already incurred opportunity costs, which makes it rational for us not to want to lose this investment. While this is a valid element of what I think is the best explanation, it is only a part of it, as I will propose in chapter three.<sup>27</sup>

---

<sup>23</sup> Amos Tversky and Daniel Kahneman, “Rational Choice and the Framing of Decisions,” *The Journal of Business* 59 (1986), S251-S278.

<sup>24</sup> Richard Thaler, “Toward a Positive Theory of Consumer Choice,” *Journal of Economic Behavior and Organization* 1 (1980), 39-60.

<sup>25</sup> William Samuelson and Richard Zeckhauser, “Status Quo Bias in Decision Making,” *Journal of Risk and Uncertainty* 1 (1988), 7-59.

<sup>26</sup> Kahneman and Tversky, “Rational Choice,” S258.

<sup>27</sup> A potentially better explanation of the basic distinction between starting versus continuing a project that perhaps Benatar had in mind (although it is not clear that he did) would be a conservative view of the kind that G. A. Cohen takes up. He holds that something’s having intrinsic value is reason enough to conserve it even if its destruction would give way to something even more valuable. A conservative view of preserving what is valuable might be a good explanation for why we have reason to want to continue to live despite certain hardships befalling us. This explanation can happily run in parallel with the one I propose in section 2.4., namely that in decisions about ending one’s life the badness of death hangs heavily in the balance and explains why a

The second reason he cites for his distinction is the presence of interests in existing in present-life cases, but not future-life cases. He assumes that we do not begin to exist in a morally relevant sense at conception, and that coming into existence is a gradual process.<sup>28</sup> Thus there is no one point in time when we start existing as persons. But, he assumes, only beings that exist in the relevant sense have strong interests in existing. Developing persons (for example, embryos) do not yet have interests in existing. These interests develop and grow stronger during the process of our coming into existence. Benatar's appeal to interests in existing have the following role. We generally think that sacrificing someone's leg to save their life is a benefit to them. The reason is that their interest in existing trumps their interest in not being impaired. But in the case of people who do not exist in the moral sense, for example if one finds out very early during pregnancy that their child will be born without a leg, there is no interest in existing to speak of. There is no such interest that we would be protecting in bringing such a child into existence with an impairment.

If we think Benatar's distinction rests on the subject's having or lacking interests in existing, it does not seem to follow that bringing people into existence with a serious impairment is a harm to them. This is because Benatar only establishes that we do not have reason to think that future people would be benefited by being brought into existence with impairments (because they do not have an interest in existing that would be protected by causing this impairment). However, it does not establish that it would be a harm to them to be brought into existence with a serious impairment either. This is because the reason that we have in present-life cases to refrain from causing people impairments is lacking in future-life cases: namely the interest in not being caused impairments. While it is correct that future

---

bad situation needs to be so much worse if it is to make dying (as opposed to never coming into existence) preferable to continuing to live. For the conservative attitude towards what is of value see G. A. Cohen, "Rescuing Conservatism", in *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, ed. R. Jay Wallace, Rahul Kumar, and Samuel Freeman (Oxford: Oxford University Press, 2011).

<sup>28</sup> See Benatar, "*Better Never to Have Been*," 25. I will also say more about the process of coming into existence in the relevant sense in section 3.2.2.

people lack an interest in existing, they also lack an interest in not being caused impairments. Thus the lack of an interest in existing that is supposed to bring down the standard for judgments in future-life cases is cancelled out by the lack of an interest in being caused impairments, which leaves us right where we started.

#### **2.4. Continuing to Live and the Badness of Death**

The standards that we use to judge whether a life is worth starting are different from those that we use to deem a life worth continuing. But this fact is not due to there being interest in existing that we are protecting in present-life cases as opposed to future-life cases, although this consideration plays a role, as we shall soon see. Nor is this distinction entirely attributable to an inherent difference between starting a project and continuing one. While such a difference might explain some cases from everyday life in which we are reluctant to terminate a project we have already started due to loss aversion or sunk opportunity costs, I believe that judgments about present-life cases and future-life cases are importantly different. There is a rational reason for which our standards are, and should be, distinct, a reason apart from having a conservative attitude towards what is valuable. I submit that the reason why the threshold for deeming a life worth ending is higher than that for considering a life not worth starting is that in the present-life case we need to take into account the badness of death.

Death only happens to existing people, and arguably it is a real evil in person-affecting terms. That is, death is a misfortune for the person who dies. The badness of death varies from case to case: most of us believe that the death of a teenager is worse for her than the death of an 90 year-old is, and that in some cases death is not bad at all for the person who dies (for example if the alternative would have been a lifetime of pain). When one finds herself in a bad situation, such as having to sacrifice a leg in order to survive, and we ask whether this impairment would make her life not worth continuing, we are asking whether her new situation would be so bad that it would be even worse for her than death. And there are

specific losses involved in death that might make it the case that it is preferable to avoid them rather than avoid the misfortune in continuing to live without a limb. But never coming into existence does not involve the losses that the death of an already existing person does. It is this specific badness involved in death (and not in never coming into existence) that explain why the threshold for deeming a life not worth continuing is higher than for judging a life not worth starting.

As my argument focuses around the badness of death and its role in our judgments about existing as opposed to just potential people, it will proceed as follows. First I will address the fundamental question of how it can ever be the case that death is bad for the person who dies. Then I can proceed to fleshing out what it is about death that is bad, drawing heavily on the work of Jeff McMahan. Finally, once Benatar's distinction between present-life judgments and future-life judgments has been established, I turn to the issue of why future-life judgments should be the ones relevant for determining the morality of bringing new people into the world. The thesis will conclude with reflections on how the argument could provide a satisfactory solution to the non-identity problem in person-affecting terms.

## Chapter 3

### The Badness of Death

In order to see whether the different standards that we use in making judgments about present-life cases as opposed to future-life cases is explained by the weighing in of the badness of death, we need to offer an account of what exactly the badness of death is. Before we can say what is bad about dying, however, we need to address a more basic question: can death ever be a misfortune for the person who dies?<sup>29</sup>

#### **3.1. Can Death Ever be Bad for the One Who Dies?**

There is a strong tradition, going back to Epicurus, which suggests death can never be bad for the person who dies because at the moment of death there is no subject to whom the misfortune can happen.<sup>30</sup> This argument presupposes what Jeff McMahan calls *The Existence Requirement* – “A person can be the subject of some misfortune only if he exists at the time the misfortune occurs.”<sup>31</sup> This requirement seems quite intuitive and plausible: in order for me to have, say, a headache, I need to exist.

However, some of us, together with McMahan, might still feel that there is rational cause for feeling pity for someone’s premature death quite apart from the reasons we have to consider it a misfortune for her friends and relatives, or to feel that the state of affairs in the world has worsened, as it were, due to this person’s disappearance. We might still feel there is something to be said about the misfortune that an untimely death represents for the person herself. How, then, could it be the case that death is a misfortune for the person who dies? To make such a case one could take one of two routes: either reject the Epicurean argument, or

---

<sup>29</sup> I assume throughout the paper that death is a person’s ceasing to exist.

<sup>30</sup> Epicurus, “Letter to Menoeceus,” 124b-127a, in *Letters, Principal Doctrines, and Vatican Sayings*, ed. Russel M. Geer (Indianapolis: Bobbs-Merrill, 1964).

<sup>31</sup> Jeff McMahan, “Death and the Value of Life,” *Ethics* 99 (1988), 33. The Epicurean argument should be distinguished from two other similar ones. The first holds that a person cannot be the subject of some misfortune unless she experiences it as being bad. The second is that an event can be bad for someone only if it makes a difference to her conscious experience. For an explanation of the differences between these arguments see McMahan, “Death,” 33-34.

attempt to make it compatible with some of the most commonly held beliefs we have about life and death, most importantly that continuing to live can be good for the person in question.

McMahan develops arguments in line with both strategies, although he ultimately believes that the Epicurean argument, and with it the Existence Requirement, should be rejected altogether.<sup>32</sup> He first proposes a way to reconcile the Epicurean Argument, namely that death cannot be bad for us, with the claim that continuing to live can be good for us. This view aims to establish that death is not bad in a person-affecting sense, but neither is it bad in fully impersonal terms. If we take up what McMahan dubs the Reconciliation Strategy, death is said to be bad in a *quasi-impersonal* sense. This is to say that its badness is due to excluding “what would be good for a person – namely, continuing to exist.”<sup>33</sup> I cannot hope to detail this sophisticated line of argument<sup>34</sup> here for lack of space, but also because I subscribe to the belief that the Existence Requirement can be overcome altogether.

To reject the Epicurean argument that death cannot be bad for the person who dies one has two options: either reject the underlying principle’s validity altogether, or argue that while the Existence Requirement is generally valid, death is a non-standard case to which it does not apply.

McMahan seems to be ambiguous between which of these two strategies he subscribes to. He starts off by saying that we should feel compelled to reject the Existence Requirement altogether in cases such as the following. Imagine that a person goes on holiday on a remote island, where he gets killed by a shark. Right after his death, his life’s work collapses back home. McMahan thinks the destruction of this person’s work is a misfortune to her even if she did not find out about it and, crucially, even if it occurred after she no longer existed.<sup>35</sup> If this is plausible, it means that the Existence Requirement should be rejected.

---

<sup>32</sup> McMahan, “Death,” pp. 38-39.

<sup>33</sup> McMahan, “Death,” p. 38.

<sup>34</sup> See McMahan, “Death,” pp. 34-38 for the full account of the Reconciliation Strategy.

<sup>35</sup> *Idem*, 39.



That someone's losing their life's work or their legacy after they die is a misfortune to them is quite controversial. If one spends her life building a house that is then demolished in an earthquake after her premature death, is this bad for her? Would the untimely death of one's child be a misfortune to the parent if it occurred after her own death? Some might find these cases enough to show that the Existence Requirement is false, but it is not obvious that they are, in the absence of further compelling arguments. To be sure, it is not absurd to imagine that such further arguments exist, as we shall see shortly. For the time being, however, the mere intuition that one's losing her legacy after her death is a misfortune to her, by itself, does not seem persuasive enough to reject the Existence Requirement.

Then, McMahan suggests that even if one does not believe that one's losing their life's work after they die is bad for them, death itself should be a sufficient counterexample to the principle's validity.

Apart from suffering great pain, it is hard to think of a clearer example than death of something that most people believe to be in most cases bad for the person to whom it happens. The Epicurean simply denies what most of us believe. Death, he claims, cannot be bad for us because when it occurs we will not exist at all. But that is precisely what we object to: that we will no longer exist when we might otherwise be enjoying the benefits of life.<sup>36</sup>

This line of argument does not seem very compelling either. To say that what is bad about death is precisely that we will no longer exist is begging the question against the Epicurean. The Epicurean argument offers a reason for which death cannot be bad for us, namely that, by logical necessity, we need to exist in order to suffer any kind of misfortune. This reply simply says "stopping to exist is itself the misfortune," but it does not seem to do anything to dislodge the argument since it just assumes the conclusion it is supposed to show.

---

<sup>36</sup> *Ibidem.*

Finally, McMahan seems to slide into the other strategy for saying that death is a misfortune to us, namely keeping the validity of the Existence Requirement intact, but suggesting that death is a nonstandard case to which it does not apply.

Indeed, [the Existence Requirement] seems,... to be a simple misgeneralization from standard cases to an importantly different nonstandard case. In most instances it is a necessary truth that a person must exist to be the subject of some misfortune: I cannot, for example, suffer the pain of a toothache unless I exist. But death is obviously a special case. To insist that it cannot be an evil because it does not meet a condition that most if not all other evils satisfy is tantamount to ruling it out as an evil simply because it has special features.<sup>37</sup>

In this paragraph it appears that he is accepting the general validity of the principle with respect to most cases, but suggests that death has particular features. It is not clear what these special features are, as he only he says two things in relation to this. One is that death involves a special kind of “not existing” as compared to the “not existing” of never existing in the first place. The “not existing” involved in death only happens to actual people, while never existing does not. So it deprives actual people of what would otherwise be good for them in continuing to live. And here is the second feature of death: its badness is privative. As opposed to suffering or pain, death has no intrinsic negative features. When it is bad this is due to what it deprives us of.<sup>38</sup> But the fact that death only happens to actual people and that its badness is of privative nature alone still does not suffice to show that it is a nonstandard case to which the Existence Requirement does not apply. There are many instances of privative evils for which the principle holds true. For example, not affording to go on holiday is bad for me because it deprives me of all the good that such a trip would do me. Not going on holiday is not an evil in itself, although it is bad in a privative sense. And it could not constitute a privative evil for me unless I existed.

---

<sup>37</sup> *Ibidem.*

<sup>38</sup> *Ibidem.*

Moreover, I believe that trying to keep the Existence Requirement in place while arguing that death is a non-standard case and hence resists its application is doomed to failure. The Epicurean argument turns specifically on the condition of someone's not existing at the time that a misfortune occurs. Death, insofar as it is defined as one's ceasing to exist, is a *typical* case that the argument addresses, not a nonstandard one.

It seems that the best option to reconsider in more detail is the possibility that the Existence Requirement is false. Thus it is a question of finding an appropriate counterexample to prove it wrong. I mentioned earlier that losing one's life work after one dies is not, by itself, an obvious contender. There is, however, something more to be said on this matter, something that McMahan mentions later in his paper and develops at length in his book *The Ethics of Killing*.<sup>39</sup> This is the idea that a later event can affect the value of a past event by affecting its meaning. If one invests all their efforts and energy into a project, but dies before getting a chance to complete it or to reap its fruits, death deprives all that past effort of its meaning and turns one's life into a tragic waste of time. Michael Lockwood expresses this point in an evocative manner: "Set against an ideal of human life as a meaningful whole, we can see that premature death can, as it were, make nonsense of much of what has gone before. Earlier actions, preparations, planning, whose entire purpose and rationale lay in their being directed towards some future goal, become, in the face of an untimely death, retrospectively pointless—bridges, so to speak, that terminate in mid-air, roads that lead nowhere."<sup>40</sup>

The idea that later events can affect the meaning of past ones is fairly intuitive. What is more controversial is that a retroactive decrease in meaning of a past event corresponds to a retroactive decrease in the value of that past event as well. David Velleman argues that "the daily well-being of your former self is a feature of the past, beyond alteration.... [Thus] when

<sup>39</sup> Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (Oxford University Press, 2002), 174-180.

<sup>40</sup> Michael Lockwood, "Singer on Killing and the Preference for Life," quoted in McMahan, *The Ethics of Killing*, 177.

subsequent developments alter the meaning of an event they can alter its contribution to the value of one's life, but they cannot retroactively change the impact that it had on one's well-being at the time."<sup>41</sup> Velleman accepts that later happenings can affect the meaning of past events, and that this can change the value of one's life overall, but they cannot change the value that they had for one's well-being at the time.

To this, McMahan has the following reply. Meaning and some kinds of value are alike in that they have a relational dimension: the meaning and value of something that is happening now is influenced by what has happened in the past. So if I get a certain job now, the meaning of getting this job is different in relation to what has happened in the past: it can be the coronation of my past work and sacrifices, or it can be a fluke or else something that I never even desired. According to which of those scenarios is my real past, the meaning of my getting this job differs. So does the value that it adds to my well-being: it seems that the value it adds to my well-being is greater if the job is the result of my planning and efforts as compared to it being like manna from heaven. If this is true, there is no reason to think that the relational dimension of meaning and of this kind of value would not work in the other direction as well. Thus getting the desired job now renders my past efforts meaningful, and adds to the value they had for my well-being. As this is precisely what Velleman denies, McMahan writes:

According to Velleman, the value of an event in a person's life, or the way that the event affects the person's well-being at the time, is not relational. The problem with this, however, is that it divorces the value that an event has for a person at the time it occurs from the meaning that the event has within the person's life. Whether a success is a reward or a mere windfall must, therefore, be irrelevant to how it affects the person's well-being at the time. But this seems implausible. A success is better than a windfall, even if the tangible aspects of the event are the same in either case.<sup>42</sup>

<sup>41</sup> J. David Velleman, "Well-Being and Time," quoted in McMahan, *The Ethics of Killing*, 179.

<sup>42</sup> McMahan, *The Ethics of Killing*, 180.

So if we are willing to accept that well-being is not just a matter of a person's intrinsic properties moment by moment (in which case the well-being at past moments cannot be altered), and allow that well-being is multidimensional in that it comprises, among other things, relational components such as meaning and value, then we can accept that later events can alter the value that past events added to our past well-being. Call this the Principle of Retroactive Effects.

If the Principle of Retroactive Effects is plausible, we are now in the position to counter the Existence Requirement more effectively. For we now seem to have two principles that result in conflicting evaluations of the same case. The Existence Requirement suggests that losing's one's life work after one dies cannot be a misfortune for him, while the Principle of Retroactive Effects implies that it is, since it renders his life pointless. In such a situation one needs to decide which principle to keep, and which to discard. A good reason to discard the Existence Requirement is that it goes against some of our deep-seated beliefs and intuitions about life and death. These include the belief that part of the wrongness of killing consists in death being bad for the victim, and the belief that suicide can be rational in some cases but irrational in others. Furthermore, if we really believed that death can never be bad for us, we would have no objection against dying prematurely, provided that we could be assured that there would be no suffering or bad consequences for the ones left alive. Thus out of the two principles, the Existence Requirement seems to be the more counter-intuitive one.

Some might worry that the Principle of Retroactive Effects is not that easier to accept either. This is understandable since, after all, it seems to go against our commonsense intuition about the past being fixed. The thing to keep in mind, however, is that the principle does not claim that present events can alter past events or alter that part of well-being that has to do with the subjective character of a person's mental states at certain times. What it proposes is that there are other elements that affect our well-being as well, and these elements

are defined relationally, and can therefore influence one another over time. If we find it intuitive to say that, for example, our past strife is rendered either pointless or meaningful by our present achievements or lack thereof, we can see the appeal of the Principle of Retroactive Effects.

Suppose we accept the Principle of Retroactive Effects. We might still find it odd to say that there is no truth at all to the Existence Requirement, since, in itself, it looks quite plausible. It is worth clarifying at this point that in rejecting the Existence Requirement my aim has been to reject the claim that one can be the subject of a misfortune *only* if one exists *at the time* the misfortune occurs. I have been rejecting the idea that it is necessary for one to exist at the time the misfortune happens for it to be bad for him. But I am certainly not rejecting the idea that one must at least have existed at some point for a misfortune to be bad for them. Thus, while I reject the strict understanding of the Existence Requirement, I endorse a weaker (and indeed, almost trivial) formulation such as the following. A person can be the subject of some misfortune only if she exists at the time the misfortune occurs or has existed at some point – call this the Weak Existence Requirement. Once again, it is true that in order for me to suffer a stomach ache I need to exist at the time it occurs. But this is only one type of misfortune. Other types, such as losing my legacy after death and having my previous efforts rendered meaningless, do not require my existing at the time they happen precisely because of their retroactive nature. In the case of potential but yet non-existing people, neither the Weak Existence Requirement, nor the Principle of Retroactive Effects can apply. Therefore whatever badness is involved in death, potential people cannot suffer it.

### **3.2. What Is the Badness of Death?**

Once we accept that death can be bad for the person who dies, we can turn our attention to the question of what exactly it is that makes death bad for us. We have already seen in the previous chapter that death is not bad intrinsically, like pain and suffering are.

Instead, its badness is privative. Thus our task is to identify what it is that death deprives us of. For this purpose it seems that the best way to proceed is by starting with our commonsense beliefs about what we stand to lose through death in various hypothetical scenarios, and, from these, try to abstract the general elements that make up the badness of death.<sup>43</sup>

### 3.2.1. *Deprivation of future goods*

We generally tend to think that the earlier someone's death occurs, the worse it is for them. For example, it is generally accepted that it is worse to die at the age of 30 rather than at 70, and it would be even worse to die at 15. This suggests that a death is worse the more years of life it excludes. However, we want to also account for cases in which life would be very bad, perhaps even so bad that it would not be worth living. Surely one's death would not be so bad, or would not be bad at all, if the alternative would be a life of continuous pain, or more generally speaking a life devoid of value. This means that death is only bad when it curtails a life that would have been a good one to lead; in other words, death is worse the more good it excludes, broadly understood as whatever makes life worth living. It seems like we can already identify the first general element of the badness of death, namely the deprivation of future goods that one would have enjoyed had death not occurred.

Indeed, many authors have put forth accounts of the badness of death that focus around the loss of future goods. Notably, Thomas Nagel suggests that "[d]eath, no matter how inevitable, is an abrupt cancellation of indefinitely extensive possible goods.... If there is no limit to the amount of life that it would be good to have, then it may be that a bad end is in store for us all."<sup>44</sup> Nagel seems to suggest that insofar as we can imagine living forever death is bad for all of us no matter when it happens.

It is important to note that I, together with McMahan, consider that what we should

---

<sup>43</sup> I remain agnostic with regard to whether life itself, irrespective of the experiences it contains, has intrinsic value or is neutral, as the thesis does not depend on it.

<sup>44</sup> Thomas Nagel, "Death," in *Mortal Questions* (Cambridge: Cambridge University Press, 1979), 4.

have in mind for the purposes of comparison is not the possibility that death could have never occurred, but the possibility that a particular “token”<sup>45</sup> of death could not have occurred. If it were the “type” of the occurrence that was bad for us, i.e. death in general, we would have to say that death, no matter when it happens, is *equally* bad for all of us since it seems to deprive us of the infinite amount of good that we would have enjoyed had we been immortal. And for this reason this kind of view could not support our belief that death is worse the earlier it occurs. Nagel’s view is that if we take the possibilities for future goods to be infinite, in the sense that they are imaginable, death is always bad for us.<sup>46</sup> Conversely, McMahan rightly contends that what counts is not what is imaginable, but what is actually likely to have happened:

...[T]he possibilities for good of which a person is deprived by death are limited by the fact that, had he not died when and how he did, he would have been condemned by his biology and circumstances to die within a certain limited period of time thereafter. The relevant alternative to death for purposes of comparison is not continuing to live indefinitely, or forever, but living on for a limited period of time and then dying of some other cause. ...[W]e measure the badness of death in terms of the quantity and quality of life that the victim would have enjoyed had he not died when and how he did.<sup>47</sup>

Moreover, on this kind of view how bad death is for someone is measured by comparing the value of the entire life the person would have had had death not occurred with the value of the (shorter) life they actually had. This is what McMahan calls the Life Comparative Account.<sup>48</sup> If we aim to evaluate how bad one’s death is we need to consider how much good she would have enjoyed had this token of death not occurred. Of course, this account raises a number of very important and complex questions. One that springs to mind immediately is of

<sup>45</sup> See McMahan, *The Ethics of Killing*, 104 on the “Token Comparison”.

<sup>46</sup> Nagel, “Death,” 7-10.

<sup>47</sup> McMahan, “Death,” 41.

<sup>48</sup> McMahan, *The Ethics of Killing*, 105. McMahan himself eventually rejects the Life Comparative Account in favour of what he terms the Time-Relative Interest Account. The latter account judges the badness of someone’s death by considering how much it matters for the person, at the time of her death, that she should continue to live (based on how strong her “prudential unity relations” are, namely to what extent the psychological relations that ground egoistic concern for her own future hold between her at the time of death and her future self, had she continued to live). The two accounts, although different, are closely related and yield similar judgments in most cases. The main argument of this thesis can go through even if one takes up the Time-Relative Interest Account instead of the Life Comparative Account.



epistemological nature: how could we possibly know how good a person's life would have been had she not died when she did? There is no way to completely get around this epistemological limitation, since clearly no one has access to the future, let alone to a counterfactual future. McMahan suggests that one way in which we can best approximate our evaluation is by taking into account the opportunities that death deprives us of, instead of the actual goods that one would have enjoyed had she not died when she did. Furthermore, our reasoning about curtailed opportunities should be probabilistic and it should discount for uncertainty.<sup>49</sup> For example, even if we have good reasons to suppose that the death of X has robbed him of a great many opportunities for a wonderful life (based on his qualities, on how his life went so far and on how life tends to unfold generally), we should still have a margin of error to account for the possibility that her life would not have been quite as good as it seems it would have been.

Another difficulty the Life Comparative Account runs into is how we are to formulate the counterfactual situation that we need for our evaluation, namely what exactly do we exclude when we exclude one's cause of death? Consider the following case. A young man who suffers from Huntington's disease dies of pneumonia, which he developed as a complication of the first disease. If we want to establish how much of a misfortune his death was for him, it will make a considerable difference to our judgment what we take the cause of his death to be. If we consider this token of death to be caused by the pneumonia, the relevant counterfactual situation would be that in which he did not develop this complication and lived to see the Huntington's disease progress differently. He would perhaps have died one year later due to a different complication, or would maybe have lived for many more years, but would have descended into dementia. Alternatively, we can take the cause of his death to be the Huntington's, in which case the counterfactual would be to live a long, healthy life.

---

<sup>49</sup> *Idem*, 107.

Clearly, if we take the cause of his death to be the Huntington's itself, his death is a much greater misfortune for him than it is if we take the cause of his death to be the pneumonia, because in the first case he would have enjoyed many more years of a good life. McMahan calls this the Metaphysical Problem – “the problem of determining which of the many different ways in which a person might not have died when he did is the appropriate one for the purpose of comparison.”<sup>50</sup> For lack of space, I can only gesture toward the complex answer that McMahan proposes, namely that we should understand one's cause of death as “the entire causal sequence of which the immediate cause of death is a part.”<sup>51</sup> In the above mentioned case, then, the right counterfactual for the purposes of comparison would be the one in which the man would have lived a life free from Huntington's disease.

We have started this section on the assumption that the Life Comparative Account seems to support one of our deep-seated intuitions about the badness of death, namely that death is worse the earlier in one's life it occurs, and it is clearly true that dying at the age of 20 is worse than dying at 80. However, there seems to be a limit to this line of reasoning, namely that it goes against what most of us think about the death of fetuses and infants. The Life Comparative Account would suggest that the death of a late-term fetus or that of a one week-old infant is worse than the death of an 18 year-old because the fetus and the infant would lose more opportunities than the 18 year-old would, everything else being equal. But most of us believe the opposite: death is worse for a teenager than it is for a fetus or a newborn. It seems that we might need another element to account for our intuition with regard to fetuses and infants, and this element is the degree of personhood.

---

<sup>50</sup> *Idem*, p. 109.

<sup>51</sup> McMahan, “Death,” 47. See also his *The Ethics of Killing*, 107-17. Exploring the Metaphysical Problem then leads to the problem of overdetermination of causes of death, which occurs in cases where, had a person not died when and how she did, something else would have killed her immediately after. See McMahan, *The Ethics of Killing*, 117-27.

### 3.2.2. Degree of personhood

One possible explanation of our intuition that the death of a fetus (a miscarriage, for example) is not as bad as the death of a teenager is that the fetus is not yet a person in the relevant sense. Here too I will draw on McMahan's account of personhood and personal identity, which in turn is inspired by Derek Parfit's work.<sup>52</sup>

It seems that a necessary condition for someone to suffer whatever we take the badness of death to be, where death is defined as someone's ceasing to exist, they need to exist or to have existed at some point, according to the Weak Existence Requirement. The question then becomes: when do we start to exist in the relevant sense? Do we start to exist at conception or sometime during the pregnancy, or perhaps after we are born? Is there even an identifiable point when we "start" to exist, or is coming into existence a gradual process? The problem of the metaphysical status of embryos and fetuses is of course at the core of the longstanding debate about the morality of abortion. On most accounts, the permissibility of abortion seems to hinge on whether embryos and fetuses are persons just like we are, in which case abortion is impermissible in the same way that murder is.<sup>53</sup>

Here I take up Parfit's account of personhood and personal identity, a view that is quite common in current Western literature in some form or another. It is the view that we are fundamentally psychological beings. To be a person, "a being must be self-conscious, aware of its identity and its continued existence over time."<sup>54</sup> But what is involved in a person's continued existence over time? In asking this question we are already moving towards a definition of personal identity as well, since personal identity is concerned with what is necessary for "past me" to be one and the same person as "present me", or in other words

<sup>52</sup> Parfit, *Reasons and Persons*, 199-209.

<sup>53</sup> For an example of an account whereby the permissibility of abortion does not depend on the metaphysical status of the embryo or fetus, see Judith Jarvis Thomson, "In Defense of Abortion," *Philosophy and Public Affairs* 1 (1971), 47-66. Thomson argues that even if the fetus is a person and has the right to life, it does not follow that it has the right to the mother's body for its sustenance while it is still in the womb.

<sup>54</sup> Parfit, *Reasons and Persons*, 202.

what is necessary for me to have a continued existence over time.

Parfit suggests that there are two necessary criteria that must be met for personal identity. One is physical: it is necessary that enough of Y's brain continue to exist, and to now be X's brain, for Y and X to be the same person. The second condition is psychological: the criterion for personal identity is psychological continuity, and psychological continuity is defined as overlapping chains of strong psychological connections.<sup>55</sup> Examples of psychological connections are the relation between an experience and the memory of it, the relation between a desire and its later satisfaction or frustration, an intention and the later act that carries it out, and even the mere persistence over time of a belief, a desire, or any other psychological feature in someone.<sup>56</sup> These connections may hold to a stronger or a lesser degree. A person is strongly psychologically connected to her past self if, as Parfit specifies, at least half the number of connections that hold every day in the life of nearly every actual person hold in her case over any day.<sup>57</sup> For example, if we take any day in my life, I exhibit at least half of the connections that more or less everyone exhibits every day: I have memories of actions that I undertook, I see my desires being fulfilled or frustrated, I act upon my previous intentions and so on.

But connectedness in itself does not seem apt to be the condition for personal identity because it is not transitive like personal identity is. Personal identity is transitive in the following way: if at this moment I am one and the same person I was at age 18, and at age 18 I was one and the same person I was at age 10, it follows that I am now one and the same person I was at age 10. But psychological connections, no matter how strong, are not transitive in this way. If I have strong psychological connections with the person I was at age 18, and the person I was at age 18 has strong connections with the person I was at age 10, it does not follow that I now have strong connections with the person I was at age 10. I lack

---

<sup>55</sup> *Idem*, 207.

<sup>56</sup> *Idem*, 205.

<sup>57</sup> *Idem*, 206.

memory of most of the things I did when I was 10, most of the desires that I had then I do not continue to have now and do not care about fulfilling them, and I also do not have the same goals and intentions I did then. But surely I am still the same person. Parfit's explanation would be that I continue to be the same person over time because of overlapping chains of strong psychological connections even though I do not have direct connections to myself when I was 10 anymore. Thus even though the person I am now is not directly connected to the person I was at age 10, I am psychologically continuous with that person through a long chain of overlapping connections that go all the way back to that age.<sup>58</sup>

Now this account of the nature of personhood and personal identity, or something like it, explains why we think the death of an embryo or an early-term fetus is less bad than death at age 20. It is because one is not a person like you and I are when they are an embryo or a fetus or a newborn. The people they would become do not yet exist because they do not (and cannot, for their brain is not fully developed yet) support the psychological features that define a person. Coming into existence, then, is a matter of degree. As our brain develops and we acquire the necessary mental capacities that define a person, our status gradually changes from being just an organism to being a person. The embryo, the fetus and perhaps the newborn are merely potential (or developing) persons.

Thus, if a necessary condition for someone to suffer a misfortune such as death is that they exist or have existed at some point, and if an embryo or an early-term fetus does not yet exist in the relevant sense (is not possessed of personhood), its death cannot be bad for it. One might wonder why it is not enough that there exists an organism which will soon support the

---

<sup>58</sup> For personal identity it is necessary that the physical criterion take a "non-branching form," namely the brain must not be divided between two people such that each of them gets enough of the original person's brain. Branching would allow for personal survival (in the sense that the resulting two people would still be psychologically connected and continuous with the original person), but not personal identity as such, since it cannot be that both resulting people are truly one and the same as the original person, nor do we have any reason to think that one of them, but not the other, is one and the same as the original person. But at the same time the original person did not cease to exist either. Parfit calls what would happen in a brain division case personal survival, whereas for personal identity as we commonly understand it the non-branching form of the physical criterion needs to hold. See Parfit, *Reasons and Persons*, ch. 12.

mental capacities necessary for the development of a person, for death to be bad for it. The answer is that we are trying to discern whether death is bad in person-affecting terms, i.e. if death is bad for this particular person. So far we have seen that the badness of death is a function of the future goods that the person would have enjoyed had she not died when and how she did. But a fetus is not yet a person, so there is no psychological continuity between this fetus and the future person who would have enjoyed the goods had the fetus not died. There is no identifiable subject who would be the benefited party. This is not the case with the death of a 20 year-old. Had this young person not died when and how she did, there would be an identifiable person who would enjoy the goods of life, and this person would be psychologically continuous with the one who died – they would be one and the same person. By contrast, the person who would enjoy the future life had a fetus not died when and how it did would not be one and the same person as the fetus, because the fetus is not yet a full-fledged person. The fetus is arguably still an organism, and we are not numerically identical with our organisms.<sup>59</sup> Furthermore, one might think that a one week-old or even a one month-old infant is still not a full-fledged person, and hence not identical with the person who will enjoy the future goods that life would contain. This helps explain our intuition that the death of a teenager is worse than the death of an infant.

### 3.2.3. *Loss of investment*

Even if one thinks that an infant is a full-fledged person, there are still some other elements that normally make death bad for an adult, but do not apply to an infant. We have already touched upon one such important element in our discussion about the Principle of Retroactive Effects. We have seen that one reason for which death is bad is that it can render one's efforts, sacrifices and plans meaningless if one dies before getting the chance to complete them or enjoy the rewards of their strife. This suggests that one's investment in their

---

<sup>59</sup> See McMahan, *The Ethics of Killing*, 24-39 for the full explanation of why we are not numerically identical with our organisms (a person is essentially more than her organism).

future is important in evaluating the badness of their death. As McMahan succinctly suggests, “death is worse to the extent that it frustrates the efforts that a person has invested in preparing for the future that the death prevents.”<sup>60</sup> Since this concerns efforts that are consciously undertaken and goal-directed, it is clear that an infant cannot suffer this kind of badness through death.

A closely related reason for which death can be bad for someone is that it frustrated the victim’s desires. Think of a young man who wants to provide his family with a home and spends his free time and his money building it from the ground up, but dies before it is finished and the family never gets to live there. The man’s death is bad for him not only because it turns out his efforts turned out to be pointless, but also because he never got to satisfy one of his most important desires. Once again, it seems that an infant’s death is not so bad in comparison as it does not have any desires that death can frustrate.

To summarize, we have seen that what is bad about death is that it deprives people of the good that their life would have contained had they not died when and how they did. But before the lost future goods can count towards this evaluation, one must first have existed, i.e. one must have been a full-fledged person to whom those future goods would have been connected. Embryos, fetuses and perhaps even infants are not full-fledged persons and the lost goods belong only partially to them. Hence their death is less bad than the death of an older child, or a teenager, or an adult. There are some other reasons for which this is so, namely that death is bad for people because it can retroactively affect the value of their past efforts and frustrate their desires for the future. This account explains not only why the death of an embryo, a fetus, or an infant is less bad than that of a 20 year-old, but also why death at age 20 is worse than at age 80. This is because the amount of future goods that the 20 year-old would lose is much higher than the amount of goods the 80 year-old can still expect from a

---

<sup>60</sup> *Idem*, 177.

normal life. Furthermore, the 80 year-old has, on average, seen more of his investments in the future pay off than the 20 year-old has, and more of his desires satisfied.



## Chapter 4

### What Is Wrong about Starting A Life that Is Not Worth Starting?

At this point I hope to have established that there is a rational reason for which we can judge a bad condition to be bad enough so as to make life not worth starting, but not necessarily bad enough to make life not worth continuing once it has started. It is that in present-life cases the prospect of living in a particular condition needs to be even worse than the prospect of suffering the badness of death. Potential, non-existing people cannot suffer the badness of death, however, which explains the lower threshold for judging their life not worth starting. But is the distinction between future-life and present-life judgments useful for overcoming the non-identity paradox in person-affecting terms? One needs to show what the harm or the wrong is in having one's life started even though it was not worth starting, given that it was nevertheless worth continuing. Why should it matter for our judgments about which people to bring into the world that their lives would not be worth starting, if they would be worth continuing? In this final chapter I make some suggestions about how the distinction might take the non-identity debate forward.

It will be helpful to remind ourselves of the argument that we need to rebut if we are to solve the non-identity problem in person-affecting terms. Using as a case study the story of Wilma who will conceive and give birth to a blind child called Pebbles instead of taking a pill for two months and conceiving a different, sighted child, David Boonin lays out the argument thus.

P1: Wilma's act of conceiving now rather than taking a pill once a day for two months before conceiving does not make Pebbles worse off than she would otherwise have been.

P2: If P's act harms Q, then P's act makes Q worse off than Q would have been had P not done the act.

C1: Wilma's act of conceiving now rather than taking a pill once every two months before conceiving does not harm Pebbles.

P3: Wilma's act of conceiving now rather than taking a pill once every two months before conceiving does not harm anyone else.

C2: Wilma's act of conceiving Pebbles does not harm anyone.

P4: If P's act does not harm Q, then P's act does not wrong Q.

C3: Wilma's act of conceiving Pebbles does not wrong anyone.  
P5: If P's act does not wrong anyone, then P's act is not wrong.  
C4: Wilma's act of conceiving Pebbles is not morally wrong.<sup>61</sup>

Recognizing that P4 is the most vulnerable of the premises, as it is widely accepted that one can be wronged without being harmed, Boonin revises P4 as follows.

P4'': If P's act does not harm Q *and* P's act is not done with the intent to harm Q *and* is not done with extreme indifference to whether or not it harms Q, and if P's act does not violate Q's rights, then P's act does not wrong Q.<sup>62</sup>

It is this premise that we can seek to reject in light of the present-life versus future-life judgments distinction. Wilma's act does not harm Pebbles because it does not make her worse off than she would have been, since she could not have existed at all had Wilma not thrown the pills away. We can also assume, as Boonin suggests, that Wilma does not forgo the medication in order to harm Pebbles or with extreme indifference to whether or not it harms her. He then goes on to reject some of the arguments that have been made to the effect that Wilma violates Pebbles' rights. He contends that it is hard to identify what actual right Pebbles could have that Wilma is infringing by throwing away the pills, given that Pebbles is merely a potential person. He can only see one such right that a potential person could have, namely the right not to be conceived into an actual person.<sup>63</sup> But he believes there would be grounds for attributing such a right to potential people only if their life would foreseeably not be worth living. But Pebbles is not in such a position, so there is no right we could conceive of that Wilma would be violating.

This seems to be the place where Benatar's distinction could make a difference. Boonin clearly had in mind a right not to be conceived when one's life would foreseeably not be

---

<sup>61</sup> Boonin, "How to Solve," 129-31.

<sup>62</sup> *Idem*, 139.

<sup>63</sup> *Idem*, 140.

worth living, in the sense of continuing. But one could argue that there is also a right not to be brought into existence when one's life would not be worth starting.

I believe there are reasons to ground such a right. Most of us think that it is regrettable when someone has had a bad start in life because their opportunities for well-being have been severely restricted from the outset. If we think it is regrettable to have bad luck in the natural lottery, we will also find it reasonable to think that knowingly causing someone to exist in a bad condition that seriously limits their possibilities for well-being is wrong.

If one can imagine Pebbles' life as a chess game, for which she gets only one opportunity to play (and it is either this game, or nothing), Wilma's throwing away the pills is equivalent to her knowingly letting Pebbles start a game that is manipulated from the beginning to Pebbles' disadvantage. Upon finding herself thrown into the game, Pebbles can choose to leave it by paying a prohibitively high price, or try to make the best out of it, play as best as she can and find satisfaction in it. But even when she chooses the latter option, and even when she does reach certain levels of well-being along the way, most of us would still say that the position Wilma has put Pebbles in – either play this rigged game and try to make the best out of it, or pay an exorbitant price to leave it – is in some sense wronging Pebbles.

The idea is that knowingly placing someone in a difficult situation that irreversibly limits their range of opportunities for well-being from the outset, even if there are possibilities to achieve some levels of well-being within that situation, is wronging them. We can cash out this idea in what Ronald Dworkin calls the “parameters” for someone's possibilities for living a good life. In *Sovereign Virtue*,<sup>64</sup> he takes up the question of what a good (or successful) life is in order to show that liberalism, and in particular his conception of liberal equality, does not obviate or threaten or neglect the good in people's lives, as so many critics of liberalism have

---

<sup>64</sup> Ronald Dworkin, *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, Mass.: Harvard University Press, 2002).

charged.<sup>65</sup> He adheres to the “model of challenge”<sup>66</sup> of what a good life is, which, in a nutshell, holds that “life is *itself* a performance that demands skill, that it is the most comprehensive and important challenge we face, and that our critical interests consist in the achievements, events and experiences that mean that we have met the challenge well.”<sup>67</sup> By critical interests he is referring to those interests that are important regardless of whether we happen to want them. For example, having healthy familial relationships contributes to our well-being even if we do not happen to care about them. Critical interests thus contribute to what he calls “critical well-being”, as opposed to volitional interests, which comprise our desires and our self-chosen pursuits, the satisfaction of which count for our “volitional well-being”. These two kinds of interests overlap in many cases, as, on the one hand, often times we want what is in our critical interest, and on the other hand having some degree of success in obtaining what we happen to want is part of our critical interest. But when the critical interests conflict with the volitional ones, Dworkin stipulates that the critical interests should take precedence if we want to live a good life.<sup>68</sup>

Dworkin’s view of a good life as being, broadly speaking, “the exercise of skill in the face of challenge”<sup>69</sup> becomes directly relevant to our discussion when he introduces the concepts of limitations and parameters of a good life. Opposing the views that would advocate one objective, universal standard that defines the good life, the “challenge model” suggests that the standard of a good life is indexed to one’s particular circumstances which define the challenge she must face in order to have a good life. As Dworkin suggests, just like the challenge of aesthetic performance in painting was different for a Renaissance painter than it was for a 20<sup>th</sup> Century artist, so too all the complex circumstances in which we happen to find ourselves provide the framework in which we form and pursue our own conception of

---

<sup>65</sup> Dworkin, *Sovereign Virtue*, 237-38.

<sup>66</sup> *Idem*, 253.

<sup>67</sup> *Ibidem*.

<sup>68</sup> *Idem*, 244.

<sup>69</sup> *Idem*, 255.

the good life. Some of these circumstances, however, act as limitations for how good our lives will be, and some act as parameters. Circumstances including the age in which we live, the political framework, our culture, nationality, gender, health, social status and so on can act either as a limitation or as a parameter for what a good life can be for us. Dworkin maintains that for most of these each of us discriminates automatically which is a parameter and which is a limitation since what counts as one or the other may vary from case to case. I may think that not being athletic is a limitation for me insofar as my idea of the good life is to become a world class athlete. Alternatively, I could consider my not being athletic as a parameter for my well-being in that becoming a world class athlete cannot feature among the ideals towards which I can direct my life, and there is nothing problematic about this, as there are so many other valuable ideals of the good life that I could achieve. Thus, Dworkin understands parameters to be those circumstances that fix the pool of what the challenges of my life might be and what the possibilities for a good life are for my particular situation without necessarily being regrettable.

However, he emphasizes that many parameters are *normative*, and is worth quoting in some detail:

[Normative parameters] define our ethical situation not in terms of our actual situation but in terms of our situation as we suppose it should be. Our lives may go badly not just because we are unwilling or unable properly to respond to the circumstances we have, but because we have the wrong circumstances. We do not even face the challenge we identify as the right one. Even if we do the best we can in the circumstances we do face, we do badly measuring our success against the chance we believe we ought to have been given, and it is the latter that defines a good life for us.<sup>70</sup>

Now we can perhaps better understand what is the wrong in the rigged chess game analogy and in the original case of Wilma. Even though Pebbles may perform well and rise to the challenge presented by the manipulated chess game, or to the challenge of living her life with a serious disability, these are not the right sort of challenges she should have been facing

---

<sup>70</sup> *Idem*, 262.

in the first place. Some of the circumstances that define one's start in life, particularly the ones we can directly affect as potential parents, should be judged as normative parameters, if not as limitations. Although I cannot give an account here of what the normative parameters for a life worth starting should be, it is plausible that a case can be made such that a certain threshold of bodily and cognitive integrity would be one of them.

Recall that Boonin's P4'' stated that if an agent P does not harm Q (i.e. make Q worse off than she would otherwise have been), and if P does not intend to harm Q, and if P does not act with extreme indifference to whether or not she harms Q, and if P does not violate Q's rights, then P does not wrong Q. An argument to the effect that one has the right not to be conceived when their life is not worth starting, even though it would be worth continuing, could be formulated by appealing to something like Dworkin's normative parameters for well-being. Even though this potential person has a fair amount of opportunities for well-being after their life has started, in the sense that they can make the best out of the challenges they face and have a life worth continuing, the parameters of the challenge that will define how well or badly their life can go are below some threshold of adequacy, and this fact may be enough to constitute a wrong.

It is worth noting at this point that by establishing the distinction between a life worth starting and a life worth continuing I do not mean to say there is no connection between how a life starts and how a life continues. This thesis has only been focusing on why it is that the same non-trivial disabling condition can be judged to make life not worth starting, but not necessarily not worth continuing as well, in order to make the logical space for arguing that never coming into existence *can* be preferable to a life worth living, as in continuing. This, however, does not mean there are no connections at all between how a life starts and how it continues. For our purposes it is relevant to note that when the literature generally talks about a life worth living, they mean a life that contains more units of good (or positive value

construed in a rather narrow way) than it does units of bad. But certain disabling conditions pervade all areas in one's life such that a utilitarian definition of a life worth living, i.e. one in which the good outweighs the bad, is not so straightforward. The intrinsic badness of many kinds of pain and suffering pervades what would otherwise be good in a person's life. The badness of an irreversible bad condition that one is born with cannot be separated from the other areas of her life: one's different aspects of life are interconnected and affect each other in meaningful ways. I cannot enjoy a beautiful sunset if I am in pain, or if I am depressed. I cannot enjoy the full range of parental gratifications if I am missing a limb.

This is not to say that actual people who are in these situations do not have a valuable life, or cannot reach high levels of well-being. It is just to say that one should not put another into that situation from the outset, especially given the following asymmetry: it is certain that this person will face considerable hardship due to their disability, but it is not certain how well they will be able to manage this hardship. Pebbles might acquire great chess playing skills due to her being forced into the game and she might learn to enjoy each move for what it's worth, but the opposite is also possible. What is certain is that she is put into a hostile position and forced to react one way or the other. This, I would say, is impermissible. Wilma, no matter how good a mother she intends to be, and how much care and resources she will surround Pebbles with, will not be able to take on Pebbles' hardship, which she herself has caused by bringing her into existence. Reasoning in a way similar to something like the precautionary (or prudence) principle,<sup>71</sup> I believe the morally right thing for Wilma to do is take the pills and forgo conceiving Pebbles in the first place.

To summarize the last points, given that (1) it is certain that the opportunities for well-being when one starts life with a non-trivial disability are considerably limited (i.e. they fall

---

<sup>71</sup> The idea of the precautionary principle is that when there is considerable uncertainty about the risk and extent of an action or policy's harmful consequences, the act or policy should not be undertaken. See Andrew Jordan and Timothy O'Riordan, "The precautionary principle: a legal and policy history," in *The precautionary principle: protecting public health, the environment and the future of our children*, ed. Marco Martuzzi and Joel A. Tickner (World Health Organization, 2004).

below some threshold of adequacy for the parameters of well-being that we think parents should guarantee their children, relative to their own context), and (2) that even the good in one's life may be negatively skewed by the pervasiveness of the bad condition one starts life with, and (3) that by the time one realizes the full meaning of their situation (perhaps during adolescence the earliest) the price to pay for "opting out" of life (the badness of death) is already too high, and (4) the uncertainty about the extent of a serious disability's consequences for one's life, one should have the right not to be brought into existence with a life not worth starting.

One could object that Pebbles' situation is not that different from any of ours. Any of us, even when we have a life worth starting, could face hardship during our lives, perhaps to an even greater extent than Pebbles. Our parents can never be certain of how much evil our life will contain, how it will affect us, how we will be able to manage it and so on. Why is it permissible for anyone to cause someone else to exist, but not for Wilma? I believe the fact that Wilma can be certain of some specific hardship that will befall Pebbles is a relevant difference here. Consider as an analogy the fact that it is not rational to lock yourself in the house due to the possibility that a brick might fall on your head and kill you if you step outside, but it is irrational to put yourself in situations where that is sure to happen, unless you want to commit suicide. Moreover, the permanent and irreversible nature of a disability like Pebbles' adds to the badness of Wilma's knowingly causing it by bringing her into existence.

A much more serious objection might be the following. If the Weak Existence Requirement is true, namely that one needs to exist at the time a misfortune occurs or at least to have existed at some point for something to be bad for them, which implies that a potential person cannot suffer the badness of death, why would we say that a potential person has rights that we can infringe? In other words, why does the Weak Existence Requirement not apply to



considerations about whether a potential person can have rights? This is indeed a serious problem and here I can only gesture towards a possible avenue for replying to it.

First of all, it is worth keeping in mind that the right I have tried to establish is of a special kind, not on a par with all the other rights persons have – it is the right not to be caused to exist in a particular condition, i.e. the right not to be turned into an actual person, which clearly only applies to potential people. You do not need to exist in the relevant sense (i.e. to be a full-fledged person) at the time the infringement of the right occurs, as the right itself precludes the responsible agent from making you into an actual person.

How then should we understand this right? Although I do not have a fully satisfying answer at this point, as this would perhaps require much more work than there is space for here, I am inclined to think that this right flows from the duties of procreators and the morality of procreation more generally. As agents who cause other people to exist, prospective parents should ensure that their offspring do not face especially difficult challenges relative to the context they are in (the times they are living in, the geographical location, the community they are in etc.). Surely the liberty to procreate is bound by special duties to the people that we cause to exist. The question is why should these duties include more than offering children a life that will be worth continuing. Peter Vallentyne, for example, believes that the special duties of procreators are limited to ensuring that their offspring's life prospects are not negative.<sup>72</sup> But there are several other authors who argue that ensuring that children's lives are worth living is too low a threshold for making procreation morally permissible. Laura M. Purdy maintains that parents should secure "normal health" and the prospects for a "minimally satisfying life" for their children.<sup>73</sup> Bonnie Steinbock and

---

<sup>72</sup> Peter Vallentyne, "Equality and the Duties of Procreators," in *The Moral and Political Status of Children*, ed. David Archard and Colin M. Macleod (New York: Oxford University Press, 2002).

<sup>73</sup> Laura M. Purdy, *Reproducing Persons: Issues in Feminist Bioethics* (Ithaca, NY: Cornell University Press, 1996), 45-46.

Ron McClamrock talk about ensuring a “decent chance for a good life”.<sup>74</sup> Perhaps the most interesting account is Seana Shiffrin’s. She holds that imposing a harm on someone in the course of conferring a benefit to them is not necessarily permissible. She suggests that harming someone is permissible in order to avoid a bigger harm that could befall them, but is not necessarily permissible in order to bestow a “pure” benefit to them, i.e. simply to add value to their lives, even if this benefit would outweigh the harm done.<sup>75</sup> Thus, the mere fact that a child might have enough good in their life to make it worth living might not justify forcing them to face life with a severe disability. In other words, even if the life prospects of the child are positive, the fact that it contains some important burdens may make procreation impermissible. Of course, this account operates with a different account of harm from the one I have worked with here (namely, one that does not imply a “worse than otherwise” relation), but something along these lines may be promising for our purposes as well.

Arguably there are reasons to think that procreative duties include more than just ensuring that the child has a life worth living (understood as continuing). Moreover, these duties are not impersonal. Although the persons to whom these duties are owed do not yet exist, we can understand them to be owed to whomever the children turn out to be. Wilma owes whoever she causes to exist, be it Pebbles or a different child, a certain standard of life. I have suggested that this standard take into account the appropriate parameters for the possibilities of well-being at the start of one’s life. If this is accepted, it follows that whoever turns out to be Wilma’s child (in our case, Pebbles) has the right to have their life started within the appropriate parameters. In other words, they have the right not to be caused to exist with a life not worth starting.

---

<sup>74</sup> Bonnie Steinbock and Ron McClamrock, “When is Birth Unfair to the Child,” in *Morals, Marriage, and Parenthood: An Introduction to Family Ethic*, ed. Laurence D. Houlgate (Belmont, Calif.: Wadsworth, 1999), 195.

<sup>75</sup> Seana Valentine Shiffrin, “Wrongful Life, Procreative Responsibility, and the Significance of Harm,” *Legal Theory* 5 (1999): 124.

## Conclusions

In this paper I have explored David Benatar's distinction between lives worth starting and lives worth continuing, and its relevance for solving the well-known non-identity problem. The paradox is that it seems that it is wrong to knowingly cause people to exist with severe disabilities even when their disabilities are a necessary condition of their existence, and even when their lives would still be worth living.

There seems to be a necessary contradiction in saying that non-existence can be preferable to a life worth living. Dissolving this paradox has been the main focus of this thesis. One can get past this problem either by explaining how knowingly causing people to exist with a non-trivial disability harms them, in the sense that it makes them worse off than they would have been had the act not been performed, or by saying that although it does not harm them, it nevertheless wrongs them in some way. David Benatar's view is that coming into existence is always a harm, no matter how good one's life is, which entails that coming into existence with a disability is a harm as well. Although I disagree with Benatar's general view, nor do I believe that the non-identity problem can be overcome by appealing to a harm-based argument (at least where harm is understood to imply a "worse off" relation), I have found the distinction that Benatar proposes between a life worth beginning and a life worth continuing to be relevant for potentially solving the non-identity problem by using it to develop a wrong-based argument. The idea of his distinction is that if a life is worth continuing, it does not follow that it is also worth starting. If this is true, the necessary logical paradox of saying that non-existence (more precisely, never coming into existence) is preferable to a life worth living, as in continuing, is overcome.

In chapter 2 I have engaged with Benatar's arguments for coming into existence's being a harm when life is not worth living, as well as when life is worth living. A key step in

arguing for the latter conclusion is introducing and establishing the distinction between lives worth starting and lives worth continuing. While I followed Benatar in saying that coming into existence is a harm when the life prospects would be negative, I have argued that the two grounds he mentions for his distinction (that it is a basic distinction that we make in trivial cases as well, and that it is explained by there being interests in existing in the case of lives worth continuing but not in the case of lives worth starting) are either unconvincing or are only part of the explanation.

In chapter 3 I have proposed a different reason to defend his distinction, namely that the threshold for judging a life worth ending is high because of the specific losses involved in death. In other words, life is worth ending when the prospect of continuing to live with a bad condition is even worse than the prospect of suffering the badness of death. This type of evil cannot befall non-existing people, and this is what explains the lower threshold for judging a life not worth starting in the first place. After having rejected the Epicurean argument according to which death can never be bad for the person who dies, I have argued for an account of the badness of death that comprises three elements: the degree of personhood, the deprivation of opportunities for enjoying the future goods of a continued life, and the loss of what one has already invested for their own future. Assuming that an embryo, a fetus, and perhaps even a newborn are not fully developed persons, such a view explains why the death of an embryo, a fetus, or a newborn is less bad than that of a 20 year-old, but also why death at age 20 is worse than at age 80.

Having supported Benatar's distinction by appealing to the badness of death, in the final chapter I have suggested a way in which it could contribute to building a person-affecting argument in favour of the conclusion that it is impermissible to knowingly cause severely disabled children to exist when one can avoid it. I have argued that there are grounds for the right not to be brought into existence with a life not worth starting for the following

cumulative reasons. (1) It is certain that the opportunities for well-being when one starts life with a non-trivial disability are considerably limited (i.e. they fall below some threshold of adequacy for the parameters of well-being that we think parents should guarantee their children, relative to their own context). (2) Even the good in one's life may be negatively skewed by the pervasiveness of the bad condition one starts life with. (3) By the time one realizes the full meaning of their situation (perhaps during adolescence the earliest) the price to pay for "opting out" of life (the badness of death) is already too high. (4) There is considerable uncertainty about the extent of a serious disability's consequences for one's life. For these reasons I believe we should have the right not to be caused to exist with a severe disability even if our lives would be worth continuing.

However, even if one is not persuaded by there being such a right, Benatar's distinction, justified by the appeal to the badness of death, could perhaps ground an impersonal solution to the non-identity problem. One could argue that the duty not to create a child with a life not worth starting is an impersonal duty that procreators have, without linking this duty to a particular actual or potential right bearer.

## List of References

- Adams, Robert M. "Existence, Self-Interest, and the Problem of Evil." *Nous* 13 (1979), 53-65.
- Benatar, David. *Better Never to Have Been: The Harm of Coming into Existence*. Oxford: Clarendon Press, 2006.
- Boonin, David. "How to Solve the Non-Identity Problem." *Public Affairs Quarterly* 22 (2008), 127-157.
- Boonin, David. *The Non-Identity Problem and the Ethics of Future People*. Oxford: Oxford University Press, 2014.
- Cohen, G. A. "Rescuing Conservatism." In *Reasons and Recognition: Essays on the Philosophy of T.M. Scanlon*, edited by R. Jay Wallace, Rahul Kumar, and Samuel Freeman. Oxford: Oxford University Press, 2011.
- DeGrazia, David. "Is it wrong to impose the harms of human life? A reply to Benatar," *Theoretical Medicine and Bioethics* 31 (2010), 317-331.
- Dworkin, Ronald. *Sovereign Virtue: The Theory and Practice of Equality*. Cambridge, Mass.: Harvard University Press, 2002.
- Epicurus. "Letter to Menoeceus." In *Letters, Principal Doctrines, and Vatican Sayings*, edited by Russel M. Geer. Indianapolis: Bobbs-Merrill, 1964.
- Feinberg, Joel. "Wrongful Life and the Counterfactual Element in Harming." *Freedom and Fulfilment*. Princeton: Princeton University Press, 1992.
- Harris, John. *Wonderwoman and Superman: The Ethics of Human Biotechnology*. Oxford: Oxford University Press, 1992.
- Heyd, David. *Genethics: Moral Issues in the Creation of People*. Berkeley: University of California Press, 1992.
- Jordan, Andrew and O'Riordan, Timothy. "The precautionary principle: a legal and policy history." In *The precautionary principle: protecting public health, the environment and the future of our children*, edited by Marco Martuzzi and Joel A. Tickner. World Health Organization, 2004.
- Kavka, Gregory. "The Paradox of Future Individuals," *Philosophy and Public Affairs* 11 (1982), 93-112.
- Lockwood, Michael. "Singer on Killing and the Preference for Life." Quoted in McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002.
- McMahan, Jeff. "Death and the Value of Life." *Ethics* 99 (1988), 32-61.

McMahan, Jeff. "Wrongful Life: Paradoxes in the Morality of Causing People to Exist," in *Rational Commitment and Social Justice*, edited by Jules L. Coleman and Christopher Morris. New York: Cambridge University Press, 1998.

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

Nagel, Thomas. "Death." in *Mortal Questions*. Cambridge: Cambridge University Press, 1979.

Parfit, Derek. "On Doing the Best for Our Children," in *Ethics and Population*, edited by Michael Bayles. Cambridge, MA: Schenkman, 1976.

Parfit, Derek, *Reasons and Persons*. Oxford: Oxford University Press, 1984.

Petersen, Thomas Søbirk. "Being worse off – but in comparison with what?: On the baseline problem of harm and the harm principle." *Res Publica* 20 (2014), 129-214.

Purdy, Laura M. *Reproducing Persons: Issues in Feminist Bioethics*. Ithaca, NY: Cornell University Press, 1996.

Raz, Joseph. *The Morality of Freedom*. New York: Oxford University Press, 1986.

Samuelson, William and Zeckhauser, Richard. "Status Quo Bias in Decision Making." *Journal of Risk and Uncertainty* 1 (1988), 7-59.

Schwartz, Thomas. "Obligations To Posterity." In *Obligations to Future Generations*, edited by Richard Sikora and Brian Barry. Philadelphia: Temple University Press, 1978.

Shiffrin, Seana Valentine. "Wrongful Life, Procreative Responsibility, and the Significance of Harm." *Legal Theory* 5 (1999), 117-148.

Steinbock, Bonnie, and McClamrock, Ron. "When is Birth Unfair to the Child." In *Morals, Marriage, and Parenthood: An Introduction to Family Ethics*, edited by Laurence D. Houlgate. Belmont, Calif.: Wadsworth, 1999.

Thaler, Richard. "Toward a Positive Theory of Consumer Choice." *Journal of Economic Behavior and Organization* 1 (1980), 39-60.

Thomson, Judith Jarvis. "In Defense of Abortion." *Philosophy and Public Affairs* 1 (1971), 47-66.

Tversky, Amos and Kahneman, Daniel. "Rational Choice and the Framing of Decisions," *The Journal of Business* 59 (1986), S251-S278.

Vallentyne, Peter. "Equality and the Duties of Procreators." In *The Moral and Political Status of Children*, edited by David Archard and Colin M. Macleod. New York: Oxford University Press, 2002.

Velleman, David J. "Well-Being and Time." Quoted in McMahan, Jeff. *The Ethics of Killing: Problems at the Margins of Life*. New York: Oxford University Press, 2002.