Title:
New Selves, New Morals: The Ethics of
Transformative Experience

Name:
Jean-Philippe Thomas

Institution:
UCL

Degree:
PhD in Philosophy

I, Jean-Philippe Thomas, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

<u>Abstract</u>

Some of our experiences radically change us. After losing a loved one, moving to a new county, or having a child, it is common to feel that you are a 'different person' to the person you were before this experience. Sometimes these experiences give us insights that we could not have had without going through them. Sometimes these experiences also radically change some of our core values, commitments, and desires. In the philosophical literature, these experiences are often referred to as 'transformative experiences'. This thesis is about the *ethics* of transformative experience — that is, the ethical issues arising in contexts in which people go through or are otherwise affected by transformative experiences.

The main goal of the thesis is to present a 'launchpad' for engaging with new issues in the ethics of transformative experience. I expand the ethics of transformative experience in two main ways. First, I provide new evidence for the idea that one of the main reasons why transformative experiences can lead to ethical issues is that they create a gap between a person's old self and their new self. Second, I show that the ethical issues arising in the context of transformative experiences are more varied than previously recognised. I develop three puzzles which show that, if we take transformative experiences seriously, we may be unable to properly regret some historical tragedies, unable to morally evaluate transformative experiences which we have not been through ourselves, and, under certain conditions, we may have to give up some of our long-term moral commitments. I end by highlighting the importance of a more general lesson, which is that transformative experiences often make it difficult to identify a single or appropriate subject of moral concern.

Impact Statement

The impact generated by this thesis can be divided into academic and non-academic impact.

The thesis' academic impact is three-fold.

First, my thesis contributes to the philosophical field of ethics by offering a 'launchpad' for thinking about new issues in the ethics of transformative experience. I do so by offering a complementary literature review of some less-discussed existing contributions to the ethics of transformative experience. In addition, I raise three new ethical puzzles about transformative experiences which have not been discussed in the literature before. I then argue that, based on the literature review and the three new puzzles, the main source of ethical problems in the context of transformative experiences is the separation of old and new selves created by these experiences. While this insight has been acknowledged in the literature, my three puzzles provide new evidence for how important it is.

Second, my thesis contributes to philosophical debates in ethics about how our moral attitudes and practices operate. The thesis argues that some of our common moral attitudes and practices, such as regret, moral evaluation or moral commitment, rely on the clear identification of a subject of moral concern. However, I argue that, in some transformative experience cases, this identification is difficult because these experiences separate our old selves from our new selves. As a result, we do not know who should be our subject of moral concern in these cases. Overall, we are left with the conclusion that many of our common moral attitudes and practices do not seem to apply in the way we intuitively wish.

Third, the thesis shows how the philosophical fields of metaphysics and ethics can be connected in a fruitful way. Similar to other debates, such as the debate around the non-identity problem, my thesis shows that the metaphysical transformation of individuals or their selves can have important ethical consequences.

Outside of academia, my thesis has some beneficial implications for public policy design and implementation.

For example, in Chapter Three, I highlight the fact that people who undergo transformative experiences, such as pregnancy or forced migration, acquire a unique kind of phenomenological knowledge that people who have not been through the same experiences do not possess. I then argue, based on the work of Fiona Woollard (2021, 2022), that our moral evaluations of transformative experiences need to take into account this unique phenomenological knowledge. On a practical level, this means that designing public policies which directly relate to transformative experiences such as pregnancy, abortion, forced migration, etc. should also take into account the perspective of people who have been through these experiences. In other words, public policy design should always be informed by the first-person perspective of people who possess unique insights into the relevant experiences.

TABLE OF CONTENTS

Introduction

Some of our experiences radically change us. After losing a loved one, moving to a new county, having our first child or falling in love, we often feel like we are a 'different person' than before these experiences. Sometimes we feel this way because these experiences give us insights that we could not have had without going through these experiences. Sometimes we also feel this way because these experiences strongly shape and change some of our core values, commitments, and desires.

In the philosophical literature, these experiences are often referred to as 'transformative experiences', because they radically transform what we know and what we are like. This thesis is about transformative experiences. More specifically, the thesis is about *the ethics* of transformative experiences — that is, the ethical issues arising in the context of people who go through or are otherwise affected by transformative experiences. The main goal of this brief introductory chapter is to set the stage for the rest of the thesis. I offer a quick overview of what this thesis' main argument is going to be and how the thesis is structured. I keep the introduction deliberately brief as more setting up is to come in the next chapter. In the chapters following the introduction, I will discuss in greater detail what transformative experiences are and what ethical issues arise in relation to these experiences.

1. The Main Argument

Most discussions of transformative experiences focus on the potential issues these experiences raise for *decision theory*. I will discuss some of these decision-theoretic issues in the next chapter. The main goal of this thesis, though, is to help expand the discussion of transformative experiences into the domain of ethics. I will do so by providing a 'launchpad' for engaging with new ethical issues that have not yet been discussed in the context of transformative experiences. More specifically, I expand the ethics of transformative experiences in two main ways. First, I offer new

evidence in support of the view that transformative experiences give rise to ethical issues because they create a 'gap' between our old and new selves. As a result of this gap, these experiences make it difficult for us to identify who our subject of moral concern should be. This idea is already present in some of the literature[1], but my thesis shows just how important it is. Second, I demonstrate that the ethical issues to which transformative experiences give rise are more varied in kind than has previously been recognised. I do this by showing that transformative experiences give rise to ethical issues about *backward-looking* attitudes, *present* attitudes, and *forward-looking* attitudes.

The main value of this thesis does not lie in establishing any particular puzzle about moral attitudes and practices and transformative experiences. Instead, the thesis' main value lies in its highlighting the relevance of transformative experiences to ethics by providing a launchpad for thinking about new ethical issues arising in the context of these experiences. The puzzles I will raise about some of our moral attitudes and practices contribute to this larger goal.

2. Thesis Structure

The thesis is made up of five chapters.

The first chapter reviews and critically assesses some of the existing contributions to the ethics of transformative experience. The main insight that I derive from the contributions I discuss in this chapter is that many moral issues arise in the context of transformative experiences because these experiences create a 'gap' between one's 'old' and 'new' selves.

While literature reviews can sometimes be unnecessary add-ons to a thesis, the review does play an important role in this thesis. The first reason why this thesis includes a literature review is

---

[1] See, for example, Farbod Akhlaghi (2023), 'Transformative Experience and the Right to Revelatory Autonomy', *Analysis* 20 (20): 3-12.

that there are very few such reviews available. Indeed, when I began this project, there were none at all. Since then, partial overviews have appeared in both the *Stanford Encyclopaedia of Philosophy*[2] and *Philosophy Compass*.[3] I think the literature review that I offer here complements those, by focusing on material to which they give less attention. In addition, given the relatively small size of the ethics of transformative experience literature, it seems useful to have a review which gives the reader a good grounding in the literature in a relatively short space.[4]

The second, third and fourth chapters of the thesis each present a new puzzle about the ethics of transformative experience which has not been discussed before. Each of the puzzles has a specific temporal focus.

---

[2] Rebecca Chan (2023), 'Transformative Experience' T*he Stanford Encyclopaedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), forthcoming URL = <https://plato.stanford.edu/archives/spr2024/entries/transformative-experience/>.

[3] Daniel Villiger (2024), 'Transformative Experience', *Philosophy Compass* 19 (6): 1-12.

[4] Some of the material that is less discussed (or sometimes not at all) in existing reviews but which I analyse in some detail includes, for example, Nilanjan Das and L.A. Paul (2020), 'Transformative choice and the non-identity problem', in Sauchelli, Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (London and New York: Routledge), 187-208; Dana Sarah Howard (2015), 'Transforming Others: On the Limits of "You'll Be Glad I Did It Reasoning"', *Res Philosophica* 92 (2): 341-370; Amia Srinivasan (2015), 'All the Same', *The Times Literary Supplement*, no. 5854, 8; Fiona Woollard (2021), 'Mother Knows Best: Pregnancy, Applied Ethics, and Epistemically Transformative Experiences', *Journal of Applied Philosophy* 38 (1): 155-171; Fiona Woollard (2022), 'Your Mother Should Know: Pregnancy, the Ethics of Abortion, and Knowledge through Acquaintance of Moral Value', *Pacific Philosophical Quarterly* 103 (3): 471-492.

The second chapter presents a puzzle about our *backward*-looking attitude of moral regret. The puzzle starts with the observation that living through a historical tragedy, such as the Holocaust or the slave trade, is a transformative experience. By living through such an experience, your old 'self' before the experience is replaced by a new 'self' after the experience. I then argue that, when we wish to regret a historical tragedy, we also wish to regret it on behalf of the new self that emerges from such tragedies. However, I argue that we seem unable to do so, because, at least in typical cases, the 'new' self was not made worse off by the tragedy. The 'new' self usually has a life worth living and exists only because of the tragedy. Overall, we are left with the highly problematic conclusion that we cannot regret some historical tragedies, at least not in terms of how they affected the surviving victim's 'new' self.

The third chapter presents a puzzle about how we can *presently* morally assess a transformative experience which we have not been through ourselves. The chapter starts with Fiona Woollard's (2021, 2022) insight that, from a moral point of view, it is essential that our moral evaluations of transformative experiences, such as pregnancy and abortion, take into account the unique phenomenological insights of people who have been through these experiences. I then build on and generalise Woollard's insight by arguing that many of us do not actually go through these types of transformative experiences. In addition, there is not always a way of obtaining these insights other than going through the relevant transformative experiences. So, counter-intuitively, it seems that we cannot morally assess many transformative experiences that we have not been through ourselves.

The fourth chapter presents a puzzle about our long-term *future* moral commitments. The chapter starts with the epistemological view that, if we know that we will come to have a particular belief in the future, and we know that this belief will be formed rationally and on the basis of veridical evidence, then we should adopt this belief now. I then argue that moral commitments are also a form of belief and so, if we know that we will come to have particular moral commitments in the future, and we know that these commitments will be formed rationally and on the basis of

veridical evidence, then we should adopt them now. I then extend this argument by suggesting that, for certain transformative experiences, while we do not know what our future moral commitments will be after these experiences, we can nonetheless be confident that these commitments will be formed under the right epistemological conditions. As a result, we are left in a situation in which, from an epistemological point of view, we need to suspend our current long-term moral commitments. This outcome is strongly counterintuitive as many of us feel deeply attached to our moral commitments and take them to be an important part of our lives.

The fifth and final chapter concludes the thesis and pulls together the strings of the previous chapters. The main lesson I draw is that my three puzzles provide new evidence for the view that transformative experiences create a 'gap' between our old and new selves which can make it difficult to identify a proper subject of moral concern. While this lesson has been discussed in the literature before, I demonstrate its importance, by showing that the ethical puzzles to which it gives rise are many and varied.

In addition, I use the final chapter to compare my approach, and some of my arguments in this thesis, with a famous discussion by Derek Parfit (1984) in *Reasons and Persons*. Parfit's discussion focuses on the rationality and ethics of personal identity, and, like my arguments, it suggests that uncertainty about a person/self's identity reduces the importance of long-term moral commitments and potentially supports a more impersonal view of morality.

3. Thesis Style

I want to briefly address the issue of the style in which I have structured this thesis. All the chapters of this thesis, except for the concluding chapter, can be read as standalone papers. These chapters make standalone contributions to the existing ethics of transformative experience literature. As a result of this choice, each of these chapters contains its own introduction and presentation of Paul's concept of 'transformative experience'. This leads to a little repetition, but it also has some

benefits. Presenting Paul's concept in each of these chapters gives me the opportunity to highlight different aspects of her views that are relevant to each chapter.

4. Why Paul's 'Transformative Experience'?

Before moving on to the literature review in the next chapter, I want to anticipate and address the question of why this thesis focuses on L. A. Paul's concept of 'transformative experience' rather than on other accounts of transformation in philosophy.

It is true that 'transformation' as a philosophical concept is not exclusive to Paul. Discussion of rational choice-making under transformative conditions has been present in philosophy for a long time. For example, Blaise Pascal's famous 'wager' — in which Pascal tries to convince us that starting to believe in God is the rational choice to make — is an example in point (Pascal [1670], 2000; 460-467). Similarly, we might think that the emphasis that some existentialists put on making 'authentic choices' is also an instance of a discussion of choice-making under transformative conditions.[5]

In addition to some historical discussions, there are also other contemporary philosophical accounts of decision-making and transformation. These include, for example, accounts by Edna Ullmann-Margalit (2006), Agnes Callard (2018), Richard Pettigrew (2019) and Havi Carel and Ian James Kidd (2020). I will discuss some of these accounts in more detail later in the thesis. Having said that, I want to make two points about why I have focused this thesis on Paul's concept of 'transformative experience' rather than any of the other historical or contemporary accounts.

First, the contemporary ethical literature around transformation, although relatively small, has developed mainly around Paul's concepts. It therefore seems natural for me to focus this thesis, which is about the ethics of transformative experience, also on Paul's 'transformative experiences'. This will allow me to engage as directly as possible with existing contributions in the literature. So,

---

[5] See, for instance, Jean-Paul Sartre ([1943], 2017), *L'être et le néant: Essai d'ontologie phénoménologique* (Paris: Gallimard).

although the topic of transformation is not new or exclusive to Paul, recent engagement with the ethics of transformation has centred largely around Paul's work.

Second, it is worth noting that the success or value of this thesis does not depend on Paul's concept of transformation being the correct one. The thesis is not meant as a reflection on Paul's work exclusively. Instead, it is meant to be a discussion of the broader issue of the ethics of transformative experience for which Paul has provided useful philosophical vocabulary and concepts. So, as long as we believe that there are experiences which radically transform what we know and how we think about the world and ourselves, the main points of this thesis will remain relevant.

Chapter One:
Literature Review

Introduction

A lot of our decision-making in uncertain conditions seems to follow a similar pattern, at least when we are reasoning well. When we have several alternative options, we try to figure out what the possible outcomes of these options are, what each outcome would be like for us, and how likely each outcome is. For example, if we have to choose between carrying an umbrella or not carrying an umbrella for our hiking trip, then we try to take into account what it would be like to carry or not carry the umbrella given rain, what it would be like to carry or not carry the umbrella given no rain, and how likely it is that rain will come.

According to L.A. Paul (2014), this typical way of making decisions does not work in an important set of cases. On her view, the standard decision-theoretic model does not apply in many important life decisions — decisions in which we do not know what (some of) the relevant outcomes will be like nor how these outcomes will transform us. Paul calls cases of this kind 'transformative experiences' (Paul 2014, 15-16).

Traditionally, transformative experiences have been discussed in the context of epistemology and decision theory. Recently, however, some philosophers have emphasized the need for an *ethics* of transformative experience. Yoaav Isaacs (2020), for instance, writes that 'transformative experiences should not be taken to pose a problem for decision theory, but should instead be taken to pose *a topic for ethics*.' (my emphasis) (Isaacs 2020, 1082). My main goal in this dissertation is to help expand the discussion about transformative experiences into the field of ethics. This particular chapter will contribute to the literature by bringing together the main papers on the ethics of transformative experience in a systematic way for the first time.

There are two main reasons why I think it is important to include this kind of literature review in the dissertation. First, the ethics of transformative experience is a very recent field of

inquiry with comparatively few papers published on this topic. As a result, there is an opportunity to provide one of the first comprehensive overviews of the existing literature.[6] Second, given the recent and ongoing nature of this debate, a literature review will offer the reader a good grounding in this new field and thus help them see more clearly how my later contributions differ from existing ones.

I will proceed as follows. In the first part of the chapter, I will set the context of Paul's work on transformative experiences by discussing the traditional decision-theoretic model in more detail. In the second part of the chapter, I will present Paul's own view of the challenge posed to decision theory by transformative experiences. In the third part, I will clarify how I intend to use the concepts of 'person', 'self' and 'identity' in my discussion of transformative experiences. In the fourth part, I will briefly motivate the need for an *ethics* of transformative experience. In the fifth part, I will present four papers dealing explicitly with ethical questions arising in the context of transformative experiences. I will particularly focus on papers by Farbod Akhlaghi (2023), Nilanjan Das and L.A. Paul (2020) and Dana Howard (2015). Discussing these papers will allow me to bring out the general point that ethical questions often arise in this debate because transformative experiences radically separate our old selves from our new selves. In the final part of the chapter, I will discuss papers by Amia Srinivasan (2015) and Elizabeth Barnes (2015), which focus on *social and political questions* arising in the context of transformative experiences. This final discussion will allow me to emphasise some of the nuance to the claim that transformative experiences radically transform us.

---

[6] The only other existing overview is a set of paragraphs in the Stanford Encyclopeadia article by Rebecca Chan (2023) on 'Transformative Experience', published recently.

1. Contextualising Transformative Experience

Consider the following example of a simple decision situation often discussed in the literature.[7]

> *Rain*: You are about to leave the house and you need to decide whether you should take an umbrella with you or not. You know that, based on the latest meteorological report, there is a seventy percent chance that it will start to rain soon.

Many people would reason about this case in an informal way as follows. It is true that carrying an umbrella is somewhat cumbersome for me. However, I definitely do not want to get wet. Given that there is a high risk that I will become wet if I do not carry the umbrella with me, I prefer to carry the umbrella, because getting wet would be worse for me than carrying the umbrella (even if there will not be any rain after all).

I now want to dissect this informal way of reasoning a bit more. Cases such as *Rain* usually contain at least four key parameters, which are the options, the different states the world could be in, the probabilities of these different states, and the values of the possible outcomes for the relevant person. The situation can be represented by a simple table, in which we imagine the values that the relevant person assigns to the different possible outcomes and then represent these values with numbers:

	Rain (70%)	No Rain (30%)
Umbrella	90	85
No Umbrella	20	95

---

[7] See, for instance, Richard Bradley (2012) 'Decision Theory: A Formal Philosophical Introduction' in Vincent F. Hendriks and Sven Ove Hansson (eds), *Handbook of Formal Philosophy* (New York: Springer), pp.618-620.

In *Rain*, we have two options, which are either carrying the umbrella or not carrying the umbrella. These two options can lead to four potential outcomes: 1) You carry an umbrella and it rains. 2) You do not carry an umbrella and it rains. 3) You carry an umbrella and it does not rain. 4) You do not carry an umbrella and it does not rain. The different states the world could be in are the state where it rains or the state where it does not rain, and the probabilities of these different states are seventy percent and thirty percent, respectively.

In addition to the options, the different possible states of the world, and the different possible outcomes, the table shows the relevant person's preferences over the different outcomes, represented by a number between 0 for a very low preference and 100 for a very strong preference. For example, this person's strongest preference is for the outcome in which they do not carry an umbrella and it does not rain. In contrast, the outcome they prefer least is the outcome in which they do not carry an umbrella and it does rain.

In order to find out what it is rational for the person in *Rain* to do, we can make a basic calculation, often referred to as a calculation of the 'expected utility' of an option — that is, the expected value of the option for the person. Usually, an option's expected utility is calculated by taking the value of each possible outcome of the option, multiplying this value by the probability that this outcome will obtain, and then summing the results.

For example, in *Rain*, the expected utility of carrying the umbrella is (90 x 0.7) + (85 x 0.3) = 88.5, while the expected utility of not carrying the umbrella is (20 x 0.7) + (95 x 0.3) = 42.5. Based on this calculation, it seems that the rational action to take in *Rain* is to carry the umbrella. For the expected utility of carrying an umbrella is higher than the expected utility of not carrying an umbrella.

This first simple case also allows us to highlight a key feature about how we make many of our personal decisions. Our reasoning in these cases often seems to follow a pattern. We try to identify the possible outcomes of our decision and project what each outcome would be like for

us. Based on this projection, we try to identify which outcome we would prefer. This standard decision-theoretic model relies on a key assumption about first-personal experience. The model assumes that, for any personal decision, it is possible for us to assign a value to *what it would be like for us* to live through each potential outcome. As I will now explain, L. A. Paul thinks there are certain cases in which this key assumption is very problematic.

2. L.A. Paul on Transformative Experience

Consider the following case.

> *Charity Abroad*: You are considering joining a charity project in an underprivileged part of the world ravaged by political and military conflicts. Your life as a charity worker would be radically different from your life at home. You could no longer take basic necessities for granted and you would be directly confronted with the consequences of violent human conflict. Some of your friends who have already taken part in this project strongly recommend it. They claim that the project changed their lives for the better by showing them what truly matters in life and that they would never want to go back to their previous lives.

At first glance, this case might seem very similar to *Rain*. We might think that the case can be evaluated by simply calculating the expected utility of deciding to volunteer and comparing it to the expected utility of not deciding to volunteer.

However, on Paul's view, while the two cases share a similar structure, they also have one crucial difference. The difference between the two cases, according to Paul (2014), is that, unlike in *Rain*, you cannot assign sensible values to the possible outcomes in *Charity Abroad*. Specifically, Paul argues that you cannot know what value or preference you would attach to volunteering (see, for example, Paul 2014, 8-11).

The main reason for this, according to Paul, is that you cannot know in advance what it would be like for you to join this charity. Paul here distinguishes between what I will call 'phenomenological knowledge' and 'propositional knowledge' (Paul 2014, 12-15). On her view, you can have propositional knowledge about joining the charity—that is, you can know what the objective facts of joining the charity are (e.g. where you are going to live, what tasks your job will involve, etc.). However, according to Paul, you cannot have 'phenomenological' knowledge—that is, knowledge about what it would be like *for you* to join this charity— unless you have had this kind of experience before (Paul 2014, 12). On this view, joining the charity is such a radically different experience from not joining the charity at all that you cannot possibly project yourself into what it would be like for you to join this charity. And, according to Paul, if you cannot know in advance what it is phenomenologically like to join this charity, then you cannot know what value you would attach to this outcome, and therefore, in this case, the standard decision-theoretic model breaks down (Paul 2014, 13-14).

We might wonder why Paul insists so much on possessing phenomenological knowledge for making decisions in cases such as *Charity Abroad*. According to Paul, decisions in cases such as *Charity Abroad* are deeply personal and intimate. On this view, we care much more about these decisions because they affect our personal life and future much more than other more ordinary decisions. And, according to Paul, it is the personal nature of these decisions which makes phenomenological knowledge so important. On Paul's view, when we make decisions in these deeply personal cases, we should draw on the most personal and intimate source of knowledge we have, i.e. phenomenological knowledge. She argues that ignoring our first-personal perspective would be failing to do justice to the intimate and personal connection we have to these decisions (Paul 2014, 18; Paul 2015, 797-798).

Paul also argues that we cannot rely on other sources of information, such as testimony from others, to acquire the phenomenological knowledge that we need to make decisions in cases such as *Charity Abroad*. On Paul's view, other people might tell us what it was like for *them* to go

through a certain experience, but this is not sufficient to tell us what it would be like for *us* to go through the same experience (Paul 2014, 13).

According to Paul, the epistemically distinctive character makes *Charity Abroad* an 'epistemically transformative' experience (Paul 2014, 47). On her view, an experience is epistemically transformative if you can only acquire the phenomenological knowledge of what it is like by going through the experience. By going through the experience, Paul argues, you are epistemically transformed, i.e. you acquire phenomenological knowledge that you did not have access to before (Paul 2014, 8).

In addition, Paul argues that some experiences, such as *Charity Abroad*, can also be 'personally transformative' (Paul 2014, 3). On her view, an experience is 'personally transformative' if and only if it radically changes a person's core set of values, desires, commitments, etc (Paul 2014, 16-18). For example, if you join the charity, it is highly likely that some of your core values, desires, etc. will be transformed. For instance, by joining the charity and being directly confronted with people's daily struggle to escape violence and secure basic necessities, you might question your own lifestyle choices and the excessive value you attach to material goods.

The personally transformative nature of certain experiences also affects the standard decision-theoretic model. Recall that the model assumes that you can define which outcome you prefer by figuring out what it would be like for you to live through each outcome. However, if an experience is personally transformative, it will also transform your personal preferences. As a result, Paul (2014) argues, you cannot definitively say which outcome you would prefer to get because, after the experience, your preferences might be radically different (Paul 2014, 16-17). It therefore seems again that the standard decision-theoretic model cannot deal with many important life decisions.

In this section, I have set out the context in which we should understand Paul's work on transformative experience. Paul's work can be understood as a challenge to the traditional decision-theoretic model. By highlighting the epistemically and personally transformative nature of certain

important decisions within our lives, Paul seems to show that many of the most important decisions in our lives cannot be made by simply maximising expected utility.

In the literature, several philosophers have responded to Paul's challenge to traditional decision theory. Some of them accept Paul's conclusion that we cannot make rational decisions in transformative experiences but argue that we can nonetheless make 'reasonable' decisions (Ullmann-Margalit 2006, 169). Most philosophers, however, want to avoid Paul's conclusion by rejecting one or more of her key premises. Meena Krishnamurthy (2015) rejects the premise that you can know the value of a transformative experience for you only if you know what this experience is like. Elizabeth Harman (2015) and Richard Pettigrew (2015, 2019) disagree with Paul's assumption that you can know what a transformative experience is like only by going through this experience and argue that we can be relatively confident what value the experiences will have for us once we have been through them.

3. Identity and Selves

Before moving on to the ethics of transformative experience, it is worth making explicit a distinction that underlies discussions of transformative experiences. After living through experiences such as *Charity Abroad*, we often say things such as 'My experience has made me a *different person*' or 'I am not *the same person* anymore'. However, it is not entirely clear what we mean by 'different' or 'same' and by 'person' in this context. What sense of identity are we employing in these cases? And what do we mean by a 'person' in these contexts? Getting clearer on these key terms will clarify the puzzles and discussions I intend to raise in the next chapters.

3.1. Numerical Identity

Suppose you have lived through the experience described in *Charity Abroad*. Your experience has radically changed you — that is, your worldview and some of your core values, desires and

commitments have changed. It therefore seems plausible to say that, in one sense, you are now a 'different' person. We can start by clarifying what we do *not* mean by 'different' in this context.

The most direct interpretation of expressions such as 'I am not the same person anymore' is that going through transformative experiences makes you into a numerically different person. In other words, on this interpretation, there are two distinct people, one existing at time t1 before the transformative experience and one existing at time t2 after the transformative experience.

However, this does not strike me as a plausible interpretation of transformative experience cases. I want to consider two arguments for why transformative experiences do not make us into a numerically distinct person.

First, transformative experiences tend to be gradual processes, i.e. they take place over extended periods of time (Das and Paul 2020, 195-202). Different views about numerical identity can use this insight to make the case why these experiences do not make us into a numerically distinct person.

For instance, on a psychological continuity view of numerical identity[8], we could argue that you remain the same person, metaphysically speaking, because you before working for the charity remain psychologically continuous (via memories for example) with you after working for the charity. If transformative experiences are gradual processes in which there is no break in psychological continuity between the person before the experience and the person after the experience, then we have no reason to believe that you become a numerically distinct person, on this view (Das and Paul 2020, 195-202).

---

[8] Derek Parfit (1984) has famously presented one version of the psychological continuity view. On this view, X is numerically identical to Y 'if and only if X today is psychologically continuous with Y, this continuity has the right kind of cause, and it has not taken a branching form' (Parfit 1984, 207).

Alternatively, on an animalist view of personal identity[9], we might say that you remain the same biological organism throughout the charity example, thus also remaining the same person, at least metaphysically speaking. Again, if transformative experiences are gradual processes in which there is no radical break in your continuity as a physical organism, then it seems, on the animalist view, that there is no reason to believe that you have become a numerically distinct person.

Second, suppose that transformative experiences do make us into a numerically different person. This would mean that the person before the experience goes out of existence. In other words, transformative experiences would be akin to 'death' for the person before the experience. However, even if some transformative experiences might be akin to death, others clearly are not. For instance, becoming a parent often counts as a personally transformative experience because it has the potential to radically change some of our core desires, projects, commitments, etc. Surely, however, we would not argue that the people before the birth of their child cease to exist after becoming parents. They do not die through this experience. Instead, they are profoundly transformed without being replaced numerically.

So, when discussing transformative experiences, usually we are not assuming a change in the numerical or metaphysical identity of persons. If we are not assuming a change in numerical identity in transformative experience cases, then how can we make sense of the idea that you feel like a 'different' person after the experience described in *Charity Abroad*? In the literature on transformative experience, several philosophers have argued that, in addition to numerical or metaphysical identity, we need to consider a person's 'ethical' identity. Often, in order to distinguish the two senses of identity, we talk about 'persons' when referring to a person's numerical identity and we talk about 'selves' when we refer to a person's ethical identity. The main idea behind this additional, ethical sense of identity is that we need a way of talking about identity

---

[9] E.T. Olson (1997) has defended an animalist account of personal identity in *Human Animal: Personal Identity Without Psychology* (New York: Oxford University Press).

and selves that goes beyond numerical identity and that allows us to illuminate practical issues relating to rationality and ethics.

3.2. Chan on 'Practical Identity'

Different philosophers have offered different ways of spelling out what exactly is meant by a person's self or 'ethical' identity. I will discuss two of them here that seem to me particularly interesting and important.

Rebecca Chan (2019) argues that, in addition to a numerical identity, a person also possesses a 'practical' identity (Chan 2019, 7). On Chan's view, there are certain features — such as our life goals, preferences or the way we perceive the world — which we take to be essential to who we are (Chan 2019, 9). For example, it might be that the self that existed before moving abroad to work for a charity had a strong preference for a financially comfortable life and a professional life project centred around financial success. The self took these preferences and this life project to be essential to who they were. According to Chan, it is not always clear exactly which features we take to be essential and each person can take different features to be essential (Chan 2019, 11).

In transformative experiences, a number of these features that we take to be essential change. For example, in *Charity Abroad*, it might be the case that, once you have undergone the relevant experience, your life project changes to a commitment to supporting economic development in poorer geographical areas. Chan argues that if, as in the case of a transformative experience, a sufficient number of these features change, then we cannot 'first-personally' identify anymore with the past life we were leading. In other words, we do not recognise the old life as *ours* anymore. On this view, we have become a different 'self', ethically speaking (Chan 2019, 7-11).

In order to explain the idea of first-personal identification, Chan draws on the work of, amongst others, David Velleman (1996). Velleman argues that we can distinguish between visualising what Napoleon saw when looking out at Austerlitz and imagining ourselves being

Napolean looking out at Austerlitz (Velleman 1996, 39-40). It is the latter type of 'imagining' that is involved in first-personal identification. If we are able to imagine or project what it is like to be, say, Napoleon, from the inside so to speak, then, we can first-personally identify with Napoleon. The same applies to our selves. If we cannot first-personally identify with our past or future self anymore, then, we are, ethically speaking, a different self (Chan 2019, 10-11).

It might be worth dwelling a little on Chan's claim that our practical identity is made up of features *we* take to be essential rather than of features held to be objectively essential for any person (Chan 2019, 11). Chan offers three reasons why she believes this to be an important insight.

First, focusing on what the agent takes to be essential features allows Chan to accommodate the diversity among agents. On Chan's view, it is just a fact that different agents take different features of their existence to be essential to who they are (Chan 2019, 11).

Second, it is important to remember that what makes an experience transformative is the fact that the relevant agent cannot first-personally identify with their previous or future self. All that matters, on this view of first-personal identification, is that the relevant agent is able to imagine losing this feature. If the agent cannot imagine losing this feature, then, on this view, it is practically essential. Given the importance attached to *the agent's* ability to imagine losing or not losing a feature, it seems natural, on Chan's view, to focus on features that the agent takes to be essential (Chan 2019, 11).

Third, Chan wishes to respond to the following objection. Surely, the objection goes, we can sometimes be mistaken about what we take to be our practically essential features. That is, it might be the case that the agent takes a feature to be essential which is not or that the agent does not recognise a feature which is essential as an essential feature. However, the objection goes, it is not clear that Chan can accommodate the possibility of making mistakes about our practically essential features. After all, on Chan's view, it seems that whatever feature we regard as practically essential is, in fact, practically essential, no matter whether we made a mistake or not (Chan 2019, 11).

Chan's reply to this objection is to argue that, even if we can make mistakes about what features are objectively essential, this does not affect her account. Despite the possibility of mistakes, we should still regard our practical identity as made up of features that *we take to be* essential (even if we might be mistaken about them). The main reason for this, on Chan's view, is that practical rationality strongly depends on what *we believe* to be the case (rather than on what is actually the case). Acting rationally, on this view, depends in large part on conforming our actions to what we believe. So, for Chan, it therefore also seems natural to focus her view of our practical identity on what we believe to be essential features (Chan 2019, 11).

3.3. Golub on 'Biographical Identity'

Camil Golub (2019) presents an alternative account according to which we can think of an ethical identity in terms of a 'biographical' identity (Golub 2019, 81).

Unlike Chan's account, Golub's account draws inspiration from so-called 'narrative' views of identity. Roughly speaking, the idea of 'narrative' views of identity is that people build up their self-identity by creating a story or narrative of their life in which particular experiences, projects, events, etc. are of central importance. On this view, this narrative or life story forms our 'self' or ethical identity (Golub 2019, 83).

In line with this narrative view of identity, Golub refers to our ethical identity as a 'biographical' identity because it reflects the idea that who we are is in large part about the biography we have created for ourselves (Golub 2019, 81). Note that the experiences, projects or events that we weave into our life story or biography can be both positive or negative in valence (Golub 2019, 83). For example, some hardship I experienced in the past may play an important part in my life story because it strongly shaped who I am today. Similarly to Chan, Golub is non-committal about exactly what is part of our biographical identity and what is not. On his view, 'our judgments about who we are in a biographical sense are typically imprecise and shifty, and any good theory of these matters should reflect this' (Golub 2019, 83). Finally, it is worth noting that

Golub's account does not imply that we endorse everything we did in the past. According to Golub, there might be past actions which have become part of who we are but that we still regret because they were ethically problematic (Golub 2019, 82).

Golub uses his concept of 'biographical identity' to respond to a particular philosophical puzzle about regret. Sometimes, Golub suggests, we judge that certain alternative lives would have been better for us. However, at the same time, we do not regret that we did not live these lives. For example, Golub suggests that, according to his own standards, it would have been better for him if his parents had emigrated to France when he was a child. However, at the same time, he also claims that he does not regret not living this alternative life in France (Golub 2019, 72-73). According to Golub, the 'biographical identity' view can help us to explain this phenomenon. On the 'biographical view', the events or experiences we actually lived through have become part of our life story and we therefore often *affirm* them as part of our biography (Golub 2019, 82). The fact of not regretting alternative lives that might have been better for us thus expresses our affirmation of who we are — that is, our affirmation of the experiences and events that make up our actual biography (Golub 2019, 82-83).

Let me now draw the different strands of this discussion together. In transformative experience cases, such as *Charity Abroad*, there is a sense in which we stay the same person and a sense in which we 'become' a different person. Metaphysically speaking, we remain numerically identical to the person we were before the transformative experience, at least in most of these cases. However, on a practical or ethical level, we become a different self. This might be the case because transformative experiences replace a sufficient number of features we take to be essential to who we are, or because these experiences push us towards a new biographical narrative of our life.

Throughout this thesis and unless indicated otherwise, I will have in mind the ethical rather than the metaphysical kind of identity when making claims about how transformative experiences change a person or self. More specifically, I will have in mind the kind of 'practical' identity

defended by Chan (2019). I will lean towards Chan's account of 'practical identity' over Golub's (2019) because it carries less theoretical baggage about the importance of narratives in our lives, and for this reason Chan's account will be more acceptable even to people who are not entirely convinced by narrative accounts of ethical identity.

4. The Ethics of Transformative Experience

As I have shown above, transformative experiences have traditionally been discussed in the context of epistemology and decision theory. In this thesis, my goal is to extend the debate about transformative experiences into ethics.

The ethics of transformative experience is an area of inquiry in which we aim to discover the ways in which going through a transformative experience can give rise to ethical questions about, amongst other things, one's decisions, commitments, attitudes, well-being and relations with others. What makes transformative experiences particularly relevant to ethics is the fact that going through these experiences alters the context in which we typically make ethical decisions. Typically, we know, at least to some extent, what results will come from our decisions and actions and what kind of a person we will be after going through a particular experience. In the case of transformative experiences, we often cannot know what results will come from our decisions and actions and, crucially, what kind of a person we will be after going through a transformative experience.

In the next section of this chapter, I will present existing contributions to the literature on the ethics of transformative experience. I will divide these contributions into two main categories. First, in section 4, I will discuss contributions in ethical theory — that is, contributions which directly or indirectly address questions about the moral rightness or wrongness of actions. Second, in section 5, I will discuss contributions which have a broader understanding of 'ethics' as also including questions of well-being, authenticity and social justice.

In discussing these contributions to the literature on the ethics of transformative experience, my aim is not to establish a set of questions that define the topic, or even to establish a set of questions to be addressed in this dissertation. My aim is instead to make a case for giving the ethics of transformative experience sustained attention, by helping the reader get a sense of the many different ways in which transformative experiences might be relevant to ethics.

5. Transformative Experience and Ethics

In this section of the chapter, I discuss four papers dealing directly with morally right and wrong action in the context of transformative experience. I will discuss papers by Farbod Akhlaghi (2023), Nilanjan Das and L.A. Paul (2020), Dana Howard (2015) and Fiona Woollard (2021, 2022). The main lesson I wish to draw from these papers is that transformative experiences create a gap between old and new selves and this gap gives rise to a number of interesting ethical issues.

5.1. Farbod Akhlaghi on Interfering with Other People's Transformative Experiences

Farbod Akhlaghi (2023) believes that we can often have a strong influence over other people's transformative experiences. Akhlaghi therefore wants to focus on the conditions under which we are morally justified in *interfering with someone else's choice* to undergo a transformative experience (Akhlaghi 2023, 3-4).

Consider the following case inspired by Akhlaghi:

Friends: Siavash is a highly paid lawyer. He is now considering whether to leave his job and become a pottery artist instead. Shireen, Siavash's best friend, is wondering whether she should prevent Siavash from giving up his current job (Akhlaghi 2023, 4).

On Akhlaghi's view, *Friends* involves a potentially (epistemically and personally) transformative experience. By radically changing her career path, Akhlaghi argues, Siavash is, for instance, also

likely to develop new desires, commitments, values, etc. and thus develop a new self (Akhlaghi 2023, 4).

I think that many people can imagine themselves in Shireen's position as Siavash's best friend. For Shireen, Siavash's decision can raise all kinds of worries. For example, Shireen might fear that she will spend less time with Siavash in the future or that Siavash will be unhappy in his new job or that Siavash has underestimated the importance of material wealth in his life. The key question, therefore, according to Akhlaghi, is the following: Under what conditions are people such as Shireen morally justified in interfering with other people's choices to undergo transformative experiences? (Akhlaghi 2023, 4-5). As Akhlaghi specifies in his paper, 'interference' is here broadly understood to range from rational persuasion all the way to coercion (Akhlaghi 2023, 5).

5.1.1. Three Inadequate Answers

Akhlaghi wants to consider three initial answers to the above question which he judges to be inadequate.

First, we might think that we are morally justified in interfering with someone else's choice to undergo a transformative experience only if we can anticipate that this experience will be a negative experience for this person. According to Akhlaghi, a transformative experience can be 'epistemically negative' if it decreases our knowledge or understanding and it can be 'personally negative' if it changes us morally or prudentially for the worse (Akhlaghi 2023, 5). For example, on this view, Shireen might argue that Siavash's desire to become a pottery artist will somehow leave him epistemically or personally worse off.

According to Akhlaghi, the main problem with this first view is a lack of predictability when it comes to transformative experiences. He argues that, often, we cannot know whether a transformative experience would be positive or negative. In addition, even if we know that a certain transformative experience *tends to be* positive or negative, we usually do not know in what way it is

positive or negative *for any given person.* And if we cannot know whether and how a transformative experience will be positive or negative for a person, we cannot appeal to the personal value of this experience to justify any intervention with someone else's choice to undergo this experience. (Akhlaghi 2023, 5-6)

Notice that, on Akhlaghi's view, this does not mean that we are never justified in interfering with someone else's choice to undergo a transformative experience. For example, Akhlaghi claims that while we might not know what it is like for someone to decide to go on a killing spree, we surely are morally justified in interfering with this person's choice (Akhlaghi 2023, 6).

Second, Akhlaghi considers the view according to which we are only justified in interfering with someone else's choice to undergo a transformative experience when it is in this other person's 'best interests' (Akhlaghi 2023, 6). For example, maybe Shireen knows that, if Siavash decides to become a pottery artist, he will no longer have enough money to go on holidays, an activity Siavash enjoys a lot. Maybe, on this view, Shireen would be justified in interfering because she knows that becoming a pottery artist will not be in her best friend's best interests.

The third view that Akhlaghi considers appeals to standard decision theory to determine when it is morally permissible to interfere with someone else's choice to undergo a transformative experience. On this view, for any transformative choice, we need to determine the expected utilities of each potential outcome for the relevant person and, if the outcome of the transformative experience does not maximize expected utility, we are permitted to interfere (Akhlaghi 2023, 6).

According to Akhlaghi, the second and the third views both suffer from the same two problems.

The first problem is that both views assume that we can know what is in someone's best interests or which option will maximize expected utility for them *before* they undergo the transformative experience. However, on Akhlaghi's view, this assumption cannot be sustained. He argues that, given the radically transformative nature of these experiences, we often cannot know in advance what is in someone's interest and whether or not their interests will be fulfilled after

the transformative experience (Akhlaghi 2023, 6-7). For example, suppose Shireen is worried that Siavash will not have enough financial resources to go on holidays after becoming a pottery artist. Maybe, however, Siavash will be so fulfilled in his new life that he would happily sacrifice holiday time after his transformative experience.

The second problem is that both views do not specify *whose* best interests or utility should be considered. After all, Akhlaghi argues, for any person undergoing a transformative experience, there is a self before the experience and there is a different self after the experience (Akhlaghi 2023, 7). For instance, Siavash might have completely different core interests, desires, values, etc. after becoming a pottery artist. As a result, the two views above are not clear about which self — the one before or after the transformative experience — should be considered when deciding about the permissibility of interfering.

5.1.2. Akhlaghi and the Right to Revelatory Autonomy

Akhlaghi goes on to present his own view, which, he argues, avoids the problems faced by the three views discussed above.

According to Akhlaghi, it is important for us, as humans, to be the 'authors' of our own lives (Akhlaghi 2023, 8). What he means by this is not that we can always decide who we will become. Evidently, given the transformative nature of certain experiences, we cannot always decide who we will become. Instead, according to Akhlaghi, it is important for us to be able to choose to *discover* who we will become through our own choices. On this view, being the author of our own lives contributes to our sense of self by allowing us to discover how our own choices transform us (Akhlaghi 2023, 8).

Given the importance of 'self-authorship' to our lives, Akhlaghi argues that we possess a

Right to Revelatory Autonomy: The moral right to decide for ourselves to discover who we will become and how our life will turn out (Akhlaghi 2023, 8).[10]

Akhlaghi also suggests that this moral right gives rise to the following duty:

Duty of Revelatory Non-Interference: The moral duty not to interfere with others' autonomous choices to discover who they will become and how their lives will turn out (Akhlaghi 2023, 9).

On Akhlaghi's view, the right and the corresponding duty allow us to specify when it is morally permissible to interfere with someone else's choice to undergo a transformative experience. According to Akhlaghi, we are morally permitted to interfere with someone else's choice to undergo a transformative experience only if this other person's *Right to Revelatory Autonomy* is outweighed and our *Duty of Revelatory Non-Interference* is no longer binding (Akhlaghi 2023, 9).

For example, we might imagine that, if Siavash's choice to become a pottery artist was very likely to put his children into poverty, then Shireen would have a right to interfere with his choice. After all, the moral reason for preventing Siavash's children from living in poverty would plausibly be stronger than Siavash's moral *Right to Revelatory Autonomy* and therefore neutralise Shireen's *Duty of Revelatory Non-Interference.*

In addition, on Akhlaghi's view, we can explain why it is harder to justify our interference in some cases than in others. The reason is that some cases of transformative experience are more

---

[10] Akhlaghi's solution is inspired by Paul's solution to the decision-theoretic problem. According to Paul (2014), when we are faced with a transformative choice, we should not choose on the basis of what the experience will be like for us. Rather, on her view, we should choose in favour or against 'revelation'— that is, we should choose whether or not we want to discover a new experience and the new life and preferences resulting from this experience (Paul 2014, 115-123).

central to the self-authorship of our lives than others. For example, Akhlaghi argues that eating a cheeseburger for the first time does not affect our core preferences, desires, etc. as strongly as, say, choosing to attend university (Akhlaghi 2023, 9-10).

According to Akhlaghi, his view is able to avoid the problems of the other three views discussed above. His view does not depend on any assumption about the (positive or negative) valence of the transformative experience for the relevant person nor on any assumption about what is in the relevant person's best interests. In addition, on this view, it is clear that we owe our moral duty to the present self of the person who exercises their *Right to Revelatory Autonomy* and not to some future self that we cannot know about prior to the transformative experience (Akhlaghi 2023, 9).

Finally, Akhlaghi wants to address the following objection: Recall that the *Right to Revelatory Autonomy* and the *Duty to Revelatory Non-Interference* prohibit interference on the grounds of protecting the relevant person's self-authorship — that is, their right to decide to discover how their decisions will shape them. On this objection, there is an easy way to respect a person's right to self-authorship and interfere with that person's decision to undergo a transformative experience. By *rationally persuading* someone not to undergo a transformative experience you respect their right to choose because you are not coercing or forcing them into a decision. At the same time, on this view, you can also convince the relevant person not to go through with their choice (Akhlaghi 2023, 9-10). For example, on this objection, it seems that it is morally permissible for Shireen to try to rationally persuade Siavash not to become a pottery artist (because of, say, the financial consequences of this decision). This objection seems problematic for Akhlaghi's view because it opens the door to an easy way of permissibly interfering with many people's decisions to undergo a transformative experience.

In response, Akhlaghi wishes to distinguish between respect for autonomy as such and respect for *revelatory* autonomy. On his view, rational persuasion respects the former but not necessarily the latter. By trying to rationally persuade someone you respect their status as a

'competent reasoner' (Akhlaghi 2023, 10). However, Akhlaghi argues that you do not necessarily respect their status as someone who should make their *own* transformative decisions (Akhlaghi 2023, 10). For example, if Shireen tries to rationally persuade Siavash because she has already made a similar decision in the past, she might undermine Siavash's ability to take his decision *for himself*. She does not give Siavash the space to discover for himself where his decision will lead. So, on Akhlaghi's view, while rational persuasion can respect revelatory autonomy, it does not always do so.

5.1.3. Two Problems for Akhlaghi's Argument

I can see two potential problems with Akhlaghi's revelatory autonomy argument. I suggest that Akhlaghi can mitigate the force of the first problem but not necessarily that of the second problem.

First, Akhlaghi argues that we are morally permitted to interfere with someone's choice to undergo a transformative experience only if that person's right to revelatory autonomy is outweighed by some other consideration(s). For example, in Akhlaghi's case of someone who decides to go on a transformative killing spree, it is clear that this person's right to revelatory autonomy is outweighed by the need to protect other people's lives (Akhlaghi 2023, 6, 9).

This argument assumes that we can usually know what the wider effects of a certain transformative experience are going to be before the relevant person undergoes this experience. However, this does not always seem to be the case. Often, we might be able to know the wider effects of someone's transformative experience only once they have gone through this experience. Recall Akhlaghi's (2023) *Friends* case in which Siavash decides to radically change his career. It seems plausible that we cannot always know what the wider effects of this decision are going to be *before* Siavash has gone through the experience. It might turn out to be a good choice that benefits both Siavash and his family. However, it might also turn out to be a catastrophic choice that leads Siavash's children into poverty. And if we do not know how the transformative

experience is going to turn out before the experience, we do not know whether or not we have the right to interfere with it.

There might be a way for Akhlaghi to, at the very least, mitigate the force of this first criticism. On his view, we can appeal to the *expected value* of interfering with, say, Siavash's career decision, to judge whether we are justified in doing so or not. If, on this view, we deem the expected value of our interference to be high enough, we are justified in interfering and thus overriding someone's right to revelatory autonomy. For example, if we can reasonably predict that, due to the high financial risk associated with Siavash's career choice, his children are very likely to fall into poverty, then we are, on this view, justified in interfering with his career decision. (Of course, we have to be aware that there might be other cases in which it is difficult to judge if the expected value of interfering is high enough or not to justify our violation of someone's right to revelatory autonomy.)

Second, suppose that there is some coherent notion of self-authorship underpinning Akhlaghi's argument. This still leaves open the question of why we should attach importance to this kind of self-authorship to the point that there is a right to revelatory autonomy. According to Aklhaghi (2023), 'some degree of self-authorship in this sense is crucial for us and others to see ourselves as *our*selves – selves we have become at least partly through transformative choices we have made' (Akhlaghi 2023, 8). This might seem puzzling to some readers. How can a transformative experience — an experience which, by definition, makes you into someone else — contribute to a unifying sense of '*our*selves'? If anything, it seems that transformative experiences undermine our sense of self by splitting our existence into radically different selves.

More importantly, why should this desire for self-authorship be so important that it should be protected by a right to revelatory autonomy that can be outweighed only under strict circumstances? Akhlaghi (2023) does not directly address this question, but it seems important for his argument, and I expect pursuing it would be an interesting further avenue of inquiry.

5.2. Das and Paul on Non-Identity

The second paper I want to discuss was written by Nilanjan Das and Paul (2020). It is worth noting that, unlike the previous paper, Das and Paul's paper deals mainly with issues of *rational* decision-making, as opposed to *moral* decision-making, in the context of transformative experiences. Nonetheless, I want to discuss the paper in an ethical context for two reasons. First, Das and Paul (2020) explicitly draw on the literature of the 'non-identity problem' in population ethics. In other words, they draw clear parallels between their views in rational choice theory and relevant ethical literature. Second, I will draw on Das and Paul's paper later on in my thesis to develop one of my own puzzles in the ethics of transformative experience.

To present Das and Paul's main argument, I need to take a step back and briefly discuss the so-called 'non-identity problem' in population ethics.[11]

Consider the following case.

Doctor: A couple decide to have a child. Before conceiving the child, the couple attend a pre-conception screening. The doctor tells the couple that if they conceive now, their child will have a life worth living but it will be born significantly impaired. However, the doctor adds that, if the couple follow a minor medical treatment for one month, their child will be born without any impairment.

Intuitively, many people believe that the couple should wait the one month before conceiving their child. However, it turns out that it is more difficult to defend this intuition than we initially think.

If the couple decide to wait one month before conceiving, they will conceive a numerically different child — child 2— to the child they would have conceived had they conceived

---

[11] For a classic statement of the non-identity problem, see Derek Parfit (1984) *Reasons and Persons* (Oxford: Clarendon Press), pp.351-380.

immediately — child 1. The reason for this is that child 1 is formed by a different egg-sperm combination than child 2. If child 1 and child 2 are numerically different, we cannot say that child 1 was made worse off by being conceived right away. After all, child 1 has a life worth living and had it not been for the parents' decision to conceive right away, child 1 would not have existed. And, usually, we think that our decisions are morally permissible if they do not make anyone worse off — that is, worse off than they would have been if we had made a different decision. So, counterintuitively, it seems that conceiving child 1 right away in an impaired condition is morally permissible.

The non-identity problem relies on three key principles that Das and Paul draw on in their paper.

First, the problem relies on the principle of

Act-Dependence: Non-identity acts create people that would not otherwise have existed.

In the *Doctor* case, conceiving now brings into existence a person — child 1 — who would not have existed if the couple had waited one month before conceiving. In contrast, if the couple decide to conceive one month from now, they will bring into existence another person — child 2— who would not have existed if they had conceived right away. Again, the reason for this is that depending on the moment of conception, a different egg-sperm combination will come together, thus creating either child 1 or child 2.

Second, the non-identity problem relies on the so-called

Comparative Account of Harm: An act harms a person only if the act makes the person worse off than they would otherwise have been.

In the *Doctor* case, the *Comparative Account of Harm* implies that the couple is not harming child 1 because conceiving child 1 now does not seem to make it worse off than it would have been otherwise. After all, child 1 has a life worth living and it would not have existed at all if the couple had decided to wait one month before conceiving.

Third, the non-identity problem also relies on the

Person-Affecting Principle: An act is morally wrong only if it is bad for someone (Parfit 1984, 358-363).

If we consider the *Doctor* case, it seems the *Person-Affecting Principle* implies that the couple's act/choice to conceive now is not morally wrong. Given that the couple's choice does not harm child 1, it also does not seem to be bad for child 1. As a result, it seems that the couple's decision to conceive child 1 right away is not morally wrong.

5.2.1. Three Analogous Principles

In their paper, Das and Paul aim to develop a problem involving transformative experience that relies on three principles that are analogous to the three principles of the traditional non-identity problem (Das and Paul 2020, 187). The main differences with the traditional non-identity problem are that Das and Paul focus on *selves* instead of persons, on *transformative experiences* instead of *non-identity choices*, and on *rational choice theory* instead of ethics.

Das and Paul start their main argument by assuming that we all possess a 'self'. They define a person's self as the 'core values, beliefs, desires, commitments, ideals, character traits, etc., that make her who she is at that time, including the first person phenomenology that is realized by having these core values, etc.' (Das and Paul 2020, 190).

Das and Paul continue their argument by offering three principles about selves and transformative experiences that are analogous to the three principles discussed above in the context of the non-identity problem.

First, they offer the principle of

Act-Dependence*: All else equal, personally transformative acts create selves that would not otherwise have existed (Das and Paul 2020, 190).

The main rationale behind this revised principle is metaphysical. On Das and Paul's view, the self I possess is constituted by mental states such as my core desires, values, commitments, etc. Let us refer to these mental states as 'self-identity conferring mental states'. If these mental states change radically, then, Das and Paul argue, my self also changes into a new self (Das and Paul 2020, 190-191). For example, eighteen-year old me might possess one self, while twenty-five year old me might possess a different self, because twenty-five year old me has radically different self-identity conferring mental states.

This view about the persistence and change of selves is the reason why transformative experiences are important to Das and Paul's argument. On this view, personally transformative events affect our core desires, values, etc. and hence the selves that correspond to them. Specifically, every time we go through a personally transformative event, a new self is created which corresponds to a new set of self-identity affecting mental states. Crucially, on Das and Paul's view, if we had not gone through the personally transformative experience, the new self would not have been created. So, it is this metaphysical view about persistence which, according to Das and Paul, can motivate the principle of *Act-Dependence** (Das and Paul 2020, 191-192).

Second, Das and Paul offer a revised comparative account of harm:

Comparative Account of Harm*: A transformative act harms a self only if the act makes the self worse off than it would otherwise have been (Das and Paul 2020, 194).

Das and Paul motivate this second revised principle by appealing to the lived experience of many disabled people. They report that many disabled people believe that, in the absence of disability, they would not possess the same self — the self they currently possess and that is so valuable and integral to who they are (Das and Paul 2020, 194-195). Instead, as these people see things, in the absence of disability, they would possess a completely different self to the one they currently possess. Now, if, in the absence of disability, disabled people would possess a different self, we cannot say that their current self is harmed by its disability. After all, their current self has an existence worth living and there is no disability-free alternative in which the current disabled self would be better off.

In general, the lesson that Das and Paul seem to draw from the lived experience of disabled people is that any claim about harm to a current self must assume a comparison with an alternative scenario in which this same self is worse off (Das and Paul 2020, 195). And this is exactly the message encapsulated by the *Comparative Account of Harm**.

Third, Das and Paul offer the

Self-Affecting Principle: An agent can rationally prefer not to perform a certain act A only in the following conditions: 'There exists an alternative act B and a present or future self x such that x's expected well-being conditional on the agent performing B is greater than x's expected well-being conditional on the agent performing A' (Das and Paul 2020, 193).

To explain why they believe this to be a plausible principle, Das and Paul ask us to consider the following two cases.

Wisdom Teeth 1: You will have your wisdom teeth removed in a painful but safe procedure. You have a choice between two surgeons. If you choose the first surgeon, the procedure will last, say, one hour. If you choose the second surgeon, the same procedure will last two hours. In other words, if you choose the second surgeon, you will be in pain for longer (Das and Paul 2020, 193).

Wisdom Teeth 2: You will have your wisdom teeth removed in a painful but safe procedure. You have a choice between two surgeons. No matter which surgeon you choose, the procedure will last for roughly one hour (Das and Paul 2020, 193).

According to Das and Paul, we should all agree that it is rational for you to choose surgeon one in *Wisdom Teeth 1* and to be indifferent about which surgeon performs the procedure in *Wisdom Teeth 2*. On their view, the *Self-Affecting Principle* can explain why this is so. On this principle, you have two options involving the same self in *Wisdom Teeth 1*. Given that your self's expected well-being is greater in choosing surgeon one, the principle plausibly tells us to choose surgeon one. In contrast, the principle tells us to be indifferent in *Wisdom Teeth 2*. After all, your self's expected well-being is the same no matter which surgeon you choose. So, according to Das and Paul, the fact that the *Self-Affecting Principle* gives us the right verdict in these two cases lends it overall credibility (Das and Paul 2020, 194).

5.2.2. A New Non-Identity Problem for Rational Decision-Making

Das and Paul use the three principles just discussed to present a problem about rational decision-making that is analogous to the ethical problem in the non-identity literature. They ask us to consider the following case:

Cataract Surgery: According to your doctor you will soon become blind unless you undergo cataract surgery.[12] Becoming blind would not be devastating for you as you have a supportive family and live in a society that accommodates blind people's needs. However, becoming blind would mean that you could not enjoy many activities you currently enjoy (Das and Paul 2020, 195).

According to Das and Paul, it is not rationally *obligatory* for you to choose to undergo surgery in this case. However, in their opinion, it is intuitively correct to say that it is rationally *permissible* for you to undergo surgery. After all, in this case, becoming blind would deprive you of a significant number of activities you currently enjoy (Das and Paul 2020, 196).

Das and Paul argue that we cannot uphold this intuition — the intuition that it is rationally permissible for you to prefer to undergo surgery — if we accept the three revised principles discussed above (Das and Paul 2020, 196).

Becoming blind in *Cataract Surgery* counts as a personally transformative event. After all, undergoing this experience is highly likely to change many of your core desires, values, etc. If undergoing this experience is personally transformative, then, following the principle of *Act-Dependence**, it gives rise to a self that would not otherwise have existed (Das and Paul 2020, 196).

Crucially, on Das and Paul's view, we cannot pretend that this new self was harmed by being brought into a blind existence. On the *Comparative Account of Harm**, the new self is not harmed because it is not made worse off than it would otherwise have been. It has an existence worth living and, if it had not been for the blind existence, the new self would simply not have existed (Das and Paul 2020, 196).

And, following the *Self-Affecting Principle*, if the new self is not harmed by its blind existence, it seems that it is irrational for you to prefer to undergo cataract surgery. On this principle, you

---

[12] We can assume here that the surgery is low-risk and almost pain-free.

can rationally prefer to undergo an act such as having cataract surgery only if undergoing surgery makes your self's expected well-being greater than in some alternative act. However, given that becoming blind counts as a personally transformative event, you would be equipped with a new self after becoming blind and there would be no alternative act in which your initial self's expected well-being was greater. Hence, it seems that it would be irrational for you to prefer to undergo cataract surgery. In fact, according to Das and Paul, given their principles, you ought to be indifferent between undergoing or not undergoing cataract surgery (Das and Paul 2020, 196-197).

On Das and Paul's view, this conclusion is problematic because it conflicts with the intuitively correct view — the view that you can rationally prefer to undergo surgery in *Cataract Surgery*. In other words, Das and Paul conclude that we must either give up the intuitively correct view about the rationality of choice in cases such as *Cataract Surgery* or we must give up at least one of the three plausible revised principles on which their puzzle relies (Das and Paul 2020, 197).

Das and Paul do not directly address the question of whether rejecting one of these revised principles to solve their puzzle also implies rejecting the corresponding principle in the original non-identity problem. However, they do write that their puzzle has a structure 'that is exactly analogous to the non-identity problem' and that the non-identity problem is part 'of a more general class of problems' (Das and Paul 2020, 203). From this, it seems relatively safe to assume that Das and Paul think that any solution to their puzzle will have direct implications for the original non-identity problem.

5.2.3. Three Unsuccessful Solutions

Das and Paul explore three solutions to their new puzzle.[13] In what follows, I will briefly present these solutions and explain why Das and Paul think that each of them is unsuccessful.

---

[13] The three solutions mirror three classic solutions to the traditional non-identity problem in ethics.

First, Das and Paul suggest that we could deny the intuition that it can be rational in *Cataract Surgery* to prefer undergoing the procedure to avoid becoming blind. On their view, the most plausible motivation for denying the intuition is the following: When we reason about cases such as *Cataract Surgery* we might often suffer from an irrational status quo bias. Concretely, we might irrationally prefer undergoing the procedure to remain sighted because we are biased towards the sighted state of affairs we currently find ourselves in. On this objection, if we were to rid ourselves of this bias, we would see that the only rational option in *Cataract Surgery* is to be indifferent about whether we undergo the procedure or not (Das and Paul 2020, 198-199).

In response, drawing on the work of Derek Parfit (2011), Das and Paul argue that we need to distinguish between a subjective and an objective view of preferences and that, on both of these views, denying the intuition is implausible (Das and Paul 2020, 199).

On a subjective view of preferences, Das and Paul argue that we have no reason for our preferences, except when they help us to fulfil some of our other preferences. On this view, it might still be rational for us to prefer to undergo the procedure in *Cataract Surgery* and remain sighted. After all, it could be that your subjective preferences are simply such that you prefer your current sighted way of being (that allows you to pursue many activities you currently enjoy) to a blind existence. According to Das and Paul, the subjective view of preferences simply does not give us a reason to dismiss such a preference for the status quo (Das and Paul 2020, 199).

On an objective view of preferences, Das and Paul argue that the reason for our preferences for some outcomes over others stems from features of these outcomes themselves (rather than from subjective features about us) (Das and Paul 2020, 199). Das and Paul suggest that we could apply this objective view of preferences to *Cataract Surgery* as follows. On their view, being sighted gives you certain 'capacities' that being blind does not give you (Das and Paul 2020, 199).[14] It might be precisely those capacities that, on an objective view of preferences, make a

---

[14] Of course, the reverse might be true as well.

sighted self so well off that any benefits of a blind existence could not compensate the loss of sight (Das and Paul 2020, 199).

If Das and Paul are correct in their application of the objective view of preferences to *Cataract Surgery*, then the objective view implies that it can be rational for you to prefer to undergo the procedure to remain sighted.[15] It follows that, given the objective view, we cannot reject the intuition that it is rationally permissible to undergo the surgery. And if we cannot reject this intuition, then the puzzle remains (Das and Paul 2020, 199).

The second possible solution is to deny the *Comparative Account of Harm*. Recall that this principle states that an act harms a self only if it makes this self worse off than it would otherwise have been. On this second solution, we need to find a way to argue that an act can harm a self even if it does not make this self worse off than it would otherwise have been. One way of doing this, according to Das and Paul, is to adopt a form of so-called 'welfare objectivism'. On this theory, there are certain objectively definable things (such as certain capacities, liberties, etc.) the possession of which make your life go well. If, on this theory, you do not possess these things, your life does not go as well as it could go (Das and Paul 2020, 199-200).

The second solution could thus claim that an act might not make a self worse off than it would otherwise have been but it can still harm the self by depriving it of one or more of these objectively good capacities, freedoms, etc. For example, if there is a self that could have existed only in a disabled body, it is true that this self is not worse off than it would otherwise have been. However, this self can still be harmed because it might lack certain capacities or freedoms that are objectively good for its well-being.

According to Das and Paul, this second solution is problematic because it cannot provide the right verdict in cases such as *Cataract Surgery* (Das and Paul 2020, 200). Recall that, ideally, we

---

[15] Presumably, Das and Paul think that the objective view of preferences might also support our starting intuition in the original *Doctor* non-identity case.

want a solution which tells us that it is rationally permissible to undergo cataract surgery. Suppose there is a list of objectively good things for your well-being and a blind person who refuses to undergo cataract surgery lacks some of these things. If welfare objectivism is true, it seems that this blind person should feel quite harmed by her blind condition. However, if we take seriously the testimony of many blind people, it seems that most of them do not feel harmed by their condition. In other words, from many blind people's own perspective, it seems that the objectively good things that they are lacking are not a harm to them at all. In fact, from their perspective, the only potential harm is that of undergoing cataract surgery. And, if the only potential harm for them is that of undergoing cataract surgery, then it is not rational for them to undergo cataract surgery in the first place (Das and Paul 2020, 200).

The third possible solution is to deny the *Self-Affecting Principle*. According to this principle, you can rationally prefer not to undertake a course of action only if this course of action harms or makes your present or your future self worse off than it would have been with an alternative act. According to Das and Paul, the main idea of this third solution is to argue that there is a more plausible theory of rational choice which is incompatible with this principle (Das and Paul 2020, 200-203).

In total, Das and Paul discuss four theories of rational choice which are incompatible with the *Self-Affecting Principle*. I will here focus on the theory they deem most plausible.

> *Averagist Theory of Prudence*: An agent can rationally prefer to perform a certain act A rather than an act B 'iff the expected average well-being of the agent's possible current and future selves conditional on her performing A is greater than the expected average well-being of the agent's possible current and future selves conditional on her performing B' (Das and Paul 2020, 202).

The main idea behind this principle is that act A and act B give rise to different current and future selves. On this principle, we can calculate the expected average well-being of the selves resulting from act A and the expected average well-being of the selves resulting from act B. Ultimately, on this view, the agent should then prefer the act with the highest expected well-being average.

Das and Paul suggest that this theory's emphasis on *average* expected well-being might solve the problem in *Cataract Surgery*. On this view, it can be rational for you to prefer undergoing cataract surgery because your expected average well-being as a sighted person (after the surgery) might be higher than as a blind person (Das and Paul 2020, 202).

However, ultimately, Das and Paul argue that the *Averagist Theory of Prudence* is not a plausible alternative to the *Self-Affecting Principle* (Das and Paul 2020, 202). On their view, this average-based theory does not provide the correct answer in cases such as the following:

The Life-Prolonging Drug: Suppose you know that the rest of your life will be absolutely amazing. You are now offered the opportunity to take a drug which prolongs your life by one day more than you would otherwise have lived. During this additional day, your life will be *slightly less* amazing than before, though it will still be amazing (Das and Paul 2020, 202).

Das and Paul ask us to imagine that you undergo a personal transformation each year — that is, you get a new self each year. You now have two possibilities: either you take the life-prolonging drug or not. If we follow the *Averagist Theory of Prudence* it seems that it is irrational for you to take the life-prolonging drug. After all, if you were to take the drug, the average well-being of your current and future selves would be lower than if you refused to take the drug. The additional day provided by the drug would lower your average because this day is slightly lower in quality of life (Das and Paul 2020, 203).

On Das and Paul's view, this verdict cannot be right. Surely, they think, it is rational for you to take the life-prolonging drug. For the additional day you gain through the drug will still be excellent even if not as amazing as the previous days. So, it seems perfectly rational to take the drug and offer yourself an additional excellent day (Das and Paul 2020, 203).

5.2.4. A Problem for Das and Paul's Puzzle

As we have just seen, Das and Paul explore three potential solutions to their new puzzle and argue that none of these solutions offers a plausible way forward. However, there is a potential problem for Das and Paul's overall argument that they do not address.

As noted above, according to Das and Paul, we cannot rationally choose against undergoing transformative experiences in many cases. For example, Das and Paul argue that you cannot rationally choose against becoming blind (and thus in favour of undergoing surgery) in *Cataract Surgery* (Das and Paul 2020, 195-197). The main reason for this seems to be that, on their view, transformative experiences, such as becoming blind, do not lower your present self's expected well-being. After all, according to Das and Paul, your present self is simply replaced by a new self after the transformative experience (Das and Paul 2020, 195-197).

Consider now the following problem. On Das and Paul's view, going through a transformative experience means the end for your current self. Your current self 'dies' or, at least, goes out of existence. Does this leave your current self worse off than persisting? Das and Paul have two options, both of which are problematic for their view. If they think that going out of existence leaves your current self worse off, their puzzle disappears. After all, this would imply that you could rationally decide to have cataract surgery because not having surgery would leave your current self worse off, by forcing it out of existence.

If Das and Paul instead think that going out of existence does not leave your current self worse off, then they would get the highly counter-intuitive result that you cannot rationally prefer not to commit suicide. After all, suicide also means the end of your current self and, on this version

of Das and Paul's view, this would not leave your current self worse off than persisting. This is a highly problematic result. Surely, most people agree that you can rationally refuse to kill yourself.

It thus seems that, to generate their puzzle, Das and Paul must make an assumption about rational decision making in transformative contexts that has highly implausible results. If instead we reject this problematic assumption, the puzzle disappears.

Despite this apparent problem with the argument offered by Das and Paul, I do think that transformative experiences generate a problem structurally very similar to the non-identity problem. However, in my view, this problem has to do with our *backward-looking attitudes*, rather than with our *forward-looking decisions*. I develop this new problem in Chapter Two.

5.3. Dana Howard on Transforming Other People

Dana Howard starts her 2015 paper with the claim that, sometimes, whether we undergo a transformative experience or not is not under our control (Howard 2015, 342). In these cases, it is often other people who can decide whether or not we undergo a transformative experience and thus become a different self. With this in mind, Howard sets the main question of her paper as follows: Under what conditions is it morally justified to *make* someone else undergo transformative experiences (Howard 2015, 342)?

I think that Howard's emphasis on *our transformative influence on others* is an important one. In the traditional decision theory literature, transformative experiences are often discussed exclusively from the perspective of a single individual. In reality, however, many of the transformative experiences we undergo are not the result of our own choices but are strongly influenced by other people's choices. For instance, many of the medical, educational, religious and cultural choices that our parents make on our behalf can have a transformational impact on us. Similarly, large-scale social and political policies that are not under our control can have a strong transformative impact on our lives.

Howard continues her paper by asking us to consider the following two cases.

Cochlear Implants: The doctor tells you that your very young child risks developing a significant hearing impairment. However, according to the doctor, this can be avoided via the implantation of cochlear implants. As a parent, you now have to decide whether or not your child should get the cochlear implants (Howard 2015, 342).

Clown: Your family has worked as clowns for several generations. You want your son to continue this tradition. However, to be effective as a clown, your son needs to walk in a funny manner. So, you decide to break your son's legs just as your father did with you. As a result, your son gets the perfect clown walking style (Howard 2015, 346).

On Howard's view, both cases are potentially transformative. In *Cochlear Implants*, if the parents choose against the implants, their child will develop a significant hearing impairment. In contrast, if the parents decide in favour of the implants, their child will not develop the hearing impairment. Depending on this decision, the child's life will be radically different. Depending on how their parents choose, they will probably perceive the world differently, develop different desires, commitments, etc. and become part of different communities. As a result, it seems that no matter how the parents choose, their choice will be transformative (Howard 2015, 342-343).

According to Howard, *Clown* is also a potentially transformative case. By deciding to break your son's legs, you not only change your son's physical constitution, but you are also highly likely to transform your son's core commitments, desires, etc (Howard 2015, 346-347). For example, while your son might have liked to have a professional sports career, he will not be able to do this anymore. As a result of your decision, your son will also become part of the clown community, a community he might not have become part of if you had not broken his legs.

We can anticipate what Howard takes to be the correct verdict in the two cases just presented. According to Howard, either choice is morally permissible in *Cochlear Implants* while the

father's decision in *Clown* is morally impermissible (Howard 2015, 364). However, the key question for Howard is: How do we justify either of the parents' decision in the first case while rejecting the father's decision in the second case (Howard 2015, 363)?

5.3.1. A First Attempt at a Justification

Howard starts by considering a first potential solution. This first solution appeals to a commonplace type of reasoning. Often, when we take decisions on someone else's behalf, we think about whether this person will endorse our decision in the future (Howard 2015, 350). According to Howard, this intuitive type of reasoning relies on the following principle:

> *Predictive Glad*: If I can predict that you will endorse my current decision on your behalf on reasonable grounds in the future, my decision is morally permissible (Howard 2015, 350).[16][17]

Howard suggests that we can try and apply this type of reasoning in *Cochlear Implants*. On this view, we need to consider whether or not the child will endorse either of the parents' decisions in the future. If we can predict that the child will endorse either or both decisions, then, on this view, either or both decisions are morally permissible (Howard 2015, 360).

While this type of reasoning reaches the correct verdict in *Cochlear Implants*[18], Howard thinks that it reaches the wrong verdict in *Clown*. On Howard's view, we can easily imagine how your clown son could endorse your decision in the future. By breaking his legs, you might have

---

[16] It is important to add the qualification 'on reasonable grounds'. This allows Howard to avoid cases in which you endorse a decision in the future based on, for example, misleading evidence (Howard 2015, 350).

[17] This principle tells us whether or not a decision is morally permissible and not whether it is the best moral decision all things considered.

[18] Even though the principle arrives at the correct verdict, Howard will later reject it even for *Cochlear Implants*.

'allowed' your son to pursue a fulfilling clown career, be part of a supportive clown community and bring joy to other people. For all of these reasons, it is not impossible to imagine that your son would endorse your decision in the future. And, if your son would endorse your decision, it seems that, on this commonplace type of reasoning, your decision is morally permissible. According to Howard, this conclusion is unacceptable (Howard 2015, 360-361).

Howard thinks that we should reject *Predictive Glad* when reasoning about the moral permissibility of the parents' decision in *Cochlear Implants* and *Clown* not only because it reaches the wrong verdict in the latter case. Instead, Howard thinks that the transformative character of cases such as *Cochlear Implants* and *Clown* makes it morally inappropriate to use *Predictive Glad* in these cases. In *Cochlear Implants*, Howard argues, the parents' present decision shapes the desires, commitments, preferences, etc. that their child will develop in the future. In other words, the parents seem to predetermine the child's endorsement of their own decision. So, Howard concludes from this that the future endorsement of the parents' decision cannot justify the decision on its own (Howard 2015, 353, 357).

Howard advances similar reasoning for *Clown*. In this case, Howard argues, you are predetermining your son's endorsement of his clown life. By breaking your son's legs, you force your son into a certain lifestyle that he is later likely to endorse. Again, on Howard's view, the fact that your son will endorse your decision in the future cannot justify this decision on its own (Howard 2015, 353, 357).

5.3.2. An Alternative Justification

As noted in the previous subsection, Howard thinks that there is a fundamental problem with principles such as *Predictive Glad*. In cases in which our present decision will strongly shape some other person's future desires, preferences, etc., Howard thinks that we cannot justify our decision by simply appealing to this person's future endorsement (Howard 2015, 353). To get a better

understanding of why this is the case, Howard asks us to dig a little deeper into the different types of preferences that a person can have.

According to Howard, some of our preferences are

Adaptive Preferences: A preference is adaptive only if it results from a decision that makes you prefer an option from a limited set of feasible options (Howard 2015, 361).

For example, if I decide to indoctrinate you with extreme political views and you later come to prefer/endorse those political views, it is likely that your preference for these views is adaptive. In other words, through my indoctrination, I limited your ability to freely choose among all existing political views. According to Howard, this kind of preference is morally problematic as it limits the set of feasible options from which you can choose (Howard 2015, 361-362).

Having said that, Howard does not claim that we should ignore people's future preferences when making decisions on their behalf. Instead, she argues that people's future preferences/endorsements can and should play two roles in our decision-making on others' behalf.

First, on Howard's view, people's future preferences can play a negative role in helping us to identify which decisions on others' behalf are *not* morally permissible. For example, if we can project that our decision on someone else's behalf will lead to an adaptive preference by this person, then, other things equal, according to Howard, our decision is not morally permissible (Howard 2015, 363).

Howard argues that it is precisely this strategy which helps us to explain why the father's decision to break his son's legs in *Clown* is morally impermissible. On her view, the father's decision to break his son's legs clearly limits the set of feasible options from which the son can choose. By inhibiting his son's physical abilities, the father in fact also inhibits his son's ability to choose lifestyles (such as jobs requiring full physical fitness) which rely on these fully functioning physical abilities (Howard 2015, 363).

Second, and more importantly, Howard argues that people's future preferences can play a positive role in helping us to identify which decisions on others' behalf *are* morally permissible. On her view, when we take a decision on someone else's behalf, we need to a) check whether the person on whose behalf we are taking the decision will endorse it in the future and b) check whether this endorsement is not the result of an adaptive preference. According to Howard, our decision on someone else's behalf is (objectively) morally permissible only if conditions a) and b) are fulfilled (Howard 2015, 367).

Howard argues that either decision in *Cochlear Implants* fulfils the two conditions and this is the reason why either decision is also morally permissible. If the parents decide against the implants, it is likely that their child will endorse this decision in the future. After all, the child will have grown into a deaf lifestyle and community which provide many benefits. In addition, Howard also argues that the child's endorsement will not be an adaptive preference — that is, it will not be the result of a choice from a restricted set of options (Howard 2015, 366-368). There are two reasons for this on Howard's view.

First, there is nothing inherently 'suboptimal' about deafness. Instead, many people nowadays think of a disability such as deafness as a mere difference rather than a lack of some necessary quality.[19] As a result, many people also believe that all lifestyles should be open to deaf people. Second, the practical challenges arising for deaf people are the result of society's failure to accommodate their needs rather than inherent problems due to their condition. So, on Howard's view, it would be reasonable for the child to endorse their parents' decision in the future because

---

[19] Elizabeth Barnes has famously defended this position in several of her writings. These writings include 'Disability, Minority, and Difference' (2009), *Journal of Applied Philosophy* 26 (4): 337-355; 'Valuing Disability, Causing Disability' (2014), *Ethics* 125 (1): 88-113; *The Minority Body: A Theory of Disability* (2016), Oxford and New York: Oxford University Press, pp.78-118.

being deaf does not restrict your set of feasible options in the way breaking your son's legs does (Howard 2015, 363).

Similarly, Howard thinks that, if the parents decide in favour of the implants, it is likely that their child will endorse their decision in the future and that this endorsement will not be an adaptive preference. According to Howard, by choosing in favour of the implants, the parents do not restrict their child's set of feasible options (Howard 2015, 363). For instance, on this view, gaining the ability to hear does not mean that the child will have to choose its, say, lifestyle, from a restricted set of feasible options.

5.3.3. A Problem for Howard's Argument

I can see at least one potential problem with how Howard has presented her argument. The argument relies strongly on the claim that the clown child's set of feasible options is restricted in a morally problematic way while the deaf/hearing child's set of feasible options is not restricted in a morally problematic way. However, it is not fully clear to me why this is the case. It seems that we could make the same argument Howard makes in the case of the deaf/hearing child also in the case of the clown child.

On this objection, there is nothing inherently suboptimal about people with a walking disability and their disability should also be seen as a mere difference rather than the absence of a necessary quality. In addition, on this view, maybe the challenges faced by people with a walking disability are also the result of society's failure to adapt to these people's needs. So, on this view, the clown child's set of feasible options is not restricted differently than that of the deaf/hearing child.

And if the clown child's set of feasible options is not restricted differently than that of the deaf/hearing child, Howard has not given us a justification for why we should morally condemn the clown child's father while accepting the parent's decision in *Cochlear Implants*. In other words,

Adaptive Preference does not allow us to morally distinguish the kinds of cases that Howard wants us to distinguish.

5.4. Fiona Woollard on Pregnancy's Transformative Character

Given that I will engage in detail with Woollard's (2021, 2022) two papers later on in my thesis, I will keep my presentation of her main argument much briefer than for the other three papers.

Woollard starts her argument by arguing that pregnancy is an epistemically transformative experience. On her view, the radical physiological and psychological changes resulting from pregnancy put the pregnant person in a unique epistemic position (Woollard 2021, 157-159). While everyone can have some form of propositional knowledge about pregnancy, Woollard argues that only people who are or have been pregnant can have a full 'appreciative' knowledge of pregnancy — that is, knowledge of what it is like to live through pregnancy from a first-person perspective (Woollard 2022, 472).

According to Woollard, the main reason why only people who are or have been pregnant can have this 'appreciative' knowledge of pregnancy is that only the lived experience of pregnancy can transmit this kind of knowledge. On this view, even the most detailed engagement with, say, narrative literature or testimony about pregnancy cannot recreate the complexity and intensity of the pregnancy experience (Woollard 2021, 159-160).

Crucially, Woollard also argues that this appreciative knowledge that only people who are or have been pregnant can have is necessary for *morally evaluating* practices related to pregnancy, such as abortion (Woollard 2021, 165-166). On this view, abortion involves costs to the pregnant person the complexity and intensity of which only the pregnant person can fully appreciate. In other words, according to Woollard, standard moral evaluations of abortion which ignore people's actual lived experience of pregnancy are, morally speaking, inadequate (Woollard 2021, 166).

Importantly, though, Woollard is not arguing that only pregnant people can morally evaluate abortion. Instead, she is arguing that pregnant people should play an integral part in the

moral evaluation process (Woollard 2021, 166). On this view, philosophers engaging with issues such as abortion need to collaborate with and draw on the knowledge and testimony of pregnant people when morally evaluating these issues.

As I mentioned at the beginning of this subsection, I do not wish to discuss Woollard's argument at great length as I will do so in Chapter Three of this thesis. But to anticipate slightly, there seems to me a serious tension between the idea that testimony is inadequate for transmitting 'appreciative' knowledge of pregnancy and the claim that everyone can morally evaluate pregnancy by drawing on testimony. I will later build on this idea to argue that, if we follow Woollard's reasoning to its natural conclusion, we get the result that, quite generally, transformative experiences can be morally evaluated only by people who have lived through them.

5.5. A General Remark

Before considering some contributions to the literature on transformative experience that raise issues regarding social/political philosophy, I want to briefly draw out a couple of general themes from the preceding discussion.

One concern shared by several of the philosophers just discussed seems to be that transformative experiences create a 'gap' between people's old selves and their new selves. The term 'gap' is here to be understood metaphorically as referring to the epistemic and metaphysical separation between different selves. Having said that, we need to be clear that, while the philosophers discussed might all accept the importance of this metaphorical 'gap', they nonetheless have different views about why and how this gap matters ethically.

On one view, the 'gap' is to be understood epistemically. On this understanding, transformative experiences lead to ethical problems because of the epistemic barriers they create between our earlier and later selves or between our self and some other person's self. For example, these experiences can make it hard for us to know what would be best for someone in the future, morally speaking. In these cases, the relevant person has an old self and a new self, each self

possessing different preferences. Crucially, as emphasized by Akhlaghi (2023) for example, we cannot know in advance what preferences the future self will possess. Hence, we seem unable to project or predict what would be best for this person, morally speaking. With this in mind, we can see why Akhlaghi cautions us not to focus on a future person's best interests when interfering with someone's decision to undergo a transformative experience. On his view, we cannot always know what will be in a future person's best interests (Akhlaghi 2023, 4-5).

Transformative experiences can also make it hard for us to acquire knowledge about the experience of someone who is in a different epistemic state from our current self. For example, Woollard raises a question about how we can evaluate a pregnancy experience of another person's future self when that future self is in a completely different epistemic state from our current self (Woollard 2021, 157-160).

On another view, the 'gap' is to be understood metaphysically.[20] In Chan's (2019) terms, we might say that transformative experiences break the continuity of our 'practical identity', thus effectively creating two selves (Chan 2019, 7). By separating our old selves from our new selves, transformative experiences can make it unclear who our moral subject should be. This, I suggest, is another source of ethical problems in the papers discussed above. For example, Howard argues that we cannot justify making someone else undergo a transformative experience simply by considering what their *present* self would say about this decision. Nor, on Howard's view, can we rely uncritically on what the *future* self of the relevant person will endorse (Howard 2015, 341-342; 360-361). Finally, Das and Paul argue that the separation of our earlier and later selves can also prevent us from justifying the decisions that we take to be intuitively rational, just as the non-identity problem in population ethics seems to prevent us from justifying decisions that we take to be intuitively morally correct (Das and Paul 2020, 196-197).

---

[20] As I mentioned above in section 3, the term 'metaphysical' is not to be understood as referring to numerical identity in this context. Instead, it refers to the 'selves' that form our ethical identity.

So, the overall point I wish to make in this subsection is that the 'gap' between our current and our future selves created by transformative experiences plays an ethically important role. Different philosophers offer different ways of explaining how and why this gap matters ethically. In the remainder of this thesis, this general point — that the epistemic and metaphysical separation of selves by transformative experiences raises important ethical issues — will play an important role. I will discuss one new puzzle in which it is unclear which moral unit our notion of *moral regret* should focus on. In addition, I will discuss two further new puzzles which arise because the separation of old and new selves leads to problems for our notions of *moral evaluation* and *moral commitment*. Finally, in the last chapter of this thesis, I will use this point to start a broader discussion around how transformative experiences make it difficult for us to identify who our subject of moral concern should be.

6. Transformative Experience in a Broader Ethical Context

In the previous section, I discussed four papers dealing with questions about morally right and wrong action in relation to transformative experiences. In this section, I will focus on two papers that emphasize ethical questions that go beyond the issue of morally right and wrong action.

6.1. Srinivasan on Authenticity

In 2015 Amia Srinivasan published a brief review of Paul's book on transformative experience in the *Times Literary Supplement*. In her review, Srinivasan briefly summarises and assesses Paul's main decision-theoretic points. In addition, and more importantly for our purposes, Srinivasan also briefly discusses some potential ethical implications of Paul's ideas (Srinivasan 2015, 8). However, Srinivasan understands 'ethics' broadly in her paper, and not as limited to questions about what the morally right or wrong action is. Instead, she includes questions about well-being, existence,

authenticity and what we should value in life within a broader concept of 'ethics' (Srinivasan 2015, 8).[21]

On Srinivasan's interpretation, Paul has conveyed a deeper message which goes beyond the decision-theoretic puzzle raised by transformative experiences (Srinivasan 2015, 8). Recall that, in the case of transformative experiences, we usually do not know whether we will enjoy our future state of being and, in the case of personally transformative events, what our new self will be like. Srinivasan argues that, on Paul's view, this insight about transformative experiences conflicts with how humans typically want to take important decisions (Srinivasan 2015, 8).

Specifically, on this view, humans typically do not want to rely on third-party evidence or testimony about relevant outcomes when taking important life decisions. Instead, on Srinivasan's interpretation of Paul (2014), humans want to 'own' their decisions by projecting themselves into their expected future self (Srinivasan 2015, 8). As I explained in section 2, Paul (2014) seems to believe that, if we have to make decisions that deeply affect our personal lives, we need to feel an intimate and strong connection to this decision, and this is the reason why we want to be able to project ourselves into the potential outcomes of our decision.

Of course, this need to own our decisions seems problematic in the case of transformative experiences because we cannot rely on our own evidence or project ourselves into our expected future self. Srinivasan continues by arguing that, according to Paul (2014), this leaves humans in a difficult predicament. In the case of transformative experiences, humans cannot make important decisions in the usual way. But if humans were to rely instead on third parties to make these important decisions, they would be inauthentic or 'self-alienated' from their choices (Srinivasan 2015, 8).

---

[21] Given this broader definition of 'ethics', I discuss Srinivasan's paper in this section rather than in the previous section.

It is not fully clear from Srinivasan's paper what the consequences for humans are if they choose inauthentically. Srinivasan writes that on Paul's view such choices are 'existentially problematic' because they would be 'an act of self-alienation' (Srinivasan 2015, 8). Presumably, this means that by relying on third-party evidence, we cannot 'own' these decisions in the way we would want. The idea seems to be that, given the life-changing importance of the decisions, we want to feel an intimate and strong connection to them that we cannot have by merely relying on third parties.

In response to what she takes to be Paul's deeper message, Srinivasan makes two ethical remarks, one about well-being and authenticity and one about ethical decision-making (Srinivasan 2015, 8).

First, she expresses some scepticism about Paul's concept of authenticity. Srinivasan argues that Paul's (2014) concept of authenticity assumes that humans are so different from each other that they could not possibly rely on each other's evidence to make important life decisions when undergoing transformative experiences (Srinivasan 2015, 8). However, on Srinivasan's view, this is implausible. In contrast to Paul (2014), Srinivasan suggests that authenticity is better understood as admitting to ourselves that we are not as different from other people as we sometimes believe (Srinivasan 2015, 8). In other words, on this view, being authentic is not believing that we are special but rather recognising that we resemble other people more than we often want to admit. And, Srinivasan argues, if we are not as different from other people as Paul (2014) thinks, then transformative experiences lose some of their importance. After all, if we are more similar to others than Paul (2014) thinks, then we can rely on other people's testimonial evidence to make important decisions, even in the case of transformative experiences (Srinivasan 2015, 8).

I find Srinivasan's first reply to Paul (2014) only partially convincing. It is surely correct to recognise that humans sometimes have a tendency to overestimate their own uniqueness. Having said that, this should not make us overlook the radical character that some transformative experiences can have. In some of the next chapters of this thesis I will discuss transformative

experiences such as being a refugee or living through a historical tragedy like the Holocaust. These kinds of transformative experiences strike me as radically different from many people's lived experiences. If we are faced with making a decision in these phenomenologically unique cases, it is not straightforward that we can draw on the testimony of other people, particularly those who have never been through such a radical experience.

So, I think that Srinivasan's point is a good one, but it can be put more cautiously. It is true that we sometimes overestimate our own uniqueness and we can therefore more often rely on other people's testimony than we might think. However, we should also not forget about the radical transformative experiences which epistemically and personally separate different people's lived experiences. When faced with decisions in these cases, it seems overly optimistic to argue that we can always rely on other people's testimony to find out how we ought to act.

Second, even assuming Paul (2014) is right about authenticity, Srinivasan raises a further question about how we make decisions in the face of transformative experiences. Srinivasan starts by arguing that personally transformative experiences sometimes equip us with a new self and we can sometimes know, in advance, what this new self will be. She then asks us to consider a case in which this new self holds values which conflict with the values we currently hold. This case leads Srinivasan to the following question: How should we value the selves we become, particularly if these selves hold values that conflict with the values we currently hold (Srinivasan 2015, 8)?

Srinivasan does not attempt to answer this question herself. However, she does caution the reader that 'the significant issue here is not about the rationality of decision-making, but about the ethics of making decisions rationally' (Srinivasan 2015, 8). In other words, according to Srinivasan, the main issue here is not, as Paul (2014) suggested, that we can never know what we will be like after our transformative experience and hence cannot take decisions rationally. Instead, on Srinivasan's view, the main issue is about how we can adjudicate between two sets of conflicting values (Srinivasan 2015, 8). The difficulty resides in the fact that we currently hold one set of those values but we also know that we will come to hold a conflicting set of values in the future, if we

make the relevant decision. According to Srinivasan, decision theory cannot tell us which of these sets of values it is appropriate to hold, and hence it cannot help with this choice. Instead, we need ethics to tell us what a person in these circumstances should do (Srinivasan 2015, 8).

6.2. Elizabeth Barnes on Social Identities and Transformative Experiences

Barnes (2015) starts her paper by making three observations about transformative experiences and their social context (Barnes 2015, 171).

6.2.1. Barnes' Three Observations

First, according to Barnes, it is important to note that the 'transformativeness' of a particular transformative experience is contingent. Barnes argues that whether a particular experience is transformative depends not just on the experience itself but also on factors external to the experience. On her view, there are two factors external to the experience itself which can influence whether this experience is transformative or not (Barnes 2015, 171-172).

The first factor is a person's character or 'what sort of person you are' (Barnes 2015, 172). In other words, on this view, your psychological dispositions, inclinations, etc. can influence whether or not a particular experience is transformative for you or not. For example, Barnes wants us to consider the case of people who receive their first pet. Before receiving a pet, many of these people did not feel any emotional attachment to animals. However, once they have their new pet, many of these people are emotionally strongly attached to their pet and they even organise their entire lives around their pets. On Barnes' view, this is good evidence that acquiring a pet has been a transformative experience for these people (Barnes 2015, 172).

However, Barnes argues that this transformation does not happen for every new pet-owner. Instead, whether or not this experience is transformative depends on the pet-owner's psychology. If the pet-owner already had, say, animal-friendly inclinations, then the experience is

much more likely to be transformative than if the pet-owner never had any such inclinations (Barnes 2015, 172).

The second factor is a person's social context. The main idea here is that the social structure in which you exist influences whether an experience can be transformative or not. Concretely, Barnes asks us to imagine a traditional and hierarchical society. Barnes argues that being poor or being rich in this kind of society shapes your perception of what is possible and what is not. If you are rich in this society, Barnes suggests, you find it normal to have a lot of 'social and economic prospects' (Barnes 2015, 172). In contrast, if you are poor in this society, Barnes argues that you think it impossible to have any genuine social or economic prospects (Barnes 2015, 172).

Now imagine someone who is poor in this society and who suddenly acquires a lot of wealth. On Barnes' view, this experience is likely to be very transformative. After all, transitioning from a financially very poor life to a wealthy life constitutes a radical transformation. However, the main reason why this experience is so transformative is the social context in which the poor person existed. The acquisition of money and, more importantly, the *prospects* that this acquisition opens up are highly likely to make this experience transformative (Barnes 2015, 172). The fact that there is such a big gap in wealth and related expectations and prospects between the different layers of society makes the transition from poor to wealthy a transformative experience. If the wealth gap was not as significant, then, on Barnes' view, the experience would probably also be much less transformative (Barnes 2015, 172).

Second, Barnes argues that the transformativeness of different kinds of experience comes in degrees. On her view, different types of experience can be more or less epistemically or personally transformative both in terms of how much they change a person and in terms of how strongly they change a person (Barnes 2015, 173). Barnes supports her claim with, amongst others, the following two examples.

Barnes argues that becoming a God (as happens to the main character of NK Jemisin's novel *The Hundred Thousand Kingdoms*) and trying a completely new soft drink (such as Irn Bru) both count as epistemically transformative experiences. However, she also argues that the two experiences are not epistemically transformative to the same extent. After all, becoming a God gives you access to much more new phenomenological information than trying a new soft drink (Barnes 2015, 173).

Barnes' second example compares her own initiation to philosophy with St Paul's religious experience in the Book of Acts. According to Barnes, both experiences can be personally transformative — that is, change one's core desires, commitments, etc. However, she also suggests that St Paul's experience is likely to be more strongly personally transformative because it has a more profound impact on St Paul than philosophy has on Barnes herself (Barnes 2015, 173-174).

Third, Barnes argues that the way in which people are epistemically and personally transformed by certain experiences can vary. While, on this view, an experience can be transformative for two people, it is often the case that the experience does not transform the two people in the same way. On Barnes' view, one of the two people might acquire more or less phenomenological information than the other or might be transformed to a greater or less extent than the other (Barnes 2015, 174-175).

For example, Barnes asks us to consider the case of two people who experience and survive horrible car accidents. For both people, this experience was personally transformative. However, Barnes claims that it could have been transformative in a different way for each of them. Maybe, one of these people will be transformed in such a way that they now believe that life is about taking risks and enjoying it to the fullest. In contrast, maybe the other person now believes that life is about being more careful, risk-averse and organised. In any case, on Barnes' view, it is likely that the same transformative experience often transforms people who undergo it in different ways (Barnes 2015, 175).

6.2.2. Transformative Experience and Social Justice

Barnes wants to use her three observations above to show that 'whether, how, and to what extent an experience is transformative can sometimes be a matter of social justice.' (Barnes 2015, 171). In order to do that, she will focus mainly on how the social structures we live in either *facilitate* or *hinder* the existence of transformative experiences.

Barnes starts by arguing that social conditions can either be 'easy' or 'hard' for transformative experiences (Barnes 2015, 175). Social conditions are 'easy' for certain types of transformative experience if, under these conditions, these types of experience are often or always transformative and, in the absence of these conditions, these types of experience do not tend to be transformative (Barnes 2015, 176).[22] For example, recall the earlier case of a young poor person's transformation into a wealthy person. In this kind of case, the social structures of the time — a hierarchical society characterised by a strong wealth gap — facilitate or make it easy for the young person's journey to be transformative. If the society were more egalitarian, then it would be very hard for an acquisition of wealth to be transformative because people had similar living standards to begin with.

It is not clear to me that this particular example can support Barnes' wider point. Suppose you live in an egalitarian society and you suddenly acquire a lot of wealth. This could mean one of two things for you.

First, it could mean that you become much richer than everyone else in this egalitarian society. If this is how we are supposed to understand Barnes' example, it is hard to see why this acquisition of wealth could not be transformative. After all, your sudden acquisition of wealth transforms your life from a life that is similar to everyone else's life in the egalitarian society to a life that is radically different.

---

[22] Barnes uses the concept of 'nearby worlds' for her definition (Barnes 2015, 176).

Second, your acquisition of wealth could mean that you are now as rich as the 'richest' people in the egalitarian society. However, given that the society is egalitarian to start with, this acquisition is not a dramatic change for you. Surely, this acquisition would not be very transformative because your initial economic position has not changed a lot after your wealth acquisition. In other words, it seems that what prevents this acquisition of wealth from being transformative is not the social context as such. Instead, it is the fact that you acquire only very little wealth compared to your initial starting point.

So, it seems that Barnes' example can be read in two ways, each of which is problematic for her wider point. Either your wealth acquisition makes you much richer than everyone else and then there is no reason why this experience should not be transformative, even in an egalitarian society. Or your wealth acquisition is not really a 'wealth' acquisition at all and this is the reason (rather than the social context) why the experience is not transformative.

There might be an alternative example which illustrates Barnes' point more effectively and in the same spirit. Suppose you live in a traditional early modern society in which travel and information sharing are much less developed than today. If you were to suddenly move to a new part of the country, it is likely that this experience could be much more transformative than it would be today.[23] The reason for this might be, as Barnes suggests, the social context in which your relocation takes place. As I noted above, given the lack of travel opportunities and the comparatively few ways of acquiring information about new parts of a country in early modern times, you would not have access to a lot or any information about your destination. It is likely that moving to a different part of the country would be therefore a much more radical, and hence potentially transformative, experience than today.

---

[23] This is not to say that relocation experiences can never be transformative in today's world. As I showed in *Charity Abroad*, there are clear cases where relocation is a profoundly transformative experience. Having said that, the context of an early modern society might make it even 'easier' for these experiences to be transformative.

As I just noted, Barnes' example about 'easy' social conditions might not be entirely convincing and an alternative example might be more effective at illustrating her point. Having said that, Barnes offers a different, more convincing example in her discussion of social conditions that can make it 'hard' for certain types of experience to be transformative. On her view, conditions make it 'hard' if, under these conditions, these types of experience do not tend to be transformative and, in the absence of these conditions, these types of experience tend to be transformative more often (Barnes 2015, 177). For example, Barnes gives us the case of a pregnant seventeen-year-old girl in an agrarian society in the 1800's. On Barnes' view, social conditions make it hard for this girl to experience her pregnancy as transformative or transformative to the same extent as someone would today. The reason for this might be, for example, that, in an agrarian society in the 1800's, pregnancy was not imbued with the same meaning or significance as it is today (Barnes 2015, 177-178).

According to Barnes, social conditions can make it not only 'easy' or 'hard' for experiences to be transformative but also 'too easy' or 'too hard' (Barnes 2015, 179, 181).

Social conditions make it 'too easy' for an experience to be transformative when they make certain experiences transformative that should not be transformative or, at least, not in this way. On Barnes' view, these social conditions are morally harmful because they have a potentially negative impact on the affected people's well-being (Barnes 2015, 179). For example, Barnes argues that patriarchal social conditions can make marriage too easily transformative for women. By getting married, these women experience a radical transformation whereby they are forced into distinct and rigid roles and responsibilities that they did not experience before, or at least, not to the same extent (Barnes 2015, 179-180).

Barnes argues further that social conditions make it 'too hard' for an experience to be transformative when they do not allow an experience to be transformative for someone or some group of people or, at least, not in a way that it should be transformative. Again, on Barnes' view, this constitutes a moral harm because of the potentially negative impact on the affected people's

well-being (Barnes 2015, 181). For example, consider the case of someone who has recently become disabled. According to Barnes, becoming disabled is always epistemically transformative, that is, you acquire some phenomenological knowledge about what it is like to be disabled that you would not have acquired if you had not become disabled (Barnes 2015, 181). In addition, Barnes argues that becoming disabled can also be positively personally transformative. By becoming disabled you gain new capacities and sensibilities and you often become part of a new and supportive community. However, according to Barnes, current social norms often make it too hard for this positive personal transformation to take place. On Barnes' view, many societies still view a disability as something to be overcome rather than embraced. This makes it hard for many disabled people to undergo a personal transformation which results in a positive and confident attitude about their own disability (Barnes 2015, 181-182).

Barnes concludes her argument by noting an important connection between her remarks about transformative experiences and social identities (Barnes 2015, 185-186). As already noted by Paul (2014), (personally) transformative experiences transform our self — that is, who we are and how we think of ourselves. Crucially, as Barnes (2015) explained above, this transformation always happens in a social context. This context, through its norms, customs and conventions, influences what ways of thinking about ourselves are available and which ones are not. Sometimes, the availability of these different self-conceptions becomes a matter of social justice. By making some self-conceptions too easily available or too hard to come by, Barnes argues that many societies risk harming individuals or groups of people. On Barnes' view, then, one key moral lesson to draw from this is that, as a society, we should try to make some self-conceptions less readily available and some self-conceptions much more easily available (Barnes 2015, 185-186).

6.3. A General Remark

Before concluding, I want to draw out one general point from the two presentations I have made in this section. When discussing transformative experiences, we are sometimes tempted to think

of these experiences as being all *radically* transformative in the *same way* for everyone. However, both Srinivasan and Barnes try to nuance this general characterisation of transformative experiences. They do so in different ways.

On Srinivasan's view, we should be cautious not to overestimate the apparently *radical* nature of transformative experiences. On her view, we humans are all much more similar to each other than we might think. As a result, Srinivasan argues that transformative experiences should be seen as less 'radical' in their way of separating different people's lived experiences from each other (Srinivasan 2015, 8). In contrast, on Barnes' view, transformative experiences are often radical but *not in the same way* for everyone. As noted above, Barnes argues that the social context in which we find ourselves strongly influences in what way and how easily transformative experiences shape our lived experience (Barnes 2015, 171).

As I already noted above in subsection 5.1, I will disagree to an extent with Srinivasan's view in the remaining chapters. Specifically, I will discuss cases of radical transformative experiences in which we cannot simply draw on testimony from other people to make our own decisions. In addition, one of my chapters, which focuses on moral commitments, will take seriously Barnes' idea that we are not all transformed in the same way. I will argue that transformative experiences can be so radically different for different people that we cannot know what our values and commitments will be after going through these experiences.

Conclusion

Some of our decisions transform us. They might transform us epistemically by providing us with phenomenological knowledge that we would not otherwise have acquired. They might also transform us on a personal level by radically reshaping our core values, desires, preferences, etc. According to Paul (2014), the transformative nature of these decisions undermines our traditional way of making decisions. I have tried to show in this chapter that, in addition, transformative experiences raise a host of ethical questions.

I have presented the main existing contributions to the ethics of transformative experience. The contributions I have discussed fall into two camps.

First, there are papers discussing issues about the moral rightness and wrongness of actions in the context of transformative experiences.

Farbod Akhlaghi's (2023) paper starts with an important insight, namely, that transformative choices can be strongly influenced by the people surrounding the decision-maker. It seems right on Akhlaghi's part, therefore, to focus on the moral conditions under which third parties are morally justified in interfering with the decision-maker's transformative choice. Akhlaghi answers his main question by arguing that we all have a right to revelatory autonomy and that we can interfere with someone else's choice to undergo a transformative experience only if this right is outweighed by other moral considerations (Akhlaghi 2023, 9). One feature to highlight about Akhlaghi's paper is that he manages to give an answer to his main question that avoids the issue that we do not possess the same self before and after transformative experiences. On his view, it is the self before the decision which possesses the right to revelatory autonomy and this is all we need to know about this self to assess their choice.

In response to Akhlaghi, I made two points. I suggested that the right to revelatory autonomy might sometimes not apply given the uncertainty surrounding transformative experiences, although I also suggested that there is a way for Akhlaghi to overcome this problem. More importantly, I also argued that it is unclear why we should possess a right to revelatory autonomy in the first place, and I suggested that this would be an interesting avenue for further inquiry.

The second paper I discussed is by Das and Paul (2020). One of the most interesting features of Das and Paul's paper is their insight that transformative experiences can lead to a new version of the non-identity problem that mirrors the traditional non-identity problem to a significant extent. The main potential problem with Das and Paul's puzzle, on my view, is that they need to assume that transformative experiences cannot be bad for someone's current self,

and, as a result, it seems that they are forced to assume that death cannot be bad for someone's current self. As noted above, this is an implausible conclusion because it would mean that we cannot rationally refuse to kill ourselves. However, in the next chapter of this thesis, I take the insight about the relationship between transformative experience and procreation further, and I present a new non-identity puzzle about moral regret in the context of transformative experiences.

I also discussed Dana Howard's (2015) paper in which she argues that we can be justified in making someone else undergo a transformative experience only if this other person endorses our decision in a non-adaptive way (Howard 2015, 367). While Howard seems right to insist on the moral importance of endorsing a decision, I also suggested that the main problem with her view is that it cannot distinguish morally problematic from morally unproblematic cases.

Finally, I briefly discussed Fiona Woollard's (2021, 2022) papers, which caution us to integrate the perspective of people who have gone through a transformative experience such as pregnancy into our moral evaluation of issues such as abortion. I will discuss these papers in much greater detail in Chapter Three of this thesis.

Overall, the main lesson I drew from these four papers was that the metaphorical 'gap' between old and new selves created by transformative experiences can be interpreted epistemically and metaphysically. While different philosophers interpret this gap in different ways, I also suggested that the 'gap' is often of central importance for ethical issues arising in this context.

Second, I presented two papers discussing ethical issues linked to transformative experiences that go beyond questions about the moral rightness or wrongness of particular actions. Amia Srinivasan (2015) presents a response to what she takes to be Paul's deeper message about how authentic choices are choices in which we 'own' our decisions. On Srinivasan's view, Paul overestimates the degree to which we are unique, and there is a clear need, on her view, for ethics to adjudicate between the conflicting values that transformative experiences can bring about (Srinivasan 2015, 8). In response to Srinivasan, I expressed some scepticism about her claim that humans overestimate their uniqueness by arguing that some transformative experiences are so

radical that we cannot rely on other people's testimony. However, I share her view that the change in values brought about by transformative experiences is morally relevant, and in Chapter Four I will discuss how this can be morally relevant for our long-term moral commitments.

Finally, Elizabeth Barnes (2015) cautions us, and rightly so on my view, to be mindful about the social context in which transformative experiences occur. Often, on her view, whether and how experiences transform us depends on broader social factors that go beyond the individual decision-maker (Barnes 2015, 171). In response to Barnes, I briefly suggested that a key example used to support her wider point about the importance of social context in facilitating or hindering transformativeness cannot actually support this point after all. Having said that, I then suggested a different example which might make Barnes' point more effectively.

Overall, I took the main lesson from the papers by Srinivasan and Barnes to be that we need to be careful not to overgeneralise the phenomenon of transformative experiences. Sometimes, we are more similar to others than we think and sometimes the social context in which we operate strongly shapes how our lives are transformed.

In the remainder of this thesis my aim will be to contribute to the emerging literature that I have presented in this chapter. As I mentioned in the introduction to this thesis, the next three chapters can be read as distinct contributions to the literature which are united by the common goal of expanding the moral debate surrounding transformative experiences. I will develop three new ethical puzzles which discuss moral topics that have not been directly addressed in the context of transformative experiences.

In the next chapter, I will raise a puzzle about transformative experiences and their potential to undermine our ability to regret historical tragedies. In the third chapter, I will build on Woollard's (2021, 2022) papers to raise a puzzle about our seeming inability to morally evaluate some transformative experiences that we have not gone through ourselves. In the fourth chapter, I will discuss the counterintuitive impact that transformative experiences can have on our long-term moral commitments.

Chapter Two:
Transformative Tragedies

Introduction

The past contains many horrible tragedies. Think, for instance, of the Holocaust or the slave trade. Naturally, we wish the past had been better for the victims of these tragedies. However, several philosophers have recently argued that there is a fundamental difficulty with this intuitive attitude.

Guy Kahane (2019; 2021) and Saul Smilansky (2020) argue roughly as follows: Wishing that the victims of these tragedies had not suffered requires wishing that these tragedies never occurred. This, in turn, requires wishing away earlier events that led to these tragedies. But, at least in many cases, these earlier events seem to be necessary parts of the causal chain leading to the existence of the victims of these tragedies.[24] So, wishing that the victims of these tragedies had better lives seems to require wishing these victims out of existence. This seems contrary to our moral concern for these victims, and it might even be incoherent. So, it seems to be morally inappropriate, or even incoherent, to wish that the past had been better for the victims of these tragedies (Kahane 2019, 165-166; Kahane 2021, 15-17; Smilansky 2020, 189-195).

In the remainder of this chapter, I will refer to the above puzzle as *No Regrets*. It is worth noting the tight connection between *No Regrets* and another famous problem in population ethics, the so-called 'non-identity problem'.[25] Like the non-identity problem, *No Regrets* relies on the

---

[24] Below, I will explain and assess in more detail the kind of necessity here referred to by Kahane and Smilansky.

[25] For a classic discussion of the non-identity problem see Derek Parfit (1984) *Reasons and Persons* (Oxford: Clarendon Press), 351-377. For more recent discussions of the problem see, for example, David Boonin (2014), *The Non-Identity Problem and the Ethics of Future People* (Oxford and New York: Oxford University Press); Johann Frick (2020), 'Conditional Reasons and the Procreation Asymmetry', 53-87; Rahul Kumar (2018), 'Risking Future Generations', *Ethical Theory and Practice* 21: 245-257; Abelard Podgorski (2021), 'Partiality, Identity, and Procreation', *Philosophy and Public Affairs* 49 (1): 51-77.

assumption that if a person had not been conceived at the time that they were in fact conceived, then this person would in fact never have existed. But whereas the non-identity problem uses this assumption to make trouble for our intuitions about the moral status of possible future acts, *No Regrets* uses this assumption to make trouble for our intuitions about the moral appropriateness of past-directed attitudes. We might therefore think of *No Regrets* as a kind of *inverse* non-identity problem.

I begin this chapter by arguing that *No Regrets* is easily resolved by challenging the assumed necessary connection between earlier historical events and actual historical tragedies. However, I go on to argue that, at least if we accept some ideas from the literature on 'transformative experience' (Paul 2014, 15-16), there is a new puzzle in the vicinity of the old puzzle.

I will proceed as follows. In Section 1, I will discuss the notion of 'regret' in more detail by distinguishing between agent-regret, evaluator-regret and what I will call evaluator-regret*. In Section 2, I will present *No Regrets* as it is usually presented in the literature, and I will raise a problem with the traditional set-up of the puzzle. In Section 3, I will argue that, if we accept some ideas from the literature on 'transformative experience' (Paul 2014, 15-16), there is a new puzzle in the vicinity of *No Regrets* which has not been discussed before. In Section 4, I will present three possible solutions to the new puzzle. Ultimately, I argue that none of these solutions offers a satisfactory way forward. I conclude that we must either look for a new solution to the new puzzle, or reject the ideas about transformative experience on which the puzzle depends.

1. Regret

Before discussing *No Regrets* as it is usually presented in the literature, I wish to get clearer on the notion of 'regret'. What does it mean to regret something and, more specifically, what does it mean to regret something on someone else's behalf, as is the case when we regret a tragedy on behalf of its victims?

We can start by distinguishing two kinds of regret. We can regret our own past actions or deliberations and we can regret an outcome brought about by someone (or something) else. In the literature, Carla Bagnoli (2000) refers to these kinds of regret as 'agent-regret' and 'evaluator-regret', respectively (Bagnoli 2000, 176). In the literature, many contributions have focused mainly on the workings and significance of 'agent-regret' — that is, on how we can regret our own past actions and deliberations.[26] Given this focus, I will also start my discussion with agent-regret. My goal is not to present my own view of agent-regret. Instead, I will analyse whether discussions of agent-regret in the literature can give us any insights about how to understand the less-discussed concept of evaluator-regret, a revised version of which I need for this chapter.

1.1. Agent-Regret

One tempting view of agent-regret is that we regret something because we made a rational or moral mistake in our past deliberation or in our past action. On this view, we wish that we had not made a mistake and that a particular decision or event had not taken place as a result. This view of agent-regret works reasonably well for some cases. For example, if I forgot to keep my promise to you and, as a result you endured some hardship, it might seem that I should regret what happened to you because I made a (moral) mistake in not keeping my promise.

However, in the literature, most philosophers seem to agree that the 'mistake' view of agent-regret is not the right way to go (e.g. Bagnoli 2000, 170-172; Setiya 2016, 1-2, 14). There seem to be clear cases in which we regret our decision or action even though we did not make a

---

[26] See, for example, Carla Bagnoli (2000), 'Value in the Guise of Regret', *Philosophical Explorations* 3 (2): 169-187; R. J. Wallace (2013), *The View From Here: On Affirmation, Attachment, and the Limits of Regret* (Oxford and New York: Oxford University Press); Elizabeth Harman (2009), 'Transformative Experiences and Reliance on Moral Testimony', *Res Philosophica*, 92 (2): 323-339; Camil Golub (2019), 'Personal Value, Biographical Identity, and Retrospective Attitudes', *Australasian Journal of Philosophy* 97 (1): 72-85.

rational or moral mistake. For instance, according to Kieran Setiya (2016), we can regret not having pursued a particular career path even though we are satisfied with our current career. In other words, we did not make a mistake when choosing our current career. However, on Setiya's view, it can still be legitimate to feel a sense of regret for a valuable career we did not choose in the past (Setiya 2016, 2-3).

If agent-regret is not based on us making a mistake, how should we understand it?

First, agent-regret is *retrospective* and *counterfactual*. Agent-regret is 'retrospective' not only because it is about the past but also because it concerns an attitude of 'looking back' at our own past life (Wallace 2013, 10; Setiya 2016, 3). In addition, agent-regret is 'counterfactual' because, in regretting, we wish that life as it happened did not happen the way it actually happened. In other words, we wish for a scenario in which things would have turned out differently.

Second, agent-regret seems to rely on the recognition that there was something *valuable* in the alternative scenarios that did not come about. If we did not recognise any value in the alternative that did not come about, we would not feel any sense of regret (Bagnoli 2000, 177-180). For instance, we regret the fact that we did not keep our promise precisely because we believe that there is something of value in promise-keeping. Similarly, sometimes we need to choose between two competing kinds of value and we feel regret precisely because we know that the alternative we did not choose also carried value, albeit of a different kind (Setiya 2016, 3). For example, suppose you choose helping your friend over looking out for your own interests in a particular situation. In this kind of case, looking out for your own interests carries practical/instrumental value that conflicts with the moral value of helping out your friend. When you regret not having pursued your own interests, you do not think that you made a mistake by helping your friend. Instead, you are simply alive to the fact that there was something of value to the option you did not choose.

Finally, notice also that feeling agent-regret does not always mean that we have an *overall preference* for the unchosen option. For example, in Setiya's (2016) example discussed above, you might feel a sense of regret for a career path you did not choose, but that does not necessarily

mean that you have an overall preference against your current career and in favour of the alternative career you did not choose (Setiya 2016, 2).

1.2. Evaluator-Regret and Evaluator-Regret*

What insights can that discussion of agent-regret give us about how we ought to understand evaluator-regret?

There are clear similarities between the two types of regret.

First, again, it seems reasonable to suppose that evaluator-regret is not based on the assumption that we regret a particular outcome because someone made a mistake. For example, if we regret that your action has bad unintended consequences that you could not have foreseen, we do not think that you made a rational or moral mistake. After all, the outcome was not the result of any mistake on your part.

In addition, both types of regret also seem to be counterfactual. It seems that, as an evaluator, we also wish that things had turned out differently than they actually did. As in the case of agent-regret, we also believe that, in the case of evaluator-regret, something of value was lost in the alternative outcome that did not actually happen. For instance, we might believe some alternative outcome would have led to more overall well-being.

Despite their similarities, there also seems to be at least one key difference between agent-regret and evaluator-regret. According to Bagnoli (2000), the object of agent-regret is different from the object of evaluator-regret. Bagnoli writes that 'the agent regrets his action or his deliberation, while the evaluator regrets some state of affairs brought about by somebody else' (Bagnoli 2000, 176). In other words, while the agent's regret seems to focus more on the process by which a particular outcome came about, the evaluator regrets the outcome itself. This distinction seems to fit common cases of regret. If I regret that my actions put me in harm's way, I tend to regret the fact that I did not choose a different *action* at the time. However, if I regret that

your actions put you in harm's way, it seems that I tend to regret the *outcome* of your choice, i.e. that you were put in harm's way.

We can usefully extend Bagnoli's concept of evaluator-regret for the purposes of this chapter. Let me refer to this extended version of Bagnoli's concept as evaluator-regret*. On this extended version, we regret not only a state of affairs brought about by someone else. Instead, we regret a state of affairs brought about by someone else *on behalf of* the people affected by this state of affairs. Concretely, for example, we regret a historical tragedy on behalf of its victims. While the object of evaluator-regret* is slightly different than the object of evaluator-regret, the former is characterised by similar features as the latter.

When we regret an outcome on someone else's behalf we cannot always assume that the people affected by the outcome made a mistake. For example, if we regret that a person became a victim of a historical tragedy, we do not think that they made a rational or moral mistake in becoming a victim. After all, people who become victims of historical tragedies usually do so through no fault of their own.

It also seems that, in the case of evaluator-regret*, we wish that things had turned out differently than they actually did for the relevant person. We also believe that, in the case of evaluator-regret*, something of value was lost in the alternative outcome that did not actually happen. For instance, we might believe the person would have enjoyed a more valuable life in terms of physical or mental well-being.

For the remainder of this chapter, I will focus on evaluator-regret* and the question of how we can, as evaluators, regret historical tragedies on behalf of their victims.

2. *No Regrets* and the Past

In this section, I will first outline *No Regrets* in the way the literature usually presents it, and I will then raise a problem for this presentation.

2.1. *No Regrets* in the Literature

Events such as the Holocaust are terrible because they were bad for particular people. We wish that these particular people had better lives—that is, that they did not have to endure any of the terrible experiences they actually endured. In other words, we wish that the tragedies they endured never took place.

According to Kahane (2019) and Smilansky (2020), the only way for certain later historical events not to have occurred is for certain earlier historical events also not to have occurred. They infer that, to wish away certain later historical events, we must also wish away some earlier historical events. For example, they suggest that, in order to wish away the Holocaust, we need to wish away World War One, given that there probably would never have been a Holocaust without World War One (Kahane 2019, 163-166; Smilansky 2020, 188-189).

But by wishing to change the past in this way, they argue, we also implicitly or explicitly wish to upset the causal patterns leading to the existence of the particular victims of these tragedies (Kahane 2019, 164; Smilansky 2020, 190). For, arguably, if World War One had not taken place, many people who actually met would not have met, and consequently they would not have conceived the children they actually conceived. And even if the same people would still have met, they almost certainly would not have conceived children at the same time they actually conceived. If the same people had not met, or if they had conceived at different times, they would not have conceived the particular Holocaust victims that they actually conceived. Instead, numerically distinct people would have existed.[27]

According to Kahane (2019) and Smilansky (2020), if the prior reasoning is correct, it is morally inappropriate to wish that these tragedies had not occurred. On their view, wishing away these tragedies also requires us to wish away the earlier historical events which led to them.

---

[27] The thought here is that actual time of conception influences numerical identity because it determines the specific sperm-egg combination which gives rise to a particular individual.

However, according to the reasoning above, wishing away these earlier events also implies wishing away these tragedies' actual victims. And wishing away these tragedies' actual victims seems contrary to our intuitive moral concern for these victims (Kahane 2019, 163-166; Smilansky 2020, 188-189).

Smilansky (2020) even goes so far as to argue that, from a purely self-interested perspective, we should *prefer* that these historical tragedies occurred. He writes:

'From the strictly personal perspective, then, we have good reason not to regret these tragic calamities. Had they not occurred, the conditions realistically necessary for *our* being born would have been precluded. Minor changes would have sufficed to prevent our parents from meeting or, had they met and given birth to a child, this would have been a different child at a different time – someone else. Similarly for our grandparents, great-grandparents, and previous ancestors. And so, the omission of even some of these historical and natural tragedies might have precluded the fragile chain of events that, realistically, is necessary for our having come into being. … *From the strictly personal perspective, then, we should prefer that these historical and natural tragedies occurred.*'[28] (Smilansky 2020, 190-191)

There might be a way of pushing Kahane and Smilansky's puzzle even further. Wishing away the actual victims of historical tragedies seems not only morally inappropriate but also to generate a kind of incoherence. By wishing away historical tragedies, we intend to express our moral concern for these tragedies' victims. But, if we simultaneously wish these victims out of existence, we seem to both care and not care about these victims at the same time.

---

[28] My italics on 'our'.

2.2. A Problem with *No Regrets*

The main problem with the preceding way of presenting *No Regrets* is metaphysical in nature. Specifically, *No Regrets* assumes an implausible necessary connection between earlier historical events and the existence of particular people.

Recall that, according to *No Regrets*, we can wish away a historical tragedy only by wishing away some earlier events which led to this tragedy, and, if we wish away these earlier events, we also wish away the tragedy's particular victims. This reasoning assumes that there is a necessary connection between the earlier events we wish away and the existence of the victims of the historical tragedy. Specifically, it assumes that the earlier events are *metaphysically* necessary for the tragedy's victims to exist — that is, that there could not have been a world in which the earlier events occurred and the victims existed, yet the tragedy did not occur (e.g. Smilansky 2020, 190).

Why should we accept this assumption?

One way of motivating this assumption might be through a causally deterministic view of history. It might seem that if history is causally determined — that is, if each historical event is determined by the conditions and events immediately preceding it combined with the laws of nature (Hoefer 2016) — wishing away historical tragedies necessarily implies wishing away earlier events which led to these tragedies.

However, to many people it seems perfectly coherent to imagine that history could have been exactly or almost exactly the same up to a certain point, except for the fact that a particular tragedy did not take place. Most of our regrets operate on the assumption that history could have been the same or almost the same up to the point where we did or said something we wish we had not. For example, if I regret not having been to the gym on Monday, I do not wish that earlier events leading up to Monday had not taken place. Instead, I wish that history had been exactly or almost exactly the same up to Monday and that I had gone to the gym on the same day.

Suppose we eliminate the assumption that earlier historical events are necessary to prevent a historical tragedy. Instead, we assume that history could have been exactly or almost exactly the same except for the fact that the tragedy did not take place.

If history could have been roughly the same, the exact same people could have existed, no matter whether the tragedy took place or not. After all, given that earlier historical events remain unchanged, the causal chain leading to the existence of the particular people also remains unchanged.

If the exact same people could have existed, we are free to say that the historical tragedies were bad for their victims. They could have existed without suffering a tragedy. And, if the historical tragedy was bad for them, it is perfectly coherent and morally appropriate for us to regret their fate.

So, if we remove *No Regrets'* main assumption — that to wish away a historical tragedy we need to wish away earlier historical events and thus wish away the victims of the tragedy — we thereby dissolve the puzzle. We are free to regret historical tragedies because they were bad for particular people.

Both Kahane (2019) and Smilansky (2020) briefly address the objection I just outlined. According to Kahane (2019), 'attitudes like gladness, sadness, and regret operate under a constraint of realism' and any claim that history could have been exactly the same except for the fact that the historical tragedy did not take place is akin to 'wishful thinking' (Kahane 2019, 165). In other words, Kahane (2019) argues that some alternative courses of history — such as the one where history is exactly the same except for the fact that the Holocaust did not take place — are so metaphysically improbable that they do not merit moral consideration. On Kahane's (2019) view, the only way of wishing away a historical tragedy implies wishing away earlier events which led to this tragedy (Kahane 2019, 164-165).

In a similar spirit, Smilansky (2020) writes:

'Logically, it is not impossible that even if (say) World War I had never happened, everyone alive today would still have come into being. In some sense, then, we could regret that *this* possibility was not realized. Realistically, however, given that our existence and that of our loved ones depends upon such calamities […] the chances we would have still come to exist even without them are so slim as to be dismissible' (Smilansky 2020, 190).

I think we should not attach much weight to Kahane and Smilansky's appeal to 'realism'. As we noted in our earlier discussion of agent-regret, evaluator regret and evaluator-regret*, to regret means engaging in a kind of counterfactual valuing. In other words, when we regret some action, or deliberation, or state of affairs, we consider some alternative possible world in which that action, deliberation, or states of affairs does not obtain and we value this alternative possibility. So, what reason is there to think that we cannot regret a tragedy by valuing the metaphysically possible world that is like the actual world save for the occurrence of the tragedy?

Based on their 'realism' defence, Kahane and Smilansky might reply that the more 'distant' a possible world is from the actual world, the less reasonable it is to dwell on it, or to compare it to the actual world. But it seems equally plausible that, as the value of a possible world goes up relative to the value of the actual world, the more reasonable it is to dwell on it. And these considerations would then pull in different directions. So, even though possible worlds in which the relevant tragedies do not occur, and yet the victims of those tragedies still exist, are very distant from the actual world, the relative values of these worlds would make them worthy of consideration. And if these distant possible worlds are worthy of our consideration, then we are able to regret that they did not come about.

In addition, it is worth pointing out that Kahane and Smilansky's 'realism' constraint on regret seems to exclude too many instances of evaluator-regret*.[29] Surely, in some instances, when we regret something on someone else's behalf, we do not consider only nearby possible worlds. Instead, we also consider more distant possible worlds, however improbable they might be. For example, I might regret that a friend's desire to become a professional footballer was not fulfilled, even if, because that friend was nowhere near skilled enough to make it, the possible worlds in which he makes it are very distant indeed.

3. A New Puzzle

In the previous section, I argued that *No Regrets* does not undermine our intuitive moral attitude towards the past. The puzzle is not metaphysically robust enough to undermine our regret of past tragedies. In this section, I argue that, at least if we accept some ideas from the literature on transformative experience (Paul 2014, 10-16), there is a new puzzle in the vicinity of *No Regrets*.[30] In order to explain this new puzzle, I need to take a step back, explain in more detail what transformative experiences are and distinguish a couple of key questions about regretting historical tragedies.

---

[29] Camil Golub (2019) makes a similar point about David Velleman's (2015) view of agent-regret. On Velleman's (2015) view, it is rational to regret something in our past only if we have 'first-personal access' to the relevant past self. This excludes the possibility of regretting something that could have happened to one of our merely possible selves (Velleman 2015, 241-242). According to Golub (2019), Velleman's view of regret is too radical because it excludes too many instances of agent-regret. On Golub's view, there are many instances where we might regret the better possible lives we could have had (Golub 2019, 77).

[30] This new puzzle is inspired by a similar-but-different puzzle discussed in Nilanjan Das and L.A. Paul (2020) 'Transformative Choice and the Non-Identity Problem', in Sauchelli, Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (London and New York: Routledge), pp.187-208. While Das and Paul present a puzzle about *rationality* and *future*-directed *choices*, I present a puzzle about *ethics* and *past*-directed *attitudes*.

3.1. Transformative Experiences

Consider the following case discussed by L.A. Paul (2014):

> *Vampire*: Suppose you have the opportunity to become a vampire. Your life as a vampire would be totally different from your life as a human. For example, you would possess several superhuman powers (such as incredible speed) and you would experience a range of sensory experiences that humans cannot experience. All of your friends have also become vampires and they absolutely recommend it, and they would never want to become human again. Now, the main question facing you is whether you should seize the opportunity to become a vampire or not (Paul 2014, 1-4).

A natural way of approaching fundamentally life-changing decisions is to think about what it would be like for us to be in the new condition. Once we know what it would be like, it seems that we are in a good position to know whether we prefer the new condition to our current condition.

According to Paul (2014), you cannot approach your decision in *Vampire* in this natural way. There are two main reasons for this.

First, before undergoing the transformation, you cannot possibly know what it would be like to be a vampire. What it is like to be a vampire is so different from what it is like to be you that you cannot know what life as a vampire would be like.[31] As a result, your choice position is epistemically impoverished — you lack crucial knowledge about what it would be like to be a

---

[31] Notice that Paul (2014) is here referring to what it is like to be a vampire from a first-person perspective. While it is true that we can know what it is like to be a vampire from an impersonal point of view (i.e. a vampire has to drink blood, etc.), we cannot possibly know what it would be like for *us* to be a vampire (Paul 2014, 2).

vampire.[32] If you lack this kind of knowledge, you cannot make a rational decision about whether to become a vampire or not. After all, you cannot compare your human condition with your potential vampire condition because you do not know what the latter would be like.[33] So, it is neither rational for you to choose to become a vampire nor is it rational for you to choose to stay human (Paul 2014, 13-18).

Paul (2014) refers to choices such as *Vampire* as 'epistemically transformative' (Paul 2014, 15). The experience transforms your epistemic position so radically that you cannot possibly make a rational decision prior to undergoing the experience itself.

Second, Paul (2014) argues that becoming a vampire is also a 'personally transformative' experience (Paul 2014, 16). Who we are, that is, the set of beliefs, commitments, desires, etc. that form our self, depends in part on our experience of what it is like to be us. If we change enough of what it is like to be us, we also change who we are. In the *Vampire* case, what it is like to be us changes dramatically. As a result, our self prior to the transformation would not be the same self as the one after the transformation.

The fact that the *Vampire* case is personally transformative also has a decision-theoretic upshot. Suppose that you somehow know what it would be like for you to be a vampire later on. Imagine further that you quite enjoy the vision of yourself as a vampire, that is, you prefer becoming a vampire to staying human. Still, even with this information, it seems you could not make a rational decision on whether to become a vampire or not. After becoming a vampire, your

---

[32] According to Paul (2014), it is important to note that relying on your friends' testimony is not sufficient for you to know whether you would like to become a vampire or not. While you can draw on your friends' testimony to improve your epistemic position, this will not be enough to bridge the epistemic gap which separates your human experience from the vampire experience (Paul 2014, 2-3).

[33] Paul is here assuming a standard decision-theoretic model according to which, to make a rational decision, you must assign values and probabilities to different outcomes and then choose the option which maximises expected value (Paul 2014, 21-22).

preferences might change radically. As a result, it could be the case that you no longer wish to be a vampire. In other words, your prior preferences might not match your later ones. Since you cannot know in advance how the experience is going to change your preferences, your initial preferences cannot act as a guide for your decision. After all, your preferences are highly likely to change after undergoing the experience (Paul 2014, 16-19; 21-22).

In the remainder of this chapter, my focus will be on personally transformative experiences.

3.2. Historical Tragedies as Personally Transformative Events

Consider some of Paul's real-life examples of personally (and epistemically) transformative experiences. She writes that 'experiencing a horrific physical attack, gaining a new sensory ability, having a traumatic accident […] participating in a revolution, having a religious conversion, having a child, experiencing the death of a parent, making a major scientific discovery' (Paul 2014, 16) can all count as personally transformative experiences. These experiences have the potential to transform a person's set of core beliefs, commitments, preferences, etc. to such an extent that the person's self prior to the experience is not the same as their self after the experience (Das and Paul 2020, 187-195).

Given the kind of examples of personally transformative experiences that Paul refers to, it seems natural to also include major historical tragedies within this category. If participating in a revolution can be a personally transformative experience, then, surely, living through a major historical tragedy can also be a personally transformative experience.

To illustrate this, I want to consider the testimony of two actual victims of historical tragedies: Consolee Nishimwe and Roger Harris.

Consolee Nishimwe, a gender activist and survivor of the Rwandan Genocide, describes her journey through the genocide in an interview for the United Nations.[34] *Nishimwe recalls her life before the genocide as follows:*

'Life in Rwanda before the genocide was beautiful. I was fortunate to have good parents and a great family. I had a *happy childhood* despite all the things that the country was enduring at that time and the discrimination against us as Tutsis.'

Living through the genocide gradually but radically changed Nishimwe's perspective on her youth and on what it means to be a teenager:

'We survived but we were crushed emotionally and psychologically, especially my mother. We didn't want to leave; we just hoped we could die too. […] It was very hard for me. I can't find words to describe how I felt. *I never thought I would be a normal teenager again.*'

After surviving the genocide, Nishimwe found a new life purpose in gender activism and the prevention of societal violence:

'I was deeply wounded. I still have nightmares. Yet, the voice of God kept telling me never to give up. Having my sisters and my mother around, someone I could speak to, helped a lot. Being able to tell my personal story to others helped me in my healing journey. People

---

[34] United Nations, Africa Renewal: 'I was tested to the limit – Rwanda genocide survivor'. Accessed on 13.10.2022 at https://www.un.org/africarenewal/web-features/i-was-tested-limit-–-rwanda-genocide-survivor. Nishimwe (2012) has also published the book *Tested to the Limit: A Genocide Survivor's Story of Pain, Resilience and Hope* (Bloomington, Indiana: Balboa Press).

are still going through tragedies around the world. […] This is why we need to speak up and *be a voice for these people.*'

In an online essay for *USA Today*, Roger Harris describes his personal journey through the Vietnam war.[35] In 1967, Harris was nineteen years old when he voluntarily joined the US Army and was deployed to Vietnam. Before his deployment, Harris had a quite instrumental view of his role as a soldier:

'I believed that it was now my turn, and if I were to die, my mom would receive a $10,000 death benefit and be able to purchase a house. I saw the war in Vietnam as a win-win situation.'

Gradually, however, the experience of war radically changed Harris' perspective on life and service:

'I *now have an appreciation for the gift of life*. Since returning home and completing college, I have devoted 42 years working in Boston schools. I see it as a tribute to my fellow Marines who paid the ultimate sacrifice.'

Overall, Harris describes the transformational impact of the war on his life as follows:

---

[35] USA Today News, 'Vietnam War: 6 personal essays describe the sting of a tragic conflict'. Accessed on 13.10.2022 at https://eu.usatoday.com/story/news/world/2017/09/11/vietnam-war-voices/105499474/.

'When I think about the Vietnam War, I am torn by personal emotions that range from anger and sadness to hope. The Vietnam War experience scarred me but also *shaped and molded my perspective on life.*'[36]

So, there does seem to be good grounds for treating historical tragedies as examples of personally transformative experiences. But what exactly is it that makes historical tragedies, such as the Rwandan Genocide or the Vietnam War, personally transformative?

Major historical tragedies strongly disrupt all aspects of their victims' lives. The tragedies affect the victims' physical and mental health, economic livelihoods, social relationships, political status, geographical location, etc. Now, given that historical tragedies affect all aspects of their victims' lives in a fundamental way, it is plausible to conclude that they change a person's core set of beliefs, preferences, desires, etc.

Consider again the cases of Nishimwe and Harris. Given the horrors Nishimwe has lived through, she might have completely different beliefs now about what it means to be human or to be treated with dignity. Similarly, Nishimwe's goals in life might have changed from wanting to construct an individual career to a desire to share her story with society at large. A similar change in perspective can be discerned in Harris' testimony. While he started his deployment to Vietnam with an instrumental and relatively detached view of life and service, he returned from the war with a genuine attachment to life and a strong desire for educational service.

If historical tragedies change a person's core set of beliefs, preferences, desires, etc., they also change a person's self. By going through a historical tragedy, a person's old self is often replaced by a new self. In the following sub-sections, I explain how — at least if we accept the claim that historical tragedies are personally transformative experiences — this change of self gives rise to a new kind of puzzle that is different from yet related to *No Regrets*.

---

[36] Emphasis mine throughout the two testimonies.

3.3. What Is Involved in Regretting a Tragedy?

Suppose you regret a transformative experience such as the Holocaust or the slave trade. What exactly do we mean by 'regretting' a historical tragedy?

One thing to note is that, often, when you regret a historical tragedy, you wish that this tragedy had not taken place because its victims would have been *better off* without living through the relevant events. They would not have been harmed.[37] With this in mind, I want to distinguish three questions that we can ask about regretting a historical tragedy.

First, we can ask: How can we regret a historical tragedy, such as the Holocaust or the slave trade, *at all*? This question has an easy answer. We can regret these tragedies because they killed many of their victims. If the tragedy had not taken place, the victims would not have been killed and hence been better off. This seems to give us reason to regret such historical tragedies.

Second, we can focus on the tragedies' survivors. How can we regret a historical tragedy *on behalf of the people who survived the tragedy*?

A similar response to the one above is still available here. As I mentioned above, historical tragedies, as personally transformative experiences, destroy and replace their victims' old selves. We might therefore argue that, while not all tragedies kill all their victims, they nonetheless destroy the selves these victims had before the tragedy. If the tragedies had not taken place, these old selves would still have existed and hence been better off. Again, this gives us reason to regret the tragedies.

While this response to the second question is plausible, it also seems incomplete. Regretting a historical tragedy because it destroyed the survivors' selves amounts to regretting the tragedy on behalf of the victims' *old* selves. But historical tragedies are transformative experiences.

---

[37] This idea is supported by the *Comparative Account of Harm* that I explained in Chapter One in my discussion of Das and Paul's (2020) new non-identity problem.

As a result, they do not only destroy their victims' old selves but also leave their victims with a new self that exists after the tragedy. This opens up a third question: How can we regret the tragedy *on behalf of the new selves that exist after the tragedy*?[38]

3.4. Regretting Tragedies on Behalf of New Selves

We might wonder why we should even consider regretting tragedies on behalf of the 'new' selves that emerge from these events. Does it not suffice to regret historical tragedies because they either killed their victims or destroyed their survivors' old selves? I will here consider two reasons why it is important to *also* be able to regret tragedies on behalf of the new selves that emerge from these events.

First, consider the following example.

New Friend: Over the last year, you have become friends with one of your work colleagues who has undergone a historical tragedy. The more you get to know your new friend, the more you learn about the horrible experiences they went through and how these experiences shaped them into the person they are now. Given everything your new friend went through, you feel regret for what happened to them.

In cases such as *New Friend*, it seems natural that you also regret the tragedy on behalf of the self that emerged from it.[39] You never knew your friend's old self—the self that was destroyed by the tragedy. Instead, it is their *new self* that you know and care about. More generally, it seems natural

---

[38] Of course, some people might suggest that regret is completely impersonal and hence not aimed at any particular individual or self. I will discuss this possibility in more detail in section 3.3. below.

[39] Notice, again, that I am not saying that you cannot or should not regret the tragedy on behalf of your friend's old self that was destroyed by the tragedy. Instead, I am saying that it is natural for you to also regret it on behalf of your friend's new self that emerged from the tragedy.

that we should also regret tragedies for the selves with whom we build up strong personal relationships.

Second, it seems natural to focus on the self that emerged from the tragedy because this self is the 'product' or 'result' of the tragedy. It is their life that was shaped by the tragedy they went through. Some of our ordinary language also seems to confirm that our regret naturally targets the self that exists after the tragedy occurred. It is natural for us to say things such as 'I regret the tragedy because it left you worse off than you could have been'. We here seem to be referring to the self who was shaped and left worse off by the tragedy—that is, the self after the tragedy. As noted above, regret tends to be comparative. When we regret an event for someone, we compare their situation or well-being *now* with how it could have been if the event had not taken place. In other words, we regret the present state of affairs in which the self who went through this event now finds themselves.

In what follows, I will focus on the question of how we can regret historical tragedies on behalf of the new selves that emerge from them. I argue that, if we focus on these new selves and historical tragedies are personally transformative experiences, it could be morally inappropriate for us to regret those tragedies. This is a highly problematic result, because it is entirely obvious that we should deeply regret horrible historical tragedies such as the Holocaust.

3.5. The New Puzzle

I noted above that historical tragedies are personally transformative events. The tragedies replace their victims' old selves with new selves. Crucially, it seems that these new selves came into existence only because of the tragedy. If the tragedy had not happened, the victims would have retained their old selves.

I also noted above that when we regret historical tragedies, we think that its victims could have been better off. This view of regret does not seem to apply to the new selves that emerge from a historical tragedy. If the tragedy had not taken place, the new selves would not have existed.

In addition, most tragedy survivors would probably claim that their current post-tragedy self is leading a life worth living.[40] As a result, it seems that we cannot claim that the tragedy harmed this new self. The tragedy did not harm the new self because it did not make it worse off than it would otherwise have been. If the tragedy had not happened, the new self would not have existed at all. If the victim's new self is not harmed by the historical tragedy, it seems there is no moral reason for us to regret the tragedy, at least not in terms of how it affected the victim's new self.

In addition, we might wonder whether our inability to regret the tragedy for the self that emerges from it also makes it incoherent for us to regret the tragedy *overall*. After all, while regret is an appropriate attitude *towards the old self* who was destroyed by the tragedy, it seems to be an inappropriate attitude *towards the new self* who was brought into existence by the tragedy. It thus seems that, while we have some reason to regret the tragedy, we also have some reason not to regret it, leaving it unclear what our attitude should be overall. As already noted above, this is a highly problematic result as we strongly believe that we should be able to regret these tragedies for what they did to their victims.[41]

There are at least two different ways of reading this new puzzle.

First, the puzzle might claim that the same person exists before and after the tragedy, and that it is only this person's 'self' that is changed by the tragedy.

If we adopt this first reading of the puzzle, it seems we cannot regret the tragedy, at least not for the self that emerged from it. The victim's new self was not made worse off by the tragedy. After all, it leads a life worth living and it exists only because of the tragedy.

---

[40] Of course, some people who survived a historical tragedy might claim that their current post-tragedy self is leading a life not worth living. In these cases, we are able to regret the tragedy for the self that emerged from it.

[41] Of course, there might be other reasons to regret a historical tragedy which are not linked to the fate of its victims. For example, we might regret that the tragedy led to the destruction of some material/cultural heritage.

Second, the puzzle might claim that the person before the tragedy is not the same person as the one existing after the tragedy. In other words, the tragedy kills the previously existing person and replaces them with a new person who also possesses a new self.

One potential motivation for the second reading stems from a psychological continuity view of personal identity. According to a classic formulation of the psychological continuity view by Derek Parfit (1984)

Psychological Continuity: X is numerically identical to Y 'if and only if X today is psychologically continuous[42] with Y, this continuity has the right kind of cause, and it has not taken a branching form' (Parfit 1984, 207).[43]

On the second reading of the new puzzle, we might use the psychological continuity view as follows. Historical tragedies fundamentally change a person's self — that is, their core beliefs, commitments, objectives, etc. Given the profound impact the tragedies have on a person's psychological make-up, it might seem plausible to claim that the person before the tragedy is not psychologically continuous with the person after the tragedy. As a result, the second reading can claim that the tragedy kills the person who exists before the tragedy and brings a new person into existence after the tragedy.

---

[42] Psychological continuity, in turn, is here understood by Parfit as overlapping chains of strong connectedness between the relevant mental states (e.g. memories and intentions) (e.g. Parfit 1984, 222).

[43] I focus here on Parfit's (1984) classic view from *Reasons and Persons*. In his later work, he explored other views of personal identity, such as the 'thinking parts' view (Parfit 2012, 26). For a classic criticism of the psychological continuity view see, for example, E.T. Olson (1997), *The Human Animal: Personal Identity Without Psychology* (New York: Oxford University Press).

If we adopt this second reading of the puzzle, there is again one way in which we can regret the tragedy and one way in which we cannot. We can regret the tragedy on behalf of the person who was killed by the tragedy. After all, this person was robbed of her potential future existence by the tragedy. However, we cannot regret the tragedy on behalf of the new person who was brought into existence by the tragedy. After all, this person is leading a life worth living and would not have existed without the tragedy and so was not made worse off by it.

In this chapter, I am going to adopt the first reading of the new puzzle for a reason discussed by Das and Paul (2020). Most historical tragedies are processes, i.e. events which take place over the course of several days, weeks, months, years, or even decades. It is therefore also plausible to assume that these tragedies transform a person' self in a progressive way. In other words, there is no clean break in the psychological life of the tragedies' victims. At any stage of the transformative process, the victim is psychologically continuous with the victim who existed at the preceding stage. As a result, the persons before and after the tragedy are the same despite possessing different selves (Das and Paul 2020, 195-202).

Notice also the following four points about this new puzzle.

First, the new puzzle differs from the one I discussed before. *No Regrets* focuses on the metaphysical non-identity of persons. In contrast, my new puzzle focuses on the non-identity of selves. By focusing on the non-identity of selves, my new puzzle also avoids the metaphysical robustness problem facing *No Regrets*.

Recall that *No Regrets* assumed that wishing away a historical tragedy means wishing away the victims of this tragedy. In response, I argued that while it is highly unlikely that the same people would have come into existence without the events leading up to the tragedy, this is not impossible. And, if it is not impossible for the same people to have come into existence, it seems morally appropriate to regret the tragedy.

The new puzzle does not rely on the same problematic metaphysical assumptions as *No Regrets*. Instead, it relies only on the assumption that major historical tragedies are personally transformative. As a result, the puzzle does not face the metaphysical robustness objection.

Second, the new puzzle expresses a quite natural attitude about the past. It is quite natural to say things like: *you should not regret the past, for it made you who you are today.* The new puzzle simply takes this natural thought to an extreme. It assumes that historical tragedies play a major role in (re-)shaping who the victims of these tragedies are, and it then infers that the victims of these tragedies would not have had the selves they actually have without these tragedies.

Third, the new puzzle avoids a problem faced by Das and Paul (2020) that I discussed in the literature review chapter. Recall that, on Das and Paul's view, you cannot rationally avoid a transformative experience in some cases because, on their view, in these cases the transformative experience does not lower your old self's expected well-being. However, as I noted in the previous chapter, this claim seems implausible. Going through a transformative experience is in many ways akin to death for your old self. And, surely, if your old self is killed by the transformative experience, its expected well-being has been lowered by the transformative experience.

My new puzzle avoids this problem as follows. I can accept that your old self is 'killed' and hence left worse off by the transformative experience. However, this is not sufficient for us to regret the tragedy. After all, as I argued above, when we regret a tragedy we also do so for the self that emerges from the tragedy, and the self that emerges from the tragedy is not left worse off by the transformative experience because it has a life worth living and would not have existed without this tragedy.

Finally, the new puzzle side-steps a key debate in the transformative experience literature. Several philosophers have argued that the decision-theoretic problem posed by Paul's concept of 'personally transformative experiences' can ultimately be overcome (e.g. Briggs 2015 189-216;

Pettigrew 2015, 768-774).[44] I do not need to take a stance on whether these responses to Paul's decision-theoretic claims succeed, because the new puzzle does not rely on those claims.

I take the new puzzle's conclusion to be highly counter-intuitive and problematic. I am *not* arguing that we should not regret past tragedies. Instead, I am arguing that, if we accept certain ideas from the transformative experience literature, it is much more difficult to defend common sense moral attitudes towards past tragedies than it might initially seem.

4. Possible Solutions to the New Puzzle

In this section, I examine and ultimately reject three possible solutions to the new puzzle. The three solutions are inspired, respectively, by the utilitarian, rights-based and 'bullet-biting' solutions to the traditional non-identity problem.

4.1. The New Puzzle and the Traditional Non-Identity Problem

The new puzzle and the traditional non-identity problem are both presented in a 'person-affecting' manner. To illustrate this, consider the following example of the traditional non-identity problem.

> *Jamie*: A couple want to have a child called Jamie. Before conceiving Jamie, the couple attend a pre-conception screening. The doctor offers the following assessment. If the couple conceive right away, Jamie will be born with a significant impairment. Jamie's life will still be worth living but it will be significantly worse than it would have been without the impairment. If the couple wait two months before conceiving, Jamie will be born without any significant medical impairment.

---

[44] For Paul's response to some of her critics, see, for instance, L.A. Paul (2015), '*Transformative Experience:* Replies to Pettigrew, Barnes and Campbell', *Philosophy and Phenomenological Research* 91(3): 794-813.

Suppose the couple decide to conceive right away. Intuitively, this decision seems morally wrong. However, it turns out to be very difficult to explain *why* this decision is morally wrong.

Since the couple conceived right away, Jamie was born and had a life worth living. If the couple had waited before conceiving, a numerically different child would have been born. Given that Jamie has a life worth living now and would not have existed two months from now, she was not made worse off by the couple's decision. If she was not made worse off by the decision, she was not harmed by it either. And, if she was not harmed by the decision, it is unclear what could make the decision morally wrong.[45]

Notice a key similarity to my puzzle discussed above. The traditional non-identity problem assumes that Jamie was not harmed by the couple's choice because she would not otherwise have existed and was therefore not made *worse off*. So, the non-identity problem seems to assume the

Person-Affecting Principle: An action is morally wrong only if it is bad for someone. (Parfit 1984, 358-363)

Similarly, the new puzzle assumes that we cannot regret a transformative historical tragedy because it did not make the historical victim's new self *worse off*. So, this puzzle seems to assume the

Person-Affecting Principle of Regret: We can regret an event for a self only if this self was made worse off by this event.

---

[45] I am here assuming that no third parties were harmed by the couple's decision to conceive right away.

There are two natural responses to the non-identity problem that both involve rejecting the *Person-Affecting Principle*. In the following subsections, I explain how these responses could also be applied to the new puzzle and the *Person-Affecting Principle of Regret*, and I argue that both responses should be rejected. Finally, I will consider and reject a third response to the non-identity problem and the new puzzle that involves accepting the *Person-Affecting Principle* and the *Person-Affecting Principle of Regret*.

4.2. Possible Solution One: Maximising Impersonal Well-Being in History

As noted above, both the traditional non-identity problem and the new puzzle are presented in a person-affecting manner. On a utilitarian view, we can solve both problems by adopting a more impersonal view of morality. Specifically, on this view, we hold that

> *Utilitarianism:* An action is morally right if and only if it maximises (expected) utility, with utility construed as total happiness or total well-being.[46]

The utilitarian view solves the traditional non-identity problem by holding that the couple's decision is morally wrong even though it does not make particular Jamie worse off. If the parents had waited two months before conceiving, Jamie—the Jamie they would have had—would have been born without a significant impairment and would have had greater well-being. So, on the

---

[46] Classic historical defences of utilitarianism include Jeremy Bentham (1970 [1789]), *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation* (Oxford: Oxford University Press) and John Stuart Mill (1987 [1863]), *Utilitarianism* (Buffalo: Prometheus Books). For a classic modern discussion of utilitarianism see, for example, Bernard Williams and J.J.C. Smart (1973), *Utilitarianism: For and Against* (Cambridge: Cambridge University Press).

utilitarian view, the couple's decision is morally wrong because it fails to maximise actual or expected utility in the world.

A proponent of utilitarianism might be able to solve my new puzzle by extending their view from actions to attitudes, as follows. On the extended view,

Utilitarianism*: An attitude is morally right if and only if it reflects a preference for utility maximisation.

On this view, the reason we ought to regret past tragedies is not because they made particular selves worse off. Instead, on this view, we ought to regret past tragedies — that is, wish that these tragedies had not occurred — because we ought to have a preference for the course of history with the highest utility, and this course of history would not contain these tragedies. For example, we ought to morally regret the Holocaust because it caused history to have significantly less utility than the best potential alternative history of the world.

4.2.1. Problems for the Utilitarian Solution

The problems for the utilitarian solution fall into two categories. First, there are three problems shared by both the utilitarian solution to the traditional non-identity problem and the utilitarian solution to the new puzzle. Second, there is an additional problem for the utilitarian solution to the new puzzle.

I will start with the three problems shared by the utilitarian solutions to the traditional non-identity problem and the new puzzle.

First, utilitarianism seems to commit us to the view that we ought to conceive as many happy children as possible. Recall that, on the utilitarian principle, we ought to make the world go best by creating as much overall utility as possible. Now, other things equal, by creating as many happy children as possible, we also create as much overall utility as possible. So, on this view, we

104

are committed to a moral requirement to conceive as many happy children as possible, other things equal.

This commitment will strike many people as counter-intuitive. While people typically believe that we have a moral reason *not* to bring *unhappy* people into existence, they also typically believe that we have no moral reason to bring happy people into existence.[47]

This first problem also extends to the utilitarian solution to the new puzzle. Recall that, on this solution, our attitude is morally right if and only if it reflects a preference for utility maximisation. On this view, our attitude towards history therefore ought to reflect this preference for utility maximisation. Presumably, utility maximisation implies that we have a preference for the course of history with the greatest number of happy people. After all, the course of history with the greatest number of happy people will be the course of history with the greatest possible utility. So, on this view, we seem committed to the conclusion that we morally ought to regret that past people did not conceive as many children as they could have.

Again, this commitment will strike many people as strongly counter-intuitive. While people typically believe that past people had moral reason *not* to bring *unhappy* people into existence, they also typically believe that past people had no moral reason to bring more happy people into existence.

Second, both solutions seem to commit us to the so-called 'repugnant conclusion' (Parfit 1984, 388). Imagine a choice between a world with a million very happy people and a world with a gazillion people who have lives just barely worth living. It is likely that the world with a gazillion people contains more overall utility than the world with a million people. On the utilitarian view, we should therefore prefer the world with a gazillion people who have lives barely worth living.

---

[47] This problem was first discussed in the population ethics literature in Jan Narveson (1973), 'Moral Problems of Population', *The Monist* 57: 62-86. It is also known as the procreation 'asymmetry' (McMahan 1981, 100).

This conclusion strikes many people as strongly counter-intuitive or even 'repugnant' (Parfit 1984, 388).

As pointed out by Kahane (2019), who considers a version of the utilitarian solution to *No Regrets*, the utilitarian solution also faces a backward-looking 'repugnant conclusion'. According to Kahane (2019), the utilitarian solution 'suggests that we should prefer alternative histories in which a far greater number of people come to exist, and, put together, add up to a greater total value, even if the lives of these additional people would have been just barely worth living' (Kahane 2019, 175). As in the case of the traditional non-identity problem, many people will find this conclusion strongly counter-intuitive or even repugnant.

Third, both solutions appear too impersonal. Recall that, on the utilitarian solution to the traditional non-identity problem, the couple's decision to conceive Jamie right away is morally wrong because it fails to maximise actual or expected utility in the world. This emphasis on actual or expected utility strikes many people as morally inappropriate. Instead of focusing on the fate of Jamie who is born with a medical impairment, the utilitarian solution seems to focus excessively on the impersonal badness of the couple's decision.

There is a similar problem for the utilitarian solution to the new puzzle. Recall that, on the utilitarian solution to the new puzzle, we ought to regret past tragedies — that is, we ought to wish that these tragedies had not occurred — because we have a preference for the course of history with the highest utility and this course would not contain these tragedies. On this view, there is no mention of the individual fates of the tragedies' victims. Again, this focus on the impersonal badness of historical tragedies will strike many people as morally inappropriate. When we regret tragedies such as the Holocaust, we seem to focus our regret not on history overall but on the tragic experiences of the Holocaust's individual victims. So, the very feature — a focus on the impersonal badness of past tragedies — that allows the utilitarian solution to solve the new puzzle fails to capture the person-affecting nature of our regret of past tragedies.

Finally, there is an additional problem for the utilitarian solution to the new puzzle. The solution seems to commit us to the view that we ought to regret almost the entire course of history. Suppose that, as a utilitarian, you believe that we ought to regret the Holocaust because the best possible alternative history without the Holocaust would have contained greater utility. Surely, on this reasoning, we should also regret tragedies that preceded the Holocaust such as World War One. For the possible history without World War One and without the Holocaust would have contained even more utility. This reasoning can be extended to all major tragedies in history, thus forcing us to regret almost the entire course of history (Smilansky 2020, 191-192).

This reasoning seems problematic because it does not fit our common attitudes towards past tragedies. When we regret the Holocaust's victims' fate, we tend to focus our concern on what could have been better for these particular victims rather than on which 'improvements' history could have contained overall. When we regret the Holocaust, we are simply not regretting the whole course of history.

4.3. Possible Solution Two: Respecting Victims' Rights

The second potential solution is inspired by the rights-based solution to the traditional non-identity problem (Parfit 1984, 364).

On one version of the rights-based solution to the traditional non-identity problem, the couple's decision is morally wrong, even though it does not make particular Jamie worse off, because it violates Jamie's right(s). Which right is violated by the couple's choice is open to debate. For example, we might think that the couple's choice to conceive Jamie now violated her right not to be conceived with a significant medical impairment.

However, this version of the rights-based view is usually seen as problematic. If the couple had waited before conceiving, Jamie would not have existed. And, if she would not have otherwise existed, then it seems that she cannot possess a right not to have been born without a significant medical impairment. After all, she could only have been born with a medical impairment.

An alternative version of the rights-based solution holds that it is wrong to create someone with a right that cannot be satisfied. On this version of the solution, the couple's choice to conceive Jamie now is wrong because it involves creating Jamie with a right (i.e. the right to a life without a medical impairment) that cannot be fulfilled. On this view, the couple's choice is morally wrong because they know that Jamie will exist with a right that cannot be fulfilled (Parfit 1984, 364-366).

The rights-based solution could perhaps be extended to solve the new puzzle as follows. We ought to regret past tragedies not because they made their particular victims/selves worse off. Instead, we ought to regret past tragedies because the people responsible for these tragedies violated their victims' rights. Again, the question of exactly which right has been violated by the perpetrators remains open. For example, we might think that the tragedies' victims had a right not to be made worse off by a historical tragedy. And, if the victims had such a right, then we seem to have a moral reason to regret their fate. After all, the victims' rights were not adequately respected.

An immediate problem with this version of the solution is that the right in question cannot be the right not to be made worse off by a historical tragedy. The whole point of the new puzzle is that the victims' new selves were not made worse off by the tragedies because the tragedies replaced the victims' old selves with new selves.

However, in response, we might claim that we should regret the tragedies because they bring into existence new selves with rights that cannot be satisfied. As in the traditional non-identity debate, it is an open question what exactly those rights are. For instance, historical tragedies often bring new selves into existence in unfavourable social and economic circumstances. Life after historical tragedies is often made harder by displacement, trauma and loss. So, maybe we should regret the tragedies because they bring into existence new selves whose right not to be born into unfavourable circumstances cannot be satisfied.

4.3.1. Problems for the Rights-Based Solution

The problems for the rights-based solution fall into two categories. First, there are two problems that apply both to the rights-based solution to the traditional non-identity problem and to the rights-based solution to the new puzzle. Second, there is a problem that applies only to the rights-based solution to the new puzzle. I will start with the first two problems.

First, the rights-based solution to the traditional non-identity problem offers counter-intuitive verdicts in other procreation cases. For example, suppose that Jamie will be born with a medical impairment no matter when the couple conceive. In this case, on the view we are considering, no matter when Jamie is conceived, she will exist with a right that cannot be fulfilled. So, on this view, it is morally inappropriate for the couple to have any child at all. After all, no matter when the couple conceive, they know that Jamie will exist with a right that cannot be fulfilled.

This verdict will again strike many people as counter-intuitive. While many people believe that we have moral reason not to conceive an impaired child if we can conceive a healthy child instead, they also believe that there is no moral prohibition on conceiving a child if the child will be born with an impairment no matter when it is conceived.

A similar problem applies to the rights-based solution to the new puzzle. Suppose you cannot avoid going through a particular transformative tragedy. In other words, no matter what you do, the tragedy will bring into existence a new self with a right that cannot be fulfilled. If the rights-based solution is correct, then it seems that you cannot avoid acting wrongly. And, on this view, we ought therefore to morally regret your wrong action.

This outcome will probably again strike many people as counter-intuitive. Many people might believe that you have moral reason not to bring a new self with a right that cannot be fulfilled into existence if you could have brought a new self with no such right into existence instead. However, most people will probably also agree that there is no general moral prohibition on bringing such a self into existence.

Second, the rights-based solution to the traditional non-identity problem does not take into account the fact that rights can be 'waived'. For example, if I offer you my watch as a gift, then I waive my property right to this watch. It seems reasonable to assume that, because Jamie cannot exist with all her rights fulfilled, she would choose to waive her right not to be born with a right that cannot be fulfilled, were she somehow given this option before being born. And, if it is reasonable to assume that Jamie would waive this right, it seems the couple cannot act wrongly on account of violating this right (Parfit 1984, 364).

The fact that rights can be waived also poses a problem to the rights-based solution to the new puzzle. It seems again reasonable to assume that, because the new selves that emerged after historical tragedies cannot exist with all their rights fulfilled, they would choose to waive their right not to be brought into existence with a right that cannot be fulfilled, were they somehow given this option before being brought into existence. So, again, if it is reasonable to assume that the new selves would waive this right, it seems that we cannot regret the tragedies on their behalf, at least not because their rights were violated.

There is a third problem that applies only to the rights-based solution to the new puzzle. The rights-based solution to the new puzzle does not apply to past tragedies, such as natural catastrophes, for which no person or group of people can be held responsible. After all, it seems incoherent to claim that a tornado or tsunami 'violated' a victim's right because the victim did not consent to living through these natural events. So, this version of the rights-based solution to the new puzzle seems to lack the scope to accommodate all of the past tragedies we might want to regret.

4.4. Possible Solution Three: Biting the Bullet

I want to consider one final solution to the new puzzle inspired by a solution to the traditional non-identity problem. The utilitarian and rights-based solutions that I discussed above both try to *solve* the traditional non-identity problem — that is, show that the parents' choice in *Jamie* is morally

wrong. According to David Boonin (2014), this strategy is misguided. On Boonin's view, we ought to bite the bullet on the non-identity problem's conclusion and accept that the parents' choice is morally acceptable. In fact, Boonin goes as far as to argue that accepting the non-identity problem's conclusion is in line with our moral intuitions (Boonin 2014, 189-192).

The main idea behind Boonin's solution is the following. On his view, none of the existing solutions to the non-identity problem can offer an entirely satisfactory solution to the problem (Boonin 2014, 189).[48] Maybe, Boonin argues, this is because we have been looking at non-identity cases in the wrong way. Rather than seeing non-identity cases as a problem, we should see them as arguments establishing their conclusion. For instance, rather than thinking that the *Jamie* case forces us to accept a problematic conclusion, we should see this case as an argument establishing the moral permissibility of the parents' decision (Boonin 2014, 190-191).

Given that such a 'bullet-biting' strategy has been offered in the traditional non-identity debate, maybe we should also consider a similar solution to the new puzzle. On this view, the new puzzle presented in this chapter should not be seen as establishing a problematic conclusion but, rather, as an argument establishing that we cannot morally regret transformative tragedies on behalf of the selves that emerge from them.

Now, prima facie, this bullet-biting solution to the new puzzle might strike many people as even less plausible than Boonin's bullet-biting solution to the traditional non-identity problem. The reason for this is that this solution to the new puzzle seems to imply that we cannot regret horrible tragedies such as the Holocaust on behalf of the selves that emerge from them. Surely, this strikes many people as too hard to swallow.

---

[48] In his book, Boonin offers a detailed exposition of most existing solutions to the non-identity problem. He then argues that none of them can satisfy a set of three plausible criteria that any convincing solution to the non-identity problem should satisfy (Boonin 2014, 19-28; 189).

Having said that, maybe there is more support for this solution than we initially think. In his solution to the traditional non-identity problem, Boonin not only claims that there is no satisfactory solution to the problem and that we should therefore simply accept its conclusion. Instead, he also offers some independent arguments which are meant to show that, in fact, the non-identity problem's conclusion is in line with our moral intuitions. I will here briefly present one of these arguments and discuss whether Boonin's strategy can be extended in such a way as to provide some support for the bullet-biting solution to the new puzzle.

Boonin proceeds as follows. He wants to present us with a case that is relevantly similar to the non-identity problem, show that we accept the conclusion of this case and that we should therefore also accept the conclusion in the non-identity problem (Boonin 2014, 195). Concretely, one of the cases Boonin asks us to consider is the following.

Lake: You are walking past a lake and you see two children drowning. One child, who is blind, is somewhat closer to you and therefore easier to save than the second, sighted child. Unfortunately, you cannot save both children and you therefore decide to save the blind child that is somewhat closer to you (Boonin 2014, 195).

According to Boonin, most of us find your decision morally acceptable on an intuitive level. In an ideal world, you would be able to save both children but, given that this possibility is excluded in this kind of case, Boonin thinks that most of us believe that you made a morally acceptable decision by saving the blind child (Boon 2014, 195).

Crucially, on Boonin's view, *Lake* is relevantly similar to non-identity cases such as *Jamie*. Similar to non-identity cases, Boonin argues, you have the choice between a world in which a disabled/blind child has a life worth living and a world in which a sighted child has a life worth living. In addition, according to Boonin, just as in non-identity cases, it also costs you a bit more effort to bring about the world in which the sighted child lives. Boonin then argues that most of

us find it intuitively acceptable not to make this extra effort in *Lake* and to bring about the world in which the blind child lives (Boonin 2014, 195-198).

If Boonin is right that most of us find it morally acceptable to save the blind child and saving the blind child is relevantly similar to the non-identity problem's conclusion, then it seems to follow, on Boonin's view, that we should also morally accept the conclusion of non-identity cases. For example, we should find it morally acceptable for the couple to conceive right away in *Jamie* and therefore to conceive an impaired child rather than an unimpaired child (Boonin 2014, 198).

Independently of how convincing Boonin's argument is for the traditional non-identity problem, it is worth exploring whether his strategy can be extended to the new puzzle about transformative tragedies. The main idea behind extending Boonin's strategy is the following. Maybe we can find a case in which someone underwent a negative transformative experience but for which we find it morally acceptable not to regret this experience for the self that emerged from it. If we can find such a case, then maybe we can show that we should also accept that we cannot regret some historical tragedies for the selves that emerge from them. One case of this kind might be the following:

> *Poverty*: Suppose you always enjoyed a decent standard of living. Suddenly, however, due to circumstances beyond your control, you fall into poverty for a few years. After this period, your financial situation improves again but you feel that you are now a different person. You feel that your new self knows better what really matters in life and attaches less importance to superficial desires and commitments.

Falling into poverty is obviously a traumatising and painful experience. Having said that, it is not unimaginable to think that some people who went through a transformative experience such as *Poverty* might nonetheless not regret this experience on behalf of their new self that emerged from

it. They might feel, for instance, that their experience made them *who they are* — that is, equipped them with a new self whose values and desires have changed for the better.

If we agree with this outcome in *Poverty*, then maybe we should also accept that we cannot regret many historical tragedies on behalf of the selves that emerge from them. As in *Poverty*, many historical tragedies force their victims' old selves through highly painful experiences and replace them with a new self that, in most cases, has a life worth living. So, maybe the idea here is that, while it is natural to regret historical tragedies on behalf of the old selves that are killed by them, it is less natural to do so for the new selves that emerged from the tragedies. Cases such as *Poverty* might show us that it is possible or even intuitive for us not to regret even highly painful transformative experiences on behalf of the selves that emerged from them and have a life worth living.

It is probably true that some victims of some historical tragedies see their experience as a part of their life that made them who they are and which they do not wish to regret on behalf of their new selves. Ultimately, however, as a general view, I find the idea of extending Boonin's solution to my puzzle unconvincing. The main problem with extending Boonin's strategy lies in the nature of many historical tragedies. Many of these tragedies, such as the Holocaust or the slave trade, had destruction, humiliation or exploitation as one of their main purposes. Given this horrible purpose, the idea that we should, in some sense, not regret or even be glad that our new self emerged from these tragedies seems too hard to swallow. Many historical tragedies are unlike *Poverty* in that it seems inappropriate to see them as some kind of 'learning experience' that we should embrace or, at least, not regret. As a result, it seems unconvincing to claim that our moral intuitions push us towards the view that we should not regret historical tragedies on behalf of the new selves that emerged from them.

4.5. A Defence of the Bullet-Biting Solution?

Suppose it is true that we should not see historical tragedies as some kind of 'learning experience'. Still, there might be a different way of understanding the bullet-biting solution that I have not yet considered. What the bullet-biting solution might show is that we have been too coarse-grained in our approach to regretting a historical tragedy. Maybe the general expression 'regretting a historical tragedy' is not nuanced enough to express the many different things we should and should not regret about a historical tragedy.

Consider again someone who lived through a historical tragedy. It seems possible that this person does not regret their particular tragedy experiences. After all, these experiences made them who they are in terms of their 'practical identity' (Chan 2019, 7). Yet, it also seems possible that this person nonetheless regrets that the tragedy, considered as a whole, occurred.

The general idea behind this solution has already been discussed in the literature around agent-regret. On this view, many of us are committed to 'affirming' our own practical identity — that is, affirming the personality, projects, etc. that make us who we are, ethically speaking. This seems to be the case even if our ethical identity is the result of past experiences that were non-ideal. And, if it is true that we are committed to our own identity in this way, then we are also committed to not regretting the non-ideal events that led to our ethical identity.[49]

Perhaps we can extend this view about agent-regret to our discussion of evaluator-regret*. On this extended view, we might also be committed to affirming the ethical identity of the tragedy's surviving victims. According to this view, then, it might be appropriate for us to not regret the

---

[49] Versions of this view have been discussed by, amongst others, Elizabeth Harman (2009), 'Transformative Experiences and Reliance on Moral Testimony', *Res Philosophica*, 92 (2): 323-339; R.J Wallace (2013), *The View From Here: On Affirmation, Attachment, and the Limits of Regret* (Oxford and New York: Oxford University Press) and Camil Golub (2019), 'Personal Value, Biographical Identity, and Retrospective Attitudes', *Australasian Journal of Philosophy* 97 (1): 72-85.

specific tragic events that led to a person's new self but, at the same time, it might be appropriate for us to regret that the tragedy, as a whole, occurred. By focusing our evaluation on *two different objects*, we can see why we might, at the same time, not regret tragedies on behalf of *individuals* whose selves were shaped by their tragic experiences and yet still regret the *tragedies as whole.*

In response, it is not clear that this defence of the bullet-biting strategy offers a *coherent* way forward. It is hard to reconcile the claims that we should regret an entire event such as the Holocaust and not regret the specific events/experiences that compose it. It is conceivable that we feel torn between affirming a victim's ethical identity and hence the Holocaust and regretting the Holocaust and thus not being committed to the victim's ethical identity. However, this should not make us overlook the apparent incoherence of holding both attitudes simultaneously. According to Setiya (2016), 'at best, I can be conflicted to the point of incoherence, preferring that the Holocaust have happened while also wishing that it had not (Setiya 2016, 6)'. So, while the fine-graining approach to regretting historical tragedies might look promising at first sight, it is not clear that it can lead to a coherent solution to my puzzle.

And there is another problem with the proposal that we are here considering. Recall that my puzzle is about regretting a historical tragedy on behalf of the new selves that emerge from it. In section 3.4, I offered two main reasons for thinking that there is something particularly counterintuitive about not being able to regret a tragedy on behalf of the new selves that emerge from it. First, I argued that it is the new selves with which we have a relationship, and it is therefore natural for us to regret the tragedy on *their* behalf. Second, I also argued that the new selves are the product or result of the tragedy and it is therefore again natural that our evaluator-regret* should focus on them.

Now suppose that there is a coherent way of regretting a historical tragedy considered as a whole while not regretting the particular experiences of the selves that emerge from this tragedy. This proposal could solve my puzzle only if the ability to regret the tragedy considered as a whole sufficiently mitigates the strong counterintuitiveness of being unable to regret the tragedy on behalf

of the new selves that emerged from the tragedy. And, for the reasons just recapped, there are good grounds to think that this strong counterintuitiveness cannot be sufficiently mitigated.

Conclusion

History contains all too many tragedies. We should wish that these tragedies never took place. But several philosophers have recently argued that there is a fundamental difficulty with this intuitive attitude. On their view, we cannot regret past tragedies without wishing away the victims of these tragedies, thus undermining our moral concern for them. I called this puzzle *No Regrets*.

In this chapter, I argued that *No Regrets* faces two main problems. First, it is not clear why, as is suggested by Kahane and Smilansky, we should focus our attitude of evaluator-regret* only on the course of history as it actually happened, or on sufficiently similar possible histories. Second, it also seems that Kahane and Smilansky's focus on history as it actually happened does not do justice to our phenomenology of regret which takes into account alternative courses of history.

However, I also argued that there is a new puzzle in the vicinity of *No Regrets* which has not been discussed before. I assumed that historical tragedies are personally transformative events — that is, events which change a victim's self to such an extent that their old self before the tragedy is not the same as their new self after the tragedy. If the victim's new self is not the same after the tragedy, we cannot claim that the tragedy made this new self worse off. Hence, it seems that we cannot regret historical tragedies, at least not in terms of how it affected the selves that emerge from these tragedies.

What are our options for responding to this the new puzzle?

First, we might try to defend one of the solutions to the new puzzle inspired by one of the extant solutions to the traditional non-identity problem, or to develop some new solution. However, as I argued above, the utilitarian, the rights-based and the 'bullet-biting' solutions face serious difficulties.

Second, we might just reject the assumption that historical tragedies are personally transformative events. If historical tragedies are not personally transformative, the victims of these tragedies have essentially the same selves before and after the tragedy. If the victims have the same selves before and after the tragedy, the puzzle goes away, because the tragedy can be said to have made the victims' selves worse off.

If we pursue this second response to the new puzzle, we might come to regard the puzzle as a reductio ad absurdum of the assumption that historical tragedies are personally transformative events. This might give us grounds to question some of the other main assumptions in the literature on transformative experience.

Chapter Three:
Only We Know: Transformative Experience and Moral Evaluation

Introduction

Some experiences radically transform us. The beliefs, desires, commitments, etc. we have before the experiences are not the same as the ones we have after living through the experiences. I will borrow a term from L.A. Paul's (2014) work and refer to these experiences as 'transformative experiences' (Paul 2014, 15-16).

In this chapter, I argue that future transformative experiences — that is, future events of which we know that they will or are highly likely to fundamentally transform us — raise a moral puzzle which has not been fully appreciated in the literature. Drawing on some work by Fiona Woollard (2021, 2022), I present an argument that seems to show that, if we have not yet lived through some transformative experiences, or it is impossible for us to live through these transformative experiences, then we cannot morally evaluate these experiences. This is puzzling because, intuitively, we *can* morally evaluate these experiences.

I proceed as follows. First, drawing on the work of Paul (2014), I explain in greater detail what 'transformative experiences' are. Second, I present an argument from Woollard that aims to show both that pregnancy is a transformative experience and that any moral assessment of abortion ought to take into account what pregnancy is like for the pregnant person. Third, I argue that, if we pursue Woollard's reasoning to its logical conclusion, we are presented with a moral puzzle concerning our ability to morally evaluate abortion without having first experienced pregnancy. Fourth, I generalise this puzzle to other transformative experiences. I then conclude by considering two ways of responding to the generalised puzzle: either we can treat the puzzle's counter-intuitive conclusion as a *reductio ad absurdum* of the idea of 'transformative experiences', or we can accept this puzzle as a serious moral issue that moral philosophers need to engage with.

1. Transformative Experience

In this section, I explain in greater detail what I mean by 'transformative experience'.[50]

Consider the following example inspired by Paul (2014).

> *Prospective Parent*: Together with your partner you are considering whether or not to have a child. On the one hand, you quite like your current life which gives you more freedom and fewer responsibilities than your imagined life as a parent. On the other hand, having a child seems like a natural step in your relationship. In addition, many of your friends who are already parents have told you that they do not regret their choice at all and would never want to go back to their childless lives.

A natural way to approach the decision in *Prospective Parent* is by imagining what it would be like to be a parent and then decide whether being a parent is better for you than remaining childless. However, on Paul's view, this standard decision-making procedure faces two problems when applied to this case.

First, Paul argues that cases like this are *epistemically* transformative (Paul 2014, 89). Cases such as *Prospective Parent* put you, as the decision-maker, into an epistemically impoverished position. On Paul's view, the reason for this is that you cannot know, or have reliable expectations about, what it is like to be a parent without having been a parent already. It is true that you can draw on general knowledge about parenthood and on your friends' testimony, but, according to Paul, this would not provide you with sufficient knowledge of what it would be like *for you* to be a

---

[50] The classic discussion of the concept of 'transformative experience' can be found in L.A. Paul (2014), *Transformative Experience* (Oxford and New York: Oxford University Press). For some critical discussion of Paul's work, see the reviews of her book published for a symposium in *Philosophy and Phenomenological Research* 91 (3): 760-813.

parent (Paul 2014, 2-3, 15-16). And, on this view, if you cannot fully know what it would be like for you to be a parent, or have reliable expectations about it, any decision about becoming a parent would be neither rational nor irrational.

Second, Paul argues that some cases are not only epistemically but also *personally* transformative (Paul 2014, 16). On this view, cases such as *Prospective Parent* are highly likely to change your core preferences and desires related to parenthood. So, even if you could know what it is like to be a parent without ever having been one, it might be the case that your preferences before becoming a parent do not match your preferences after becoming a parent (Paul 2014, 16-17). For example, it might be the case that you prefer not becoming a parent before actually becoming one, and yet you would have a strong preference to be a parent if you chose to become one. So, in other words, you cannot rely on your current preferences about parenthood to make a decision, because going through the experience of parenthood is likely to change your preferences substantially. Again, on this view, any decision about becoming a parent would be neither rational nor irrational.

According to Paul, her observations about cases such as *Prospective Parent* are likely to generalise to many other experiences. Many experiences are structurally similar to *Prospective Parent* in that the people who experience them are epistemically and/or personally transformed by them (Paul 2014, 16). Further examples of potentially epistemically and personally transformative experiences include pregnancy, losing a loved one, gaining or losing a sensory ability, and falling in love.

Below, I will use Paul's concepts of 'epistemically' and 'personally' transformative experiences to develop a new moral puzzle which has not been fully appreciated in the literature.[51]

---

[51] So far, there has been little engagement with the ethics of transformative experience in the literature. In addition to Fiona Woollard's (2021; 2022) two articles that I will discuss in this chapter, the following articles/chapters have directly or indirectly addressed ethical issues about transformative experience: Elizabeth Barnes (2015) 'Social

2. Woollard's Argument

The puzzle I want to discuss in this chapter is inspired by ideas first presented in the literature by Fiona Woollard (2021, 2022). In this section of the chapter, I explain Woollard's ideas. I then expand them into a more general puzzle in the next section.

2.1. Woollard on Pregnancy

According to Woollard (2021), pregnancy is an epistemically transformative experience — an experience through which you acquire knowledge about what it is like to be pregnant which you could not have acquired without being pregnant.[52] On this view, being pregnant gives you not only factual knowledge about pregnancy but also a unique phenomenological knowledge about *what it is like* to go through this experience (Woollard 2021, 158-159). Woollard (2021) offers four main reasons why it is impossible to acquire a grasp of pregnancy without having gone through the experience yourself and thus why it should be seen as an epistemically transformative experience.

Identities and Transformative Experiences', *Res Philosophica* 92 (3): 171-187; Amia Srinivasan (2015), 'All the Same', *Times Literary Supplement* 1-3; Dana Howard: 'Transforming Others: On the limits of 'you'll be glad I did it' reasoning', *Res Philosophica* 92 (2): 341–370); Nilanjan Das and L. A. Paul (2020): 'Transformative Choice and the Non-Identity Problem' in Sauchelli, Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (London and New York: Routledge), 187-208; Yoaav Isaacs (2020) 'The Problems of Transformative Experience', *Philosophical Studies* 177: 1065-1084; Farbod Akhlaghi (2023): 'Transformative Experience and the Right to Revelatory Autonomy', *Analysis* 20 (20): 3-12.

[52] Woollard (2021) defines a narrow and a broad sense in which experiences can be epistemically transformative. Following Paul (2014), Woollard defines an epistemically transformative experience in the 'narrow' sense as an experience which allows you to acquire knowledge about what an experience is like that you could not otherwise have acquired. In contrast, Woollard defines an epistemically transformative experience in the 'broad' sense' as an experience which allows you to acquire knowledge about what an experience is like whether or not you could have acquired this knowledge otherwise. For my purposes, we can leave Woollard's distinction aside.

First, according to Woollard (2021), being pregnant is *very different* from any other experience that we will go through.[53] She describes pregnancy as 'a complex set of interacting experiences many of which are utterly different from experiences commonly had by people who have not been pregnant' (Woollard 2021, 157). She concludes that, given the unique character of pregnancy, a person who has never been pregnant cannot possibly acquire a full grasp of what it means to be pregnant (Woollard 2021, 161).

Second, on Woollard's (2021) view, experience is always *much richer* than any form of testimony or narrative. This means, she argues, that even the most detailed and lively first-person accounts of pregnancy can never convey the full experience of being pregnant (Woollard 2021, 161).

Third, Woollard (2021) proposes that the pregnancy experience contains so many simultaneous changes to one's body, sense of self, etc. that it is impossible to convey them to someone in a finite time. And, given that it is crucial to grasp all the details in order to comprehend the *whole* experience of pregnancy, it is impossible to fully convey what it is like to be pregnant to someone (Woollard 2021, 161).

Fourth, according to Woollard (2021), lived experience is so vivid that it cannot be conveyed in a fully adequate way. In other words, no matter how lively and close to reality your narrative account of pregnancy is, a person who has never been pregnant cannot fully grasp what pregnancy is like (Woollard 2021, 161). Reason four differs from reason two in that reason two emphasizes that no narrative can convey the full *details* of the pregnancy experience while reason

---

[53] Woollard (2021) does acknowledge that severe illness is similar to the experience of pregnancy. However, on her view, experiencing severe illness is still quite different from being pregnant (Woollard 2021, 167). In addition, Woollard (2021) acknowledges that serious engagement with narrative literature about pregnancy can allow for a partial grasp of the pregnancy experience. However, again, Woollard thinks that this grasp can only ever be partial without going through the experience yourself (Woollard 2021, 156).

four emphasizes that no narrative can convey the *phenomenological vividness* or *force* of the pregnancy experience.

My sense is that reasons two and four might not be helping Woollard's (2021) argument as much as she suggests. Plausibly, almost any minimally complex experience is too rich or too vivid to be conveyed fully by testimony. However, this does not make these experiences candidates for 'transformative' experiences. So, it seems that Woollard's characterisation of pregnancy as an epistemically transformative experience really relies on reasons one and three, i.e. the claims that pregnancy is so different and multidimensional that it cannot be fully grasped, at least in a finite time, without having been pregnant.

That said, reasons two and four are likely to *magnify* reasons one and three. For example, if pregnancy is very different *and* much richer than some other experiences, it will be even harder to convey what it is like to be pregnant via testimony than it is to convey what these other experiences are like via testimony.

Woollard (2021) goes on to argue that pregnancy's epistemically transformative character has important moral implications for the debate about the (im)permissibility of abortion. Specifically, Woollard (2021) argues that, to be persuasive, any argument against abortion must engage with the experience of what it is like to be pregnant (Woollard 2021, 165). To see why this is so, she asks us to consider an abortion opponent who does not want to engage with the experience of what it is like to be pregnant. According to Woollard (2021), this opponent would have to assume that the costs that pregnancy imposes on the pregnant person do not matter morally. And, in order to warrant any such claim, Woollard argues, this opponent would need to show that a) the foetus has full moral status, b) there is no morally relevant difference between killing a foetus and killing an innocent threat[54] and c) it is always impermissible to kill innocent

---

[54] An 'innocent threat' is someone who threatens to harm you without being morally culpable for posing this threat. For example, in Robert Nozick's (1974) famous 'Ray Gun' example, you find yourself sitting in a deep well, and

threats. For, according to Woollard, the abortion opponent's claim that the costs imposed on the pregnant person do not matter morally can be sustained only if all three of these assumptions turn out to be true. However, as Woollard sees it, these three assumptions should strike us as strongly counter-intuitive and hence very unlikely to be true. So, to avoid these problematic assumptions, Woollard concludes that anyone engaging with the (im)permissibility of abortion must engage with the costs involved in being pregnant, and hence with what it is like to be pregnant (Woollard 2021, 165-166).

2.2. Clarifying Woollard's Argument

Before considering how we could build on Woollard's argument, it is worth clarifying three points about her argument.

First, it is important to emphasize what the knowledge about pregnancy is that Woollard takes to be accessible only via a direct experience of pregnancy. Clearly, pregnancy imposes certain costs on the pregnant person. These costs are both physical and mental. For example, a pregnant person can sometimes experience a unique kind of physical pain due to the presence of another living being inside them. It is precisely knowledge of these physical and mental costs that is, according to Woollard, only accessible to people who are or have been pregnant (Woollard 2021, 165-166). Woollard supports this component of her view through her descriptions of the epistemically unique and exclusive position that pregnant people find themselves in (Woollard 2021, 157-161).

---

someone then throws another person down the well at the risk of killing you. The person who is falling down on you is an 'innocent threat' because they pose a threat to your life but they cannot be held morally responsible for being a threat (Nozick 1974, 34-35). On Woollard's view, an opponent of abortion would have to show that foetuses count as such 'innocent threats' and can therefore never be killed through an abortion (Woollard 2021, 165).

Second, it is also important to clarify what knowledge about pregnancy Woollard takes to be transmissible via testimony and what knowledge about pregnancy she does not take to be transmissible via testimony. Since Woollard thinks that the physical and mental costs of being pregnant are knowable *only* via direct experience, she thinks that these costs are not transmissible via testimony (Woollard 2021, 165-166). However, Woollard does not deny that the pregnant person could transmit knowledge about the moral status of pregnancy or abortion to others via testimony. So, the moral status of pregnancy or abortion is, on Woollard's view, not something that the pregnant person has exclusive access to (Woollard 2022, 487).

Third, it is important to clarify what Woollard means by 'engaging' with the debate about the permissibility of abortion and 'engaging' with the costs involved in being pregnant. Woollard writes that we 'need to figure out how to engage in ethical reasoning when relevant knowledge is not fully accessible to us' and an opponent of abortion 'should at least engage with the question of what is involved in requiring a person to remain pregnant' (Woollard 2021, 156; 165). It is not fully clear from Woollard's own writing how she understands this 'engagement'. Nonetheless, I think we can make sense of Woollard's understanding of 'engagement' as follows.

Woollard's view of 'engagement' is rooted in a dissatisfaction with the way in which abortion is being discussed in the ethical literature. Woollard argues that the debate 'has failed to pay sufficient attention to the nature of pregnancy' and that most analogies employed in this debate 'compare pregnancy […] to interactions with minimal contact and apparently minimal effects on one's body'[55] such that 'most discussion of abortion downplays the physical burdens of pregnancy' (Woollard 2021, 164). In other words, on Woollard's view, it seems that engaging with the permissibility and costs of abortion means gaining an understanding of pregnancy that does justice to its embodied and multidimensional nature.

---

[55] One example of such an analogy might be Judith Jarvis Thomson's (1971) violinist case in which a famous violinist is 'plugged' into you (Thomson 1971, 48-49).

In addition, it has become somewhat common in recent years to talk about the need to 'take seriously' people's 'lived experiences'. I think we can see Woollard's argument as closely related to this idea. Her core insight seems to be that 'engaging with' or 'taking seriously' the lived experiences of others can mean recognising that their experiences put them in an epistemically unique and exclusive position, giving them morally relevant knowledge that cannot be transmitted to others via mere testimony.

Finally, it seems that, on Woollard's view, we should understand 'engagement' to refer not only to engagement with debate and argumentation regarding the moral permissibility of abortion but also to engagement with debate and argumentation regarding social policies around pregnancy and abortion. Woollard writes that her work has 'implications for both the activities of professional philosophers and for public discourse' and that 'public discourse about the ethics of abortion must be deeply informed by the voices of those who have been pregnant' (Woollard 2022, 489). Woollard thinks that her argument forces us to rethink our approach to social policy-making in contexts in which epistemically transformative experiences (such as pregnancy) play a central role. As a policy-maker, on Woollard's view, 'engaging with' the experience of, say, pregnancy means giving pregnant people a prominent role in the policy-making process (Woollard 2022, 489).

2.3. Building on Woollard's Argument

So far, Woollard has defended two main claims. On the one hand, she has argued that pregnancy is a transformative experience that we can only partially understand if we have not been pregnant. On the other hand, she has argued that engaging with the (im)permissibility of abortion requires us to fully understand what pregnancy is like.

Woollard's two main claims make it natural to draw the following conclusion: Given that only people who have been pregnant can fully understand what it is like to be pregnant and thus engage properly with the (im)permissibility of abortion, anyone who has not been pregnant cannot engage properly with the (im)permissibility of abortion.

Perhaps surprisingly, Woollard herself does not draw this conclusion. Instead, she argues that any debate about abortion ought to be collaborative—that is, it ought to draw on the perspectives of both people who have been pregnant and people who have not been pregnant but who are willing to fully engage with the experience of pregnancy through, for example, a close study of narrative literature on pregnancy (Woollard 2021, 166). Specifically, Woollard (2022) writes that 'first, ideally, ongoing dialogue is needed rather than a one-off transmission of knowledge; second, we require critical respect for the pregnant person's understanding of the value of their fetus, rather than simply unquestioning acceptance of what they say' (Woollard 2022, 487). So, a key part of Woollard's collaborative model is a reliance on testimony from people who have been pregnant by people who have not been pregnant.

This proposal to make use of testimony from people who have been pregnant seems puzzling. After all, Woollard has previously argued that any form of testimony is insufficient to transmit the very different and vivid experience of pregnancy to someone who has not been pregnant. Following Woollard's own reasoning, it seems that the more natural conclusion would have been that people who have not been pregnant cannot morally assess the (im)permissibility of abortion.

Woollard might reply that my puzzlement stems from ignoring the collaborative nature of moral reasoning about the (im)permissibility of abortion. On Woollard's (2021, 2022) view, a key feature of this collaborative moral reasoning is the feedback relation between the person who is or was pregnant and the philosopher who has not been pregnant and is trying to morally assess the (im)permissibility of abortion (Woollard 2022, 487-488). According to Woollard, it is fine for philosophers who have not been pregnant to engage in moral reasoning about abortion but we 'should be alive to the possibility that we have missed some key detail or underestimated the significance of some proposition' in our reasoning (Woollard 2022, 487). And it is the person who is or has been pregnant who can make us aware of these kinds of errors. So, Woollard suggests, rather than relying on pure testimony, any philosopher who engages with abortion ought to test

their reasoning against the actual experience of pregnancy with the input of people who have been through this experience (Woollard 2022, 487).

I think that, even taking these ideas into account, Woollard still draws an unnatural conclusion from her premises. Specifically, it is not sufficiently clear how the collaboration or testing of ideas that she describes allows us to avoid a problematic reliance on testimony from people who have been pregnant. On the one hand, Woollard herself writes that collaboration needs a 'kind of *testimony*' (Woollard 2022, 487) but, on the other hand, she stresses that 'the limits of even impure testimony can be seen in the appropriateness of a kind of moral deference to those who have first-hand acquaintance' (Woollard 2022, 487). So, it seems to me that Woollard has to either give up her claims about the inadequacy of testimony or accept the fact that people who have not been pregnant cannot morally assess the (im)permissibility of abortion.

2.4. An Objection and a Response

I want to consider an objection to the puzzle I raised above. On this objection, my puzzle conflates two kinds of testimony, one that Woollard can allow and one that she cannot allow. Once we are clearer on what kind of testimony Woollard can and cannot allow, the objection holds, we will see that the puzzle I have built on the basis of Woollard's argument disappears.

Transmitting knowledge of the costs involved in pregnancy via testimony can mean at least two things. First, it might mean that we can transmit knowledge about the *severity and moral importance* of these costs via testimony. Second, it could mean that we can transmit knowledge about *the exact nature and lived experience* of these costs via testimony. According to the objection, Woollard can allow the former kind of testimony even if the latter kind is impossible. And, more importantly, on this objection, the first kind of testimony is all that we need for morally evaluating transformative experiences such as pregnancy. Once we know about the moral importance and severity of the costs involved in pregnancy we can, the objection holds, start our moral evaluation even if we have not experienced pregnancy ourselves.

In response, I want to make two points.

First, if Woollard thought that knowledge about the severity and moral importance of pregnancy's costs could be transmitted via testimony, then her conclusion would seem almost trivial. For her conclusion would simply be that, in order to engage with the moral permissibility of abortion, we need to enquire about the costs involved in pregnancy. But only the very staunchest opponents of abortion would disagree with this.

Instead, a more charitable reading of Woollard's argument is that, on her view, knowledge about the moral importance and severity of the costs involved in pregnancy cannot be transmitted via testimony. While this claim might be more controversial, it is also much more substantive than that attributed to Woollard by the objection. In addition, this reading of Woollard would do justice to the fact that Woollard has spent a considerable amount of space explaining the ways in which the pregnancy experience is so different and multifaceted that we cannot simply acquire relevant moral knowledge of these experiences via testimony (Woollard 2021, 161).

Second, it is not clear that we can neatly distinguish between testimony about the severity of the costs involved in pregnancy and testimony about the lived experience of being pregnant. Specifically, it is difficult to imagine that we could get a sufficient grasp of the severity of the costs involved without also getting a good grasp of what experiencing them is like. After all, the lived experience is what shapes the relevant person's perception of the costs and their severity. In other words, if we cannot acquire knowledge about the lived experience of the costs of pregnancy via testimony, we also seem unable to acquire knowledge about the severity of these costs via testimony.

The stronger conclusion that I have drawn from Woollard's argument fits better not only with Woollard's own assumptions but also with the broader literature on transformative experience. In the literature these experiences are taken to raise decision-theoretic problems. Recall that an experience is epistemically transformative if you cannot know what this experience is like without going through the experience. Now, if you cannot know what the experience is like

without going through it, it also seems that you cannot decide in advance whether you prefer going through the experience or not. In other words, if you have not gone through a particular type of epistemically transformative experience, you are in an epistemically impoverished position in relation to this type of experience. If all this is right, it seems difficult to resist the conclusion that I have drawn about abortion. After all, if you cannot or have not been pregnant, you are in precisely this epistemically impoverished position in relation to the costs involved in pregnancy/abortion.

In what follows, I build on Woollard's argument to develop a more general puzzle about moral assessment and epistemically transformative experiences.[56] Since my puzzle will depend on the assumption that only people who have gone through a transformative experience can fully appreciate the complexity of this experience, I will refer to it as the *Only We Know* puzzle.

3. The *Only We Know* Puzzle

It seems that knowledge of what some experiences are like presupposes going through these experiences. As we have seen, Woollard (2021) argues that pregnancy is one example. But, presumably, her reasoning can be extended to other kinds of experiences. Recall that, for Woollard, what makes pregnancy an epistemically transformative experience is its unique and multi-dimensional character (Woollard 2021, 161). The same can presumably be said of experiences such as being displaced, living through a societal conflict, or living through a natural catastrophe. All of these experiences have the same unique and multidimensional character that Woollard ascribes to pregnancy. And, as Paul (2014) argues, more common experiences such as becoming a parent or losing a loved one can also be epistemically (and personally) transformative experiences (Paul 2014, 16).

Now, we also know that some of us have not or will very likely never go through some of these epistemically transformative experiences. For example, for many of us, it is highly unlikely

---

[56] Woollard (2021) hints at but does not develop such a puzzle herself (Woollard 2021, 167-168).

that we will live through a severe famine.[57] But, as Woollard has argued in the case of abortion, it seems that morally assessing an experience—answering the various moral questions that the experience raises—might sometimes require being fully informed about this experience (Woollard 2021, 165). By 'fully informed', I mean that any assessment of this experience ought to take into account the question of 'what it is like' to go through this experience. Below, I will generalise this claim beyond the case of abortion and argue that, if we want to avoid a seemingly arbitrary distinction between whose lived experience matters and whose does not, we must accept that morally assessing a wide range of experiences requires knowing what it is like to go through these experiences. If all this is right, then there are a reasonable number of transformative experiences that the majority of us are unable to morally evaluate.

3.1. The Puzzle

Intuitively, we feel like we can morally assess many experiences, including many transformative experiences. For example, we feel like we can assess whether it is morally right or wrong for a country to refuse entry to refugees or refuse aid to people suffering from severe hunger. It also seems very plausible that, in order to morally assess these experiences, we need to know what it is like to go through these experiences. However, for certain transformative experiences that we have not gone or cannot go through, it seems plausible that we cannot *really* know what they are like. It therefore seems to follow, from generalisations of the arguments considered above, and contrary to our intuitions, that we are unable to morally assess these transformative experiences.

To further illustrate this puzzle, consider another more concrete example.

Migration: A natural catastrophe caused by climate change destroys large parts of your country of residence. As a result, you decide to flee via any available route. After a perilous

---

[57] Of course, the worsening effects of climate change might make these experiences ever more common.

and traumatic journey, you finally reach the neighbouring country's border. Unfortunately for you, the neighbouring country refuses entry to refugees.

It seems probable that both Paul and Woollard would class *Migration* as an epistemically and personally transformative experience. Fleeing your country of residence under these conditions has such a profound and lasting impact on your physical and psychological make-up that it is highly likely that you could not have known what this experience would be like before going through the actual experience. It also seems plausible to assume that your core desires, commitments, and beliefs are highly likely to change as a result of your going through this experience.

Now consider someone who wants to argue that it is morally permissible to refuse entry to refugees. If this person were to refuse to take into account what is it like to be a refugee in their moral assessment of migration policy, they would have to accept that the costs imposed on the refugees in *Migration* do not matter morally. And this, in turn, would force them to accept highly implausible assumptions. They would have to accept that a) only the interests of/costs to a country's residents matter morally when making decisions about migration and that b) in migration decisions the interests of/costs to a country's residents always speak against accepting refugees. If they were to accept these assumptions, they could conclude that the costs to the refugees in *Migration* do not matter morally when assessing the neighbouring country's refusal to accept refugees.[58]

---

[58] In the literature on migration, most philosophers defend accounts which permit an intake of refugees. For arguments in defence of 'open borders' for both migrants and refugees see, for example, Chandran Kukathas (2004), 'The Case for Open Immigration' in Andrew I. Cohen and Christopher Heath Wellman (eds.), *Contemporary Debates in Applied Ethics*. (Malden, MA; Oxford: Wiley-Blackwell) pp. 376-390 and Arash Abizadeh (2008), 'Democratic Theory and Border Coercion: No Right to Unilaterally Control Your Own Borders', *Political Theory* 36 (1): 37-65. Even philosophers who oppose the idea of 'open borders' in general nonetheless recognise the need for welcoming refugees.

However, as in the case of abortion, these assumptions should strike us as highly implausible and therefore unlikely to always be true.

First, while it might be plausible that we should give greater moral weight to the interests of our co-nationals when determining migration policy, it seems highly implausible to assume that *only* the interests of our co-nationals matter morally *at all*. It seems arbitrary to use national residence (which is itself often arbitrary) as a distinguishing feature for having morally relevant interests.

Second, it also seems highly implausible to assume that the interests of a country's residents always speak against allowing entry to refugees. There seem to be many scenarios in which it would be in the residents' interest to welcome refugees. For example, on a very pragmatic level, welcoming refugees might help the residents' economic situation by supporting an ageing labour force.

So, it seems that, at least as a general matter, these two assumptions cannot be sustained. If these two assumptions cannot be sustained, the costs to refugees ought to be taken into account when morally assessing migration policy. And, if the costs to refugees ought to be taken into account when assessing migration policy, we ought to take into account what it is like to be a refugee. And, if we ought to take into account what it is like to be a refugee, and being a refugee is a transformative experience, then it seems that people who have never been a refugee cannot morally assess migration policy.

Notice that my discussion of *Migration* is likely to generalise to other transformative experiences. For the claim that we need to take into account what a transformative experience is like in order to morally assess it is likely to generalise. The main reason we need to take into account what a transformative experience is like to morally assess this experience is the need to avoid

---

See, for instance, David Miller (2005), 'Immigration: The Case for Limits' in Andrew I. Cohen and Christopher Heath Wellman (eds.), *Contemporary Debates in Applied Ethics*. (Malden, MA; Oxford: Wiley-Blackwell) pp. 193-206.

morally arbitrary distinctions between whose lived experience matters and whose lived experience does not. For example, as I argued above, in *Migration*, it would be morally arbitrary to exclude the refugees' lived experience from a moral assessment of the neighbouring country's migration policy. Similar conclusions seem plausible for other transformative experiences. For example, it seems morally arbitrary to exclude the lived experience of victims of large-scale societal conflicts such as genocides from our moral assessment of these conflicts.

So, it seems that our moral assessment of many transformative experiences ought to take into account what these experiences are like. However, we also know that some people, perhaps including us, have never and will never live through these experiences (and that third-party testimony cannot be rich enough to convey the full complexity and force of these experiences). Contrary to our intuitions, then, it seems that we are not in the right position to morally assess these transformative experiences. This is the *Only We Know* puzzle.

3.2. An Objection and a Response

Before moving on, it is worth considering an objection to my previous argument.

First, let me distinguish between phenomenological knowledge of x and mere propositional knowledge of x. Phenomenological knowledge of x here refers to the kind of knowledge that Woollard refers to in her argument about pregnancy/abortion (Woollard 2022, 487). It is knowledge about what it is like, from a first-person perspective, to experience x. In contrast, propositional knowledge refers to factual knowledge which anyone, irrespective of whether they have gone through a particular experience or not, can gain about x.

My previous argument assumes that we cannot morally assess *Migration* without possessing phenomenological knowledge of what it is like to be a refugee. However, on this objection, this assumption is excessive. Instead, the objection goes, it is sufficient to possess mere propositional knowledge about the refugees in order to morally assess *Migration*.

This objection has some force. After all, independently of what it is like to be a refugee, we seem to have propositional knowledge which tells us that it is bad to be a refugee under the conditions specified in *Migration*. On this view, there are some basic objective goods that the refugees are simply lacking. For instance, they might not have access to clean/running water or hygienic sanitation facilities. If we know all of this about the refugees' condition, why do we need to also know what it is like to be a refugee to morally assess the neighbouring country's migration policy? Our mere propositional knowledge seems sufficient to offer a negative moral assessment of the migration policy based on its negative utility.

My reply to this objection has three parts.

First, it is worth re-emphasizing the epistemic gap between the people who have been through a particular transformative experience and people who have not been through this kind of transformative experience. This gap also exists in *Migration*. While we, as people who have not been through this kind of experience, might know what preferences or burdens these refugees have, we can certainly *not* know how the refugees *experience* their preferences and burdens.

And if we do not know how the refugees experience their preferences and burdens, it is hard to correctly morally evaluate their situation. The reason for this is that, from our perspective, we cannot determine which preferences/burdens are the most important to the refugees nor how much weight each refugee attaches to individual preferences and burdens. So, the transformative nature of experiences such as *Migration* creates an epistemic gap which makes it difficult to arrive at a correctly tailored and nuanced moral evaluation. By failing to take into account the refugees' own lived experience, we risk imposing our perspective of which preferences/burdens are important and how weighty each of those preferences/burdens is.

Second, there is something morally paternalistic, or at least 'detached', about claiming that we know what is morally right and what is morally wrong for refugees without taking into account their own lived experience. It seems that ignoring this first-person experience in our moral assessment of *Migration* fails to capture something important about morality. We fail to recognise

and acknowledge that the refugees are human beings with an inner life — and not merely bundles of well-being to be added up.[59]

Third, even if the particular *Migration* example did not convince every reader, there are other examples which might be even more convincing. I want to consider one further example here.

New Disability: At the age of sixteen you suffer a terrible car accident. As a result of the accident, you lose your eyesight. Apart from becoming blind, you recover completely.

I think that *New Disability* counts as an epistemically and personally transformative experience. Losing a sensory ability radically changes how one perceives the world and it seems plausible to assume that you can know what this experience is like only by actually going through it. In addition, with a new way of apprehending the world, you might also develop new corresponding desires, preferences, etc. So, it also seems plausible to assume that losing a sensory ability might radically transform your core desires, preferences, etc.

Suppose now that someone wants to implement a new public policy which affects mainly newly blind people. In addition, this person thinks that they can morally evaluate any potential policy without taking into account the lived experience of newly blind people. This strikes me as clearly the wrong approach. Focusing simply on the propositional knowledge that we can gain about newly blind people's existence seems misguided.

Blind people and people who have become blind possess unique phenomenological insights into what it is like to live as a blind person. So, it seems that they are also best placed to understand what the typical needs of blind people are. Given that we want any moral assessment

---

[59] What I am proposing here might be one plausible way of fleshing out the so-called 'value receptables' (Chappell 2015, 322) objection to utilitarian theories of morality.

of public policies for newly blind people to take into account the real needs of newly blind people it seems that we need to take into account the unique lived experience of newly blind people for our moral assessment.

However, becoming blind is a transformative experience, which means its phenomenological character is accessible only to people who have gone through a similar experience. As a result, it seems that only people who have become blind themselves can morally assess new public policies for newly blind people. In other words, we end up with the same puzzling conclusion as in *Migration*: In many cases of transformative experiences, only people who have been through these experiences can morally assess them. In these cases, possessing mere propositional knowledge about the relevant experiences seems insufficient for our moral assessment.

3.3. Why We Might Need a Solution

Before looking for potential solutions to *Only We Know*, I want to use this sub-section to explain two main reasons why I think it is important and urgent to find a solution. The puzzle seems to have shown that some (and potentially many) people cannot morally assess certain transformative experiences. This, I argue, has potentially dangerous political consequences.

First, it seems that morally assessing a type of transformative experience can be a prerequisite for morally assessing social and political policies that legislate or regulate this kind of experience. For example, it seems plausible to assume that being able to morally assess pregnancy is a prerequisite for being able to morally assess social and political policies relating to pregnancy and abortion.

Now, morally assessing social and political policies is an important tool for participating in the democratic process. For example, if there is a referendum on a new abortion policy, it seems important that every citizen can morally assess the policy in order to cast their ballot in an informed and constructive way. However, if *Only We Know* is right, so that there are some transformative

experiences (and hence some related social and political policies) that cannot be morally assessed by everyone, this effectively undermines many people's ability to fully participate in the democratic process. These people seem unable to morally assess the relevant kind of transformative experience (and hence the related policies) and therefore cannot develop an informed opinion for participation in the democratic process. In other words, there is a risk of undemocratic votes if not everyone can morally assess the transformative experiences in question.

Notice also that we cannot brush aside this worry by pointing out that people who cannot morally assess some social and political policies can make up their mind about these policies by drawing on the advice of people who are able to morally assess them. After all, Woollard has pointed out that testimony is often insufficient for conveying the full complexity of certain transformative experiences such as pregnancy (Woollard 2021, 161). And, if this is true, it seems hard to avoid the conclusion that any form of democratic 'collaboration' between people who can morally assess some transformative experiences and people who cannot is impossible. People who have not been through the relevant kind of transformative experience simply cannot know how to use the advice they receive from people who have been through the relevant transformative experiences.

Second, there is not only a risk of excluding some people from the democratic process but also a risk of not being able to justify political decisions to governed people in certain cases.

Many people seem to think that one of the pillars of a democratic and liberal government is the government's ability and willingness to justify its political decisions to the governed people.[60]

---

[60] Many philosophers endorse this general claim under the label of 'public reason theory'. On this view, a law or policy is justified only if it is acceptable to all reasonable (idealised) citizens. For a recent overview of the debate see Paul Billingham and Anthony Taylor (2022), 'A framework for analyzing public reason theories', *European Journal of Political Theory*, *21*(4), 671-691. Public reason theories have been motivated in many different ways including, for example, by an appeal to civic friendship (see, for example, John Rawls (1993) *Political Liberalism* (New York: Columbia University Press)) or justice (see, for example, Jonathan Quong (2010) *Liberalism Without Perfection* (New York and Oxford: Oxford

On this view, the government does so by transparently and publicly stating reasons for taking certain decisions or adopting certain policies that are acceptable to the governed people. However, if *Only We Know* is right, this does not always seem possible.

Suppose the government is adopting a new policy about pregnancy or abortion and needs to justify this policy to the governed people. In other words, they try to share their reasons for adopting this policy with the governed people. If *Only We Know* is right, there is a kind of phenomenological knowledge about transformative experiences such as pregnancy which cannot be shared through simple testimony. This is problematic because many of the governed people who have not been through, say, pregnancy, cannot 'understand' the reasons for the government's policy. In other words, the reasons become 'uncommunicable', thus undermining the justification requirement on government policies.

Given these two worries about the potential political consequences of *Only We Know*, it seems even more pressing to start looking for a solution to the puzzle. What are our options? I think there are broadly two. We can a) reject transformative experiences or b) look for an acceptable solution that is compatible with them. I will shortly turn to discussing each of these options in turn.

University Press), and Quong (2013) 'On The Idea Of Public Reason' in J. Mandle and D. Reidy (eds.) *The Blackwell Companion to Rawls* (Oxford: Wiley-Blackwell), 265-280)). For some prominent criticisms of public reason theory see, for instance, Joseph Raz (1998), 'Disagreement in Politics', *The American Journal of Jurisprudence* 43, 25-52. and David Enoch (2013), 'The Disorder of Public Reason', *Ethics* 123 (1), 141-176; David Enoch (2015), 'Against Public Reason' in D. Sobel, P. Vallentyne, and S. Wall (eds.), *Oxford Studies in Political Philosophy* 1 (Oxford: Oxford University Press), 112-142.

4. The Puzzle and Standpoint Epistemology

However, before considering potential solutions to my puzzle, it is worth taking a step back and further contextualising the puzzle. In this section, I consider how my interpretation of Woollard's argument and the puzzle I built on the basis of her argument fit into relevant discussions in what is known as 'standpoint epistemology'. In addition, I consider two objections that have been levelled at standpoint epistemology to assess whether analogue objections can be brought to bear on my puzzle.

4.1. Standpoint Epistemology

The core insight of standpoint epistemology is that, epistemically speaking, people are not always equals. According to Emily Tilton and Briana Toole (forthcoming), standpoint epistemology is supported by two theses: the *situated knowledge thesis* and the *epistemic privilege thesis* (Tilton and Toole forthcoming, 1). According to the *situated knowledge thesis*, the social position we occupy 'systematically shapes and limits what we know […] what we take knowledge to be as well as specific epistemic content' (Wylie 2003, 31). Put differently, the positions we occupy in our society's various hierarchies strongly influence not only what knowledge is accessible to us but also whether we take specific insights to be knowledge in the first place.

Tilton and Toole explain that the *epistemic privilege thesis* builds on the *situated knowledge thesis* by holding that people in marginalised or powerless social positions possess some epistemic advantage(s) over other people precisely because of their marginalised social position (Tilton and Tool forthcoming, 1). Thus, according to the epistemic privilege thesis, marginalised people might be better placed to know, for instance, what is wrong with the discriminatory system of which they are victims.

Tilton and Toole further explain that some philosophers have used standpoint epistemology to defend an additional view, often referred to as 'deference epistemology'.[61] On this view, epistemically less well-placed people ought to defer to — that is, defeasibly accept — the epistemic insights of epistemically better-placed people (Tilton and Tool forthcoming, 7–8). For instance, socially powerful people ought, on this view, epistemically defer to marginalised people when it comes to knowledge about, say, discrimination.

I think that there are interesting similarities between the insights of standpoint epistemology and my puzzle, and I will discuss some of them here. Having said that, there are also some important differences that it will also help to make clear.

First, on my interpretation, Woollard would agree with standpoint epistemology's main insight — that is, humans are not always equals, at least epistemically speaking. On Woollard's view, pregnant people are epistemically privileged compared to people who have not been or cannot be pregnant. For pregnant people have a unique epistemic access to the costs involved in pregnancy and abortion (Woollard 2021, 165–166).

Crucially, though, Woollard does not argue that it is pregnant people's social position which puts them in an epistemically privileged position. Instead, on Woollard's view, it is the *kind of experience* that pregnant people go through which privileges them epistemically. It seems to be the transformative nature of the pregnancy experience which equips the people who experience it with unique epistemic insights (Woollard 2021, 158–159).

The same holds with respect to my puzzle, which I take to be a generalised version of Woollard's argument. I suggest that it is the transformative nature of certain experiences which puts the people who experience them in a uniquely privileged epistemic position to, for example, assess the costs involved in these experiences. I then add that, given these people's privileged

---

[61] It is worth noting that Tilton and Toole do not endorse 'deference epistemology' themselves (Tilton and Toole forthcoming, 10-12).

epistemic position and the fact that some people cannot acquire this knowledge that is necessary for moral evaluation of the relevant experiences, we are left with the puzzle *Only We Know*.

Second, on my interpretation, Woollard's argument also contains some elements of so-called 'deference epistemology'. As I mentioned above, Woollard writes that we need 'a kind of moral deference to those who have first-hand acquaintance' with experiences such as pregnancy (Woollard 2022, 487). According to Woollard, we ought to defer to pregnant people's understanding of the costs involved in pregnancy or abortion when morally evaluating these experiences (Woollard 2021, 165–166). Now, Woollard herself does not take this deference to be limitless. She thinks that her argument allows deference to pregnant people's understanding of the costs while still permitting that people who have not been pregnant can morally evaluate relevant experiences (Woollard 2022, 487–488). However, I have argued above that the more natural conclusion to draw is that, if Woollard's argument is correct, then we need to defer our entire moral evaluation of pregnancy and abortion to people who have lived through these experiences.

As I mentioned above in section 3.3, I take this form of necessary deference to be democratically problematic. Participating in the democratic process seems to sometimes require access to morally relevant information (such as information about the costs of pregnancy or abortion) that, on my interpretation of Woollard's argument, is not always available for people who have not gone through relevant transformative experiences.

I think this last point might reveal a broader tension between deference epistemology and some versions of political liberalism and democratic theory. Deference epistemology seems to push us to accept the views of people who are epistemically better situated than us as our own views. However, at the same time, at least if we accept some of the core principles of political liberalism, it seems important that each of us has an appropriate grasp of relevant knowledge such that we can understand public justifications based on this knowledge or make democratic choices based on this knowledge. While the intention behind deference epistemology — doing justice to

the socially situated perspectives of different social groups — is often laudable, it might also be in tension with the liberal democratic goals of political inclusion and public justification.

4.2. Objection One: Reifying Differences?

According to Tilton and Toole, one major objection to standpoint epistemology is the claim that it 'wrongfully reifies differences between social groups' (Tilton and Toole forthcoming, 2). The main worry here is that by differentiating between people on the basis of their epistemic status we also risk perpetuating the misleading view that social groups can be differentiated based on some features shared by all and only members of these groups.

As noted above, standpoint epistemology relies on the *epistemic privilege thesis*, i.e. the view that people from marginalised or powerless social groups sometimes possess epistemic advantages because of their social position. One key question, then, is what explains the fact that only people from these marginalised social groups have access to certain epistemic insights? Tilton and Toole consider two potential answers which they ultimately reject.

First, Tilton and Toole consider the view according to which people from marginalised social groups possess some essential natural or biological feature in virtue of which they possess an epistemic advantage over people who do not belong to their social group (Tilton and Toole forthcoming, 3). For example, on this view, women might possess some essential biological features which separate their experience and epistemic standpoint from those of men.

In response, Tilton and Toole argue that one of the main issues with this type of biological essentialism is the fact that it risks legitimising differences between, say, women and men which are not at all natural but rather often socially constructed (Tilton and Toole forthcoming, 3). In other words, the essentialist view risks reifying or entrenching existing socially constructed stereotypes and injustices.

Second, Tilton and Toole consider the view according to which all people from a particular marginalised social group have a similar experience of their oppression due to their shared social

location (Tilton and Toole forthcoming, 3). Unlike the biological essentialism view, this second view does not rely on the existence of some natural features shared by all and only members of a marginalised social group. Instead, this view posits the existence of a shared experience of oppression which is caused by existing social norms, conventions and injustices.

While Tilton and Toole agree with the rejection of biological essentialism, they also argue that this second view is problematic. They hold that this second view does not do justice to the heterogeneity of experiences within an oppressed social group. For example, they argue that there is no 'single' experience of oppression for women. Instead, there are many different experiences of oppression for women based not only on their gender but also on their class, race, religion, etc (Tilton and Toole forthcoming, 4).

4.2.1. Extending the Objection?

I now want to consider whether a similar objection can be brought to bear on my puzzle. We might think that my puzzle also wrongfully reifies differences between people who are or have been pregnant and people who are not or have not been pregnant. On this extended objection, my puzzle risks legitimising differences between pregnant and non-pregnant people which are not natural but rather socially constructed. In addition, my puzzle risks ignoring that pregnancy can lead to very heterogeneous experiences. I want to reply to these two points in turn.

First, it is true that Woollard's argument and the puzzle I built on it put a lot of emphasis on bodily features which make pregnancy experiences unique for pregnant people. Having said that, this assumption strikes me as much less problematic than biological essentialism's claim that there are some biological differences between, say, men and women that somehow explain their different epistemic standpoints. As Woollard has shown, pregnancy is a unique physical and mental experience that is unlike any other experience one can have (Woollard 2021, 161). It is not at all implausible that this unique physical and mental experience could ground a unique epistemic standpoint.

Second, Woollard's argument and the puzzle I built on it assume a certain homogeneity within people's pregnancy experiences. After all, my puzzle relies on the assumption that pregnancy is such a distinctive and epistemically unique experience that non-pregnant people cannot fully comprehend it. Drawing on Woollard's argument, my puzzle assumes that, despite the possible differences in pregnancy experiences, there are also significant similarities in pregnant people's experiences of their pregnancy. Again, this claim strikes me as relatively unproblematic. Consider, for instance, the wealth of written and digital resources that offer general guidance to all pregnant people throughout the different phases of their pregnancy. If there were no significant similarities in pregnant people's pregnancy experiences, these resources would not exist.

4.3. Objection Two: Deference and Laziness

Tilton and Toole consider a second objection which can be raised against deference epistemology. On this objection, the main worry is that, if the socially dominant epistemically defer too much to the socially dominated, then there is no incentive for the former to critically engage with the issues that affect the latter (Tilton and Toole forthcoming, 8–9).

Recall that, according to deference epistemology, epistemically less well-placed people ought to defer to epistemically better-placed people. For example, people from socially dominant groups ought to defer to people from socially dominated groups when it comes to insights about discrimination. The worry is that, given the socially dominant people's position in society, they can never access the epistemic position of the socially dominated. In other words, their social position definitively limits what they can know. If this were indeed the case, the objection goes, then there would be no incentive for the socially dominant to try and learn about the socially dominated people's perspective. After all, if deference epistemology is right, then there is no way for the socially dominant to even access alternative perspectives. As a result, according to the objection, the socially dominant risk becoming 'epistemically lazy' (Tilton and Tool forthcoming, 8–9). Tilton and Toole here rely on José Medina's (2013) work and define epistemic laziness as a 'persistent

refusal to probe for evidence or to consider the perspectives of others' (Tilton and Tool, forthcoming, 8).

According to Tilton and Toole, this epistemic laziness is problematic both for the socially dominant and for the socially dominated. If the socially dominant become epistemically lazy, they cannot develop a proper understanding of others' perspectives. As a result, they cannot contribute to the collective work of building a more just social order. In addition, if the socially dominant become epistemically lazy, even more epistemic burden falls on the socially dominated. The socially dominated are then solely responsible for building up the epistemic insights that can help us overcome an unjust social order (Tilton and Tool forthcoming, 9–10).

4.3.1. Extending the Objection?

Again, I will consider whether this second objection can be brought to bear on my own puzzle. On this extended objection, my puzzle risks encouraging epistemic laziness from people who cannot be pregnant. After all, based on Woollard's argument, I suggested that pregnancy, due to its transformative character, is so epistemically unique that only people who have gone through this experience can fully comprehend it. If the pregnancy experience is really inaccessible for some people, then there is a risk that these people become epistemically lazy because there is no way for them to fully engage with the pregnancy experience. This, in turn, might be problematic because, as I argued above, we need everyone to engage with experiences such as pregnancy for democratic reasons.

However, rather than being an objection to my puzzle, these ideas from the debate around standpoint epistemology help to bring out the intuitive force that drives my puzzle. I have argued that there is a clear moral and democratic need for everyone to be able to engage with moral and social debates around pregnancy. However, at the same time, I also argued that pregnancy is epistemically transformative such that not everyone can fully engage with this experience. So, just

as this extended version of the objection maintains, there is a risk that the transformative nature of pregnancy and similar experiences has ethically and politically adverse consequences.

4.4. 'Achieving' a New Standpoint

Before moving on, I want to consider one final point made by Tilton and Toole in response to the two objections to standpoint epistemology discussed above. On their view, it is wrong to claim that only marginalised people have access to the marginalised epistemic standpoint. Tilton and Toole argue that, with enough epistemic work and effort, everyone can achieve the marginalised epistemic standpoint. For example, according to Tilton and Toole, there is nothing which, in principle, prevents people from socially dominant groups from reaching the same epistemic standpoint (and hence insights) as members of marginalised groups (Tilton and Toole forthcoming, 10–12). Tilton and Toole call their view the 'achievement thesis', because the core idea is that epistemic standpoints are achieved or earned (Tilton and Toole forthcoming, 10).

According to Tilton and Toole, if you want to earn access to the marginalised epistemic standpoint, you need to commit to a process of 'consciousness-raising' (Tilton and Toole forthcoming, 10). This process involves the conscious effort of moving beyond one's current ideological frameworks and towards a less biased understanding of the social order and the groups within it.

Finally, Tilton and Toole argue that their achievement view can hold onto standpoint epistemology while avoiding the two objections above. It is still the case that socially marginalised people possess certain epistemic advantages in virtue of their social position. However, they also argue that access to the marginalised epistemic standpoint is not exclusive. As a result, there is also no need to posit any form of biological essentialism or similarity in marginalised experience that could somehow explain how marginalised people have access to a standpoint to which socially dominant people do not have access (Tilton and Tool forthcoming, 10–12).

Now, we might wonder whether Tilton and Toole's achievement view can also be extended to my puzzle. Perhaps, on this extended view, everyone can somehow achieve or earn access to the epistemic insights of pregnancy, even without having been pregnant. However, notice that this extended view would amount to denying that pregnancy is a genuine transformative experience. On this extended view, everyone can, in principle, access the insights of pregnancy, and so pregnancy is not transformative in the way suggested by Woollard and by my puzzle. I will consider this response to my puzzle below, in section 5, before turning to other potential solutions to my puzzle in section 6.

5. Rejecting 'Transformative Experiences'

First, we could take *Only We Know* to be one piece in a broader reductio ad absurdum of the concept of 'transformative experiences'. The main idea behind this strategy is that, if Paul (2014) is right that transformative experiences cause serious problems for standard decision-theory, and if transformative experiences lead to my strongly counter-intuitive moral puzzle, then this gives us reason to reject the concept of transformative experience altogether.

Recall that, according to Paul, the concept of transformative experience undermines the standard decision-theoretic model in cases such as *Prospective Parent*. Given the transformative nature of this kind of experience, we cannot know what being a parent will be like for us and, as a result, Paul argues that we cannot make a rational decision about whether we want to be a parent or not. In addition, even if we could know what it would be like for us to be a parent, Paul argues that we often do not know what our preferences will be after the transformative experience. And if we do not know whether we will prefer to be a parent or not after the experience, we again seem unable to make a rational decision about parenthood on Paul's view (Paul 2014, 8-18).

So, it seems that, on Paul's view, we have to give up on the standard decision theoretic model. But, crucially, as I have argued, transformative experiences do not seem to lead only to problems in decision theory. As I have tried to show in this chapter, accepting 'transformative

experiences' seems to generate the moral puzzle *Only We Know*. So, the challenge of 'transformative experiences' seems even more profound than initially assumed, because it forces us to give up not only on standard decision theory but also on some of our convictions about moral assessment.

This combined challenge to decision theory and morality might be too much for some people. These people might reasonably hold that this challenge should just make us sceptical about the plausibility of the concept of 'transformative' experiences. Maybe, one might argue, if the concept of transformative experiences leads to such counter-intuitive conclusions in both decision theory and morality, then it is more plausible for us to give up on transformative experiences. On this view, then, we could deny that there are genuinely 'transformative' experiences. Maybe these experiences do change us but not in the radical way described by Paul. And, if there are no genuinely 'transformative' experiences, we can both stick to standard decision theory and put aside *Only We Know*.

I think this is a reasonable response to the puzzle, but I will set it aside here so as to focus on ways forward that might be acceptable to those persuaded that transformative experiences are real.

In decision theory, several philosophers have argued that 'transformative experiences' do not generate any decision-theoretic problems that cannot be overcome by standard decision theories or, at least, adapted versions thereof. I will discuss these responses to Paul here so that we can see whether they give us any insights about how to address *Only We Know*. The responses usually proceed in at least one of four ways.

First, instead of rejecting Paul's conclusion that we cannot make rational decisions in transformative experiences, some people wish to embrace this conclusion. On their view, we accept that we cannot make rational decisions in these cases. However, this does not mean that we cannot make any choice at all or that we are forced to make a random choice. Edna Ullmann-Margalit (2006) argues that, in the absence of a rational choice, we can still make a 'reasonable' choice. On her view, a choice is reasonable only if it tries to minimise the transformative features

that lead to the decision-theoretic problem in the first place (Ullmann-Margalit 2006, 169). For instance, instead of deciding to become parents right away, Ullmann-Margalit might argue that it would be more reasonable to take this decision progressively by informing ourselves about all the aspects involved in parenthood, talking to other parents, etc. On this view, by slowing down the transformative process in this way, we can reduce the radical nature of transformative experiences and thus make more informed decisions.

Second, suppose we do not wish to embrace Paul's decision-theoretic conclusion. Recall that one of Paul's key assumptions is that we can know the value of an experience's possible outcomes only by knowing what this experience is like. For example, on Paul's view, it seems that you can know whether you enjoy being a parent or not only if you know what it is like to be a parent. According to Meena Krishnamurthy (2015), this claim is not justified. On her view, there are ways of determining the value of an experience's outcome other than knowing what this experience is like. For instance, consider again the question of whether being a parent would be a positive or negative experience for you. On Krishnamurthy's view, you can answer this question by considering whether you are, for example, someone who likes to have new experiences or whether there are any moral reasons to be or not be a parent (Krishnamurthy 2015, 174-175). In other words, on her view, you can draw on insights to determine whether being a parent would be a positive or negative experience for you other than what this experience would be like for you.

Third, suppose Paul is right that we can know the value of an experience's possible outcomes only by knowing what this experience is like. Still, some people are sceptical of Paul's further claim that you can know what a transformative experience is like only by going through this experience (and therefore you can only know the value of possible outcomes by going through these experiences). Maybe there are other ways for us to know what an experience is like. For example, Elizabeth Harman (2015) has argued that we can know what it is like to be a parent even if we have not yet been a parent ourselves. For example, Harman argues that we might experience affective relationships with our younger siblings or other people's children that are similar to those

experienced during parenthood (Harman 2015, 326). In other words, on Harman's view, we can know a lot about what it would be like to be, say, a parent even if we have not been a parent before.

Krishnamurthy (2015) also makes a similar point in her paper. On her view, we should stop thinking of an experience such as parenthood as a big, unified experience. Instead, on her view, we should think of it as consisting of many smaller experiences such as caring for a child, being angry at a child, playing with a child, etc. And once we conceive of the experience in this more fine-grained way, Krishnamurthy argues that we realise that we can have all of these experiences even if we do not have a child ourselves. For instance, we might play with our sibling's child or care for a friend's child. Crucially, on Krishnamurthy's view, this also gives us the opportunity to know whether we value these kinds of experiences or not (Krishnamurthy 2015, 176-177).

Fourth, suppose that we can know what an experience is like without going through this experience and hence what value each possible outcome has for us. This still leaves us with the problem of personally transformative experiences. In other words, the experiences are likely to change our preferences to such an extent that we cannot predict what our preferences will be after the experience. And if we cannot know what our preferences will be, we cannot base our decision on our current preferences. For example, it might be that you would rather not have children now. However, we cannot exclude the possibility that the experience of becoming a parent transforms you such that you absolutely love being a parent. As a result, it does not seem possible to take your current preference of not being a parent as a reference for your choice.

In response to this worry, Richard Pettigrew (2015, 2019) has offered a two-part reply. The first part is that, according to Pettigrew, we can know, or at least anticipate with some certainty, how our preferences are likely to change as a result of a certain transformative experience. For example, we can draw on statistics or testimonies about other people's preference changes after their transformative experiences (Pettigrew 2015, 770-771). It might be that eighty percent of

people who are relevantly similar to you love being a parent once they have become one. On Pettigrew's view, this is good evidence to suggest that you will also love being a parent.

If you are not fully convinced by this first idea, it is worth considering the second part of Pettigrew's reply. According to Pettigrew, one of Paul's assumptions is that the preferences that matter when making decisions in personally transformative cases are your current preferences (Pettigrew 2019, 149). On Paul's view, personally transformative experiences pose a problem precisely because we cannot rely on our current preferences to determine which course of action to take (Paul 2014, 16-17). After all, we are not sure whether we will have the same preferences after the transformative experience. However, on Pettigrew's view, this focus on current preferences is too narrow.

According to Pettigrew, we can divide our life into periods corresponding to different selves. For example, our self prior to a transformative experience is not the same as the self after a transformative experience. Pettigrew argues further that each of our past, present and future selves has utilities (that is, values it attaches to different outcomes) and credences (that is, levels of confidence about different ways that things could be). Pettigrew then suggests that we should weigh and aggregate these utilities and credences into a general value function. When it comes to making a decision in transformative cases, Pettigrew argues that our choice should try to maximise this general value function (Pettigrew 2019, 82-85). This has the advantage that, when we decide whether to have children, for example, we are not forced to base our decision exclusively on our current preferences. Instead, we can rely on the aggregate of all our past, present and future preferences and credences.

Given the multiplicity of solutions in the decision-theory debate, I suggest that the most constructive way forward when it comes to *Only We Know* is not to reject transformative experiences altogether but also to start looking for actual solutions to the puzzle. The solutions I discuss below will mirror those offered in the decision theory debate to some extent. Specifically, I will examine whether a) we should bite the bullet on *Only We Know*, b) we can morally assess a

transformative experience without knowing what it is like, c) we can acquire the knowledge required for moral assessment even if we do not go through the transformative experience ourselves, and d) there is a hybrid solution combining benefits from all the previous solutions.

6. Looking for a Solution

The second main option we have is to accept *Only We Know* as a genuine moral puzzle that moral philosophers need to take seriously.[62] On this view, we should try to solve the puzzle without rejecting the concept of transformative experiences.

I have already presented two solutions of this kind above. In section 2.2., I discussed and addressed Woollard's own solution according to which people who have not been through the relevant transformative experiences can nonetheless rely on the testimony of others to acquire at least part of the necessary phenomenological knowledge. In section 3.2., I discussed the solution according to which we should deny the claim that possessing phenomenological knowledge of certain transformative experiences such as pregnancy is necessary to morally assess these and related experiences. I will now set these solutions aside and discuss four additional potential solutions in turn.

6.1. Should We Trust Moral Experts?

In many areas of life we trust experts to do things on our behalf. For example, we trust a plumber to fix our shower drain if it is broken or a doctor if we suffer from a medical condition. Usually, we do not find it puzzling that we rely on these experts. There are simply people who know more or are better at certain activities and we find it normal to trust them with the activities they do on our behalf.

---

[62] Yoaav Isaacs (2020) has also emphasised the importance of philosophers paying more attention to the 'ethics' of transformative experience (Isaacs 2020, 1082).

On one possible view, we should apply the same reasoning to *Only We Know*. There are people who have gone through certain transformative experiences that we have not gone through. As a result, on this view, it seems that these people are better equipped to morally assess these experiences. Just like the doctor, these people benefit from knowledge that we do not possess. For example, we might think that people who have gone through pregnancy (and possess sufficient legal expertise) should be the ultimate decision-makers when it comes to abortion law. Given that we usually trust experts, we should also trust these experts' input when it comes to the moral assessment of certain transformative experiences.[63]

If we were to adopt this way of biting the bullet, *Only We Know* loses much of its force. On this view, there is nothing puzzling in *Only We Know*'s conclusion that there are certain experiences that cannot be morally assessed by everyone. After all, deferring to experts is common in many different areas of life and we usually do not find this practice problematic.

6.1.1. Problems with Trusting Moral Experts

In section 3.3 above, I already mentioned two political reasons why it might be problematic to leave the moral evaluations of transformative experiences entirely up to moral experts. As I argued above, this view risks excluding people from the democratic process and undermining the transparency requirement for new laws and policies. In addition to these two general political problems, I want to raise some further problems specific to the 'expertise' strategy of biting the bullet.

First, it is worth noting that we do not blindly trust experts in all areas of life. Instead, we usually draw on their advice or input to make our *own* choices. It is very rare that experts take

---

[63] Questions about moral expertise and deference to others are also relevant to discussions about transformative experiences in interpersonal contexts such as the ones discussed in Farbod Akhlaghi (2023), 'Transformative Experience and the Right to Revelatory Autonomy', *Analysis* 20 (20): 3-12.

decisions on our behalf.[64] For example, we might go to the doctor to get their advice or point of view on our injury. However, we usually do not let the doctor decide what to do about our injury. Instead, we decide what to do and whether to follow the doctor's treatment advice or not. Similarly, an analogous reasoning seems to apply in transformative experience cases. For example, as a member of parliament voting on a new abortion law, you might draw on the advice of people who have been through pregnancy. However, this does not mean that you would let these people vote on your behalf. So, if we do not usually blindly trust experts, there is no reason to blindly trust people who have gone through transformative experiences.

Second, it is unclear why we *should* trust people who possess the relevant phenomenological knowledge to make moral decisions on our behalf. While these people might have the appropriate knowledge, this does not mean that they are good at moral decision-making. Possessing phenomenological knowledge is only one part of moral assessment. In addition to this phenomenological knowledge, taking good moral decisions requires other forms of knowledge (such as propositional knowledge) and skills such as empathy, logical and critical reasoning, etc. And there is no reason to assume that people who possess the relevant phenomenological knowledge will also possess these other forms of knowledge and skills.

6.2. We All Have *Similar* Experiences

We might also argue that even people who have not been through a particular transformative experience can acquire the relevant phenomenological knowledge because these people have been through *similar* transformative experiences. The main idea is that if two transformative experiences are reasonably similar in their transformative character, then you do not need to go through both of them to acquire the relevant phenomenological knowledge.

---

[64] Rare instances where experts take decisions on our behalf might include medical procedures where we have explicitly handed over our authority to, say, a doctor.

For example, you might not have gone through a political transformation that radically changed your political values and convictions. However, maybe you have gone through a similar religious transformation from a very religious to a less religious person. In this case, it seems plausible to assume that the transformative character of both experiences is similar. Both experiences transform deep-rooted commitments that are important to your individual and social life into new and radically different commitments. And, if these transformative experiences are reasonably similar, it seems plausible to claim that you can know what it is like to go through a political transformation even if you have 'only' been through a religious transformation.

The main problem with this solution is its limited scope. It seems plausible to claim that a political and religious transformation have a lot in common. However, as pointed out by Woollard herself, it seems more difficult to find a transformative experience that is reasonably similar to pregnancy (Woollard 2021, 158). The same seems true for experiences that are not a regular part of growing older or more mature, such as *Migration*.

6.3. Can We 'Simulate' Transformative Experiences?

Suppose it is true that experiencing some transformative experiences gives you access to a unique kind of phenomenological knowledge. Is there no way of accessing, or at least sufficiently approximating, this phenomenological knowledge other than by actually going through the experience?

Woollard suggest that a way of roughly approximating the phenomenological knowledge of a transformative experience is through engagement with narrative literature about relevant experiences. However, she also stresses that even such an engagement is insufficient to acquire enough phenomenological knowledge in cases such as pregnancy (Woollard 2021, 156, 160-161).

But there might be a viable alternative that we have not yet considered. We might be able to approximate the phenomenological knowledge of certain transformative experiences by engaging in practical *simulations* of these experiences. These simulations might offer people a more

immersive understanding than reading about these experiences in narrative fiction. Crucially, practical simulations seem to be able re-create some of the details and vividness of these experiences in a much more effective way than narrative literature.

I acknowledge that the idea of 'simulating' a transformative experience such as *Migration* might seem silly or even outrageous at the beginning. Surely, one might think, there is no way in which we can simulate such complex and emotionally laden experiences. However, I think there is reason to be more optimistic about the potential of simulation. Consider the following three examples of how simulation is already being used with the purpose of recreating complex experiences.

First, a UK-based charity called *Empathy Action* offers workshops that aim to create empathy for refugees or poor people through simulation.[65] These workshops put participants in immersive environments that reconstruct the experience of, say, refugees. For instance, participants might spend time in a reconstructed refugee camp and act out the roles of refugees during the workshop.

Second, the German brand 'Produkt+Projekt' has developed an age simulation suit called *GERT Suit* (standing for Gerontological Suit) along with various accessories for the suit.[66] The purpose of the suit and its accessories is to simulate how it would feel for you to be 30 to 40 years older than you currently are. For example, the suit reduces your neck mobility through a neck brace or your visual field through a special set of glasses.

Third, new technologies such as virtual or mixed reality headsets also have the potential to simulate complex life experiences. The immersive experiences of augmented or virtual reality seem to be the closest we can currently come to actual lived experience. For instance, while surgeons

---

[65] https://www.empathyaction.org/workshops. Accessed on 10 September 2023.

[66] https://www.age-simulation-suit.com. Accessed on 10 September 2023.

might currently use these headsets to practice complex surgeries, it is not implausible to imagine that the headsets could also be used to simulate complex life experiences in the future.

When simulation tracks the lived experience of, say, refugees or elderly people to such an extent it seems plausible to assume that we can at least approximate the phenomenological knowledge of certain transformative experiences. After all, the effect of simulation workshops or age simulation suits is to make participants live through some of the psychological and physical dimensions of the relevant transformative experiences. As I noted above, new technological gear, such as augmented or virtual reality headsets, might make such simulations even more useful and accessible in the future. And, if we can roughly approximate the phenomenological knowledge of certain transformative experiences, it seems that people who are willing to engage with relevant simulations have better chances of morally assessing these experiences.

6.3.1. Problems with Simulation

Of course, this solution to *Only We Know* only goes so far.

First, it does not seem possible to offer simulations for every kind of transformative experience. For example, it seems hard to offer an effective simulation of pregnancy to people who have not been or cannot be pregnant. Pregnancy is characterised by a unique set of individual bodily features that are simply very difficult to reproduce.

Second, there might also be some transformative experiences that we can simulate but not with sufficient detail. In other words, the simulation might not be able to get close enough to the lived transformative experience. For example, it seems hard to recreate the full emotional force of a loved one's death with a simple simulation.

Having said that, these two limitations should not take away from the main point of this solution, which is that it might be possible to approximate the phenomenological knowledge of *some* transformative experiences through practical simulation. And, if we can approximate this phenomenological knowledge, it seems that people who have not lived through certain

transformative experiences can nonetheless morally assess them to some extent so long as they are

willing to engage with simulations of these experiences.[67]

6.4. A Hybrid Solution?

There seems to be no single solution to *Only We Know*. Personally, I think that the solutions

presented in sections 5 (rejecting transformative experiences) and 3.2 (denying that

phenomenological knowledge is necessary for moral assessment) are the least constructive

solutions. Instead, the most promising solution to me is to use several solutions and to create a

'hybrid' solution. Depending on the transformative experience at hand, we can draw on the

solution that is most appropriate. For example, we might think that we can simulate *some*

transformative experiences to such an extent that we can acquire the relevant phenomenological

knowledge as suggested in section 5.3. In addition to simulating some transformative experiences,

people might draw on their own transformative experiences if they are relevantly similar to *some*

other transformative experiences (as discussed in section 5.2.). Overall, the main idea behind the

hybrid solution is, in the absence of a universal solution, to be flexible and cover as many

transformative experiences as possible and thereby offer as many people as possible the

opportunity to morally assess transformative experiences.

Conclusion

Going through experiences such as pregnancy can give you unique epistemic insights into the very

different and multidimensional nature of these experiences. Amongst other things, living through

these experiences can provide you with an appreciation for the costs involved in them. On

---

[67] One interesting and perhaps unexpected upshot of this solution is that simulations which approximate personally

transformative experiences might be able to equip us with a new self in the same way that actual transformative

experiences do.

Woollard's view, this is morally significant. Woollard argues that in the case of abortion people who are or have been pregnant are in a unique epistemic position to morally assess the costs imposed on pregnant people by allowing or disallowing them an abortion. She argues further that any moral assessment of abortion ought to take into account this unique experience of people who are or have been pregnant (Woollard 2021, 163-166).

In this chapter, I developed Woollard's ideas into a general puzzle called *Only We Know*. I argued that some (or even many) of us will never go through certain epistemically and personally transformative events. I also suggested, following Woollard, that, to morally assess an experience, it is often necessary to be adequately informed about this experience. So, counterintuitively, it seems that some of us are in an epistemically impoverished position such that we are unable to morally assess some transformative experiences.

I discussed an objection to my argument according to which cases such as *Migration* can be morally assessed without appealing to the lived experience of refugees. In response, I argued that this objection might not capture a key part of moral assessment, namely the recognition and acknowledgement of people as beings with an inner life. This was made particularly clear in my discussion of *New Disability*.

In the final section of the chapter, I discussed two ways of addressing *Only We Know*, but none of these ways seems entirely convincing.

First, I considered the response which denies that there are transformative experiences. On this view, the fact that transformative experiences force us to give up on standard decision theory and lead to *Only We Know* is a reason to doubt the overall plausibility of these experiences. People inclined towards this response might argue that our experiences change us but not in the fundamental way described by Paul. As a result, on this view, there is no *Only We Know* puzzle anymore. If experiences are not radically transformative, everyone seems able to acquire the relevant phenomenological knowledge and morally assess these experiences.

However, while reasonable, I do not think this is the most constructive way forward. Instead, based on the solutions offered in the decision-theoretic debate, I suggested that we should start looking for actual solutions which take *Only We Know* seriously as a moral issue.

Second, I presented four additional solutions to the puzzle.

I started by presenting one way of biting the bullet on *Only We Know* — the possibility of deferring to moral 'experts' who possess relevant phenomenological knowledge. However, this solution has some obvious limitations as we seem unable to reliably draw on other people's moral expertise and it gives rise to political problems concerning democracy and transparency.

I then discussed two further solutions according to which people can acquire relevant phenomenological knowledge through testimony or their own similar experiences and thus morally assess transformative experiences. The main problem with these solutions is their limited scope and that testimony does not seem able to convey the full phenomenological knowledge that we need for our moral assessment.

Ultimately, I believe that a 'hybrid' solution is currently our best response to *Only We Know*. By drawing on multiple solutions, especially the simulation of transformative experiences and the use of our own similar transformative experiences, we can perhaps create a hybrid solution that has the scope to cover as many transformative experiences as possible.

Chapter Four:
I Can't Commit. Transformative Experiences and Moral Commitments

Introduction

Most of us currently have some long-term moral commitments. For example, some of us might believe that we should always support higher taxes on the wealthy, while others might believe that it is our moral duty to serve our country's army for as long as we can.

Sometimes, we can know in advance that these moral commitments will or are highly likely to change in the future. This is particularly true when we can anticipate that we will go through so-called 'transformative experiences' (Paul 2014, 15-16) — experiences which radically change our core desires, beliefs, and commitments.

In this chapter, I present a puzzle — the *I Can't Commit* puzzle — which seems to show that, if we expect to live through particular transformative experiences and we do not know *how* our beliefs will change as a result, we are rationally required to give up our corresponding long-term moral commitments. This result is puzzling because, intuitively, it is rationally permissible to retain these long-term moral commitments.

I proceed as follows.

In the first part of the chapter (sections 1 and 2), I lay the ground for my puzzle. I start by explaining L.A. Paul's (2014) concept of 'transformative experience' in more detail. I extend Paul's concept by drawing on the work of Havi Carel and Ian James Kidd (2020) and Agnes Callard (2018). I then explain an epistemological principle known as the 'Reflection Principle' (e.g. Van Fraassen 1984, 244), and I argue that a qualified and extended version of this principle applies to moral commitments. I argue that an interesting upshot of this qualified principle is that, if we expect that we will go through a certain kind of transformative experience and we know how our

moral commitments will change as a result, we are rationally required to give up some of our current moral commitments and adopt our future moral commitments instead.

In the second part of the chapter (sections 3 and 4), I present *I Can't Commit*. I argue that, if we expect that we will go through a certain kind of transformative experience and we do *not* know how our moral commitments will change as a result, rationality requires us to give up our long-term moral commitments and have only the kind of short-term moral commitments that are unlikely to be affected by future transformative experiences. I go on to defend my argument against three objections.

In the third part of the chapter (section 5), I present a way of extending *I Can't Commit* to commitments about value in general. I argue that there are certain cases involving transformative experience in which rationality requires us to give up commitments about value, such as a belief in the aesthetic value of a specific music genre.

I conclude by considering whether this puzzle should lead us to reject the notion of transformative experiences, to reject the Reflection Principle, or to try to find an alternative solution to the puzzle. Ultimately, I argue that finding a solution which takes the puzzle seriously is preferable.

1. Transformative Experiences

In this first section, I will explain the concept of 'transformative experience' in more detail.

According to Paul, we, as humans, cannot know what it is like to be some non-human creature. For example, on Paul's view, we cannot know what it is like to be an octopus (Paul 2014, 5-6). Intuitively, this will strike most readers as very plausible.

However, it is important to understand what exactly Paul means by saying that we cannot know 'what it is like to be' an octopus. On her view, this does not mean that we cannot *imagine* what it is like to be an octopus (Paul 2014, 6). Actually, many games, books and tv shows seem to

be in the business of trying to imagine what it is like to be a non-human animal. Think, for instance, of a children's tv show which includes personified animals.

Instead, Paul argues that we cannot know what it is like to be an octopus from a *first-person perspective*.[68] In other words, on Paul's view, we cannot know how the octopus experiences its own existence. The reason for this is that we, as humans, have never been and can never be an octopus (Paul 2014, 6).

Given the inaccessibility of some *non*-human experiences, Paul suggests that we should now consider whether there are also cases in which *human* experiences are not accessible to us (Paul 2014, 8). She asks us to consider the following case inspired by Frank Jackson:

> *Colour*: Mary has grown up in a black and white room and she has therefore never seen colour in her life. She now leaves the black and white room and sees a red object for the first time (Jackson 1982, 130).

On Paul's view, the case of Mary is relevant because it shows that even some human experiences are not accessible in the relevant sense. In this case, young Mary cannot know what it is like to be old Mary, at least not in terms of experiencing the colour red (Paul 2014, 9).

One of Paul's key assumptions here is that young Mary cannot know what it is like to experience the colour red because she has never experienced this colour before. On Paul's view, then, knowing what some kinds of experiences are like presupposes that we have gone through these kinds of experiences before (Paul 2014, 9).

The kinds of experiences that Paul here refers to are so-called 'epistemically transformative' experiences. Paul argues that once we have been through an experience such as seeing colour for

---

[68] For a classic statement of the idea that conscious experience consists in a first-person perspective/point of view, see Thomas Nagel (1974), 'What Is It Like to Be a Bat?', *The Philosophical Review* 83 (4): 435-450.

the first time, we are epistemically 'transformed' (Paul 2014, 15). On this view, we move from an epistemically impoverished point of view before the experience to an epistemically richer point of view after the experience. We now know 'what it is like' to, say, experience colour. Crucially, Paul argues that this transformation can occur only when we go through the experience ourselves. On her view, relying on the testimony of others is insufficient to convey the kind of knowledge that we are after. While other people's testimony can help us to better imagine what it is like to, say, experience colour, it cannot provide us with the relevant first-person perspective knowledge of colour (Paul 2014, 46-47).

Paul continues by arguing that some experiences can be not only epistemically transformative but also 'personally' transformative (Paul 2014, 48). Consider the following example.

> *Illness*: You lead a pretty regular life, and your main priorities are your professional career and leisure activities. Suddenly, you are diagnosed with a severe illness. Fortunately, after a long struggle, you survive the illness. However, you notice that being severely ill has changed you. Your values, priorities, commitments, desires, etc. in life have changed. For example, you now believe that you have a moral duty to help other people who are going through a similarly severe illness. In the same spirit, you notice that your professional career is no longer one of your main priorities.

The case described in *Illness* is a typical example of a 'personally' (as well as epistemically) transformative experience in Paul's sense. On Paul's view, we all possess a 'self' made up of our core values, commitments, desires, etc. Some experiences — those that are personally transformative — change our core values, commitments, desires, etc. to such an extent that we no longer possess the same self. After these experiences, we possess a new self (Paul 2014, 42). For example, in *Illness*, being severely ill counts as a personally transformative experience because it has

dramatically changed your core values, desires, commitments, etc. As a result, you also possess a new self which you would not have possessed had you not gone through the experience of severe illness.

Notice that, on Paul's (2020) view, transformative experiences can be processes rather than just single events (Das and Paul 2020, 205-206). For example, in the case of *Illness*, it seems less plausible to claim that you have changed as a result of a single event such as receiving the diagnosis of your illness. Instead, it seems that you have changed as a result of undergoing the lengthy process of being severely ill. The change was probably gradual and involved incremental changes to your core desires, values, etc.

Having defined epistemically and personally transformative experiences, Paul moves on to the final part of her argument. She suggests that these experiences call into question our standard view of rational decision-making. On this standard view, we decide between two alternative courses of action by assigning each possible outcome a value and a probability and then choosing the course of action which maximises expected value (Paul 2014, 20-21).

According to Paul, we cannot apply this standard model to epistemically and personally transformative experiences (Paul 2014, 30). She gives two reasons.

First, Paul argues that in order to assign two alternative courses of action a value we need to know what it would be like for us to experience the possible outcomes of these courses of action. However, as noted above, epistemically transformative experiences prevent us from knowing what it would be like for us to experience these outcomes. Consequently, we cannot know what value to attach to specific outcomes. And, if we cannot attach values to alternative outcomes, the standard decision-theoretic model cannot tell us whether choosing one outcome over the other is rational or irrational (Paul 2014, 30-31).

Second, even if we suppose that we can know what it is like to experience certain epistemically transformative experiences, Paul argues that we still face the problem of personal transformation in some cases. Recall that personally transformative experiences radically change

our core values, desires, commitments, etc. Crucially, in many of these cases, we cannot know in advance what our core values, desires, preferences, etc. will be after the personally transformative experience. As a result, even if we can attach a value to two alternative outcomes now, we cannot know which of these outcomes we will prefer *after* the personally transformative experience. So, again, the standard decision-theoretic model cannot tell us whether choosing one outcome over the other is rational or irrational (Paul 2014, 32-33).

1.2. A Multitude of Transformative Experiences

Paul's concept of transformative experiences emphasizes a particular view about how and when transformations take place. Specifically, she focuses, at least for many cases, on transformations that are the result of *major* life *decisions*. It makes sense that Paul would focus largely on cases of transformative experience involving sudden change as a result of a particular decision, because her primary interest is in drawing out the problems that transformative experiences pose for standard theories of decision making (Paul 2014, 18). However, Paul is clearly aware that some transformative experiences involve gradual rather than sudden changes, and that some do not involve any decision at all (Paul 2014, 16). We have already seen one example of this kind in *Illness*. And, as recent work by Havi Carel and Ian James Kidd (2020) and Agnes Callard (2018) makes clear, there are many examples of such gradual transformative experiences. Since the cases of transformative experience that I will focus on in this chapter involve gradual change, I will now briefly explain how Carol and Kidd, and Callard, demonstrate the ubiquity of transformative experiences involving gradual change.

1.2.1. Carel and Kidd on Transformation

Carel and Kidd's starting point is the diagnosis that human life and agency is characterised by what they refer to as three 'facts of life', namely contingency, vulnerability and subjection (Carel and Kidd 2020, 201–205). Let me briefly explain how Carel and Kidd understand each of these

concepts. In saying that human life is 'contingent', Carel and Kidd mean that humans are subject to many factors outside of their control. These factors might include our body or our social context, both of which can affect us in ways beyond our control (Carel and Kidd 2020, 202). In addition, Carel and Kidd point out that humans are 'vulnerable' to factors such as bodily ageing or illness or the breakdown of social relationships (Carel and Kidd 2020, 202–203). Finally, according to Carel and Kidd, many human lives are characterised by 'subjection', i.e. social, political or even bodily domination exercised by outsiders (Carel and Kidd 2020, 203).

Given these three facts of life, Carel and Kidd suggest that transformative experiences are varied and ubiquitous. On their view, if human life is indeed characterised by contingency, vulnerability and subjection, then we need a taxonomy to reflect these different kinds of transformative experiences. Specifically, Carel and Kidd argue that many transformative experiences are not the result of a free and rational deliberation procedure as suggested in many of Paul's examples (Carel and Kidd 2020, 205–206). In addition to the 'voluntary' (i.e. resulting from a voluntary decision) transformative experiences discussed by Paul, Carel and Kidd add the concepts of 'non-voluntary' and 'involuntary' transformative experiences (Carel and Kidd 2020, 205–206).

A non-voluntary transformative experience is an experience which transforms us in a forced way. For instance, living through a terrible illness is often a non-voluntary transformative experience for the relevant patient (Carel and Kidd 2020, 205–206). According to Carel and Kidd, a transformative experience involving a choice is 'involuntary' if the choice has unintended transformative consequences (Carel and Kidd 2020, 206). Carel and Kidd give an example in which you decide to save a child that is running in front of a car. As a result of your choice, you save the child but you also sustain life-changing injuries. In other words, on Carel and Kidd's view, while

you made the choice to save the child, your life was also transformed in an involuntary way (Carel and Kidd 2020, 206).[69]

Carel and Kidd argue that, once we see that transformative experiences need not be tied to an explicit decision-making procedure, we can make sense of how people transform over time even if they have not undergone any major transformative decision scenarios such as the ones described by Paul. On their view, many of our mundane daily experiences can transform us in a 'cumulative way'. In other words, while each of these individual experiences might not be epistemically or personally transformative, their combination over time can lead to profound transformations (Carel and Kidd 2020, 209–210).

For example, think about how our political values change over time. We might start by meeting people who hold different values from our own. We might then have interesting and stimulating conversations with these people and we might start being more aware of relevant political discussions in the news. After some time, we might even attend some political meetings that we would never have attended before. None of these individual experiences transformed our political values. Rather, on the picture presented by Carel and Kidd, it is the *accumulation* of these smaller, sometimes even 'mundane' experiences that can substantially transform our political values (Carel and Kidd 2020, 210).

In the remainder of this paper, many of the examples that I use do not resemble Paul's major life decision scenarios. Instead, my focus will be on the non-voluntary or involuntary

---

[69] Carel and Kidd also highlight the fact that transformative experiences can be positively or negatively valenced. On their view, a transformative experience is *epistemically* positive if it increases our knowledge or helps us develop good epistemic attitudes and habits (e.g. openness towards new evidence) and it is *epistemically* negative if it decreases our knowledge or causes us to develop bad epistemic attitudes (e.g. lack of openness towards new evidence); and a transformative experience is *personally* positive if it makes us into a better person (e.g. a more caring person) and *personally* negative if it makes us into a worse person (e.g. a more self-centred person) (Carel and Kidd 2020, 207).

transformative experiences which can transform us in a cumulative way as described by Carel and Kidd. I will show how a plausible epistemic principle can lead to a new moral puzzle in cases involving these transformational processes.

1.2.2. Callard on Transformation

As noted above, the cases I will discuss later in this paper will follow Carel and Kidd's model of transformation. Having said that, Agnes Callard (2018) has also identified, and extensively analysed, a model of gradual transformation over time. Callard's 'aspiration' model of transformation focuses on cases in which agents aim to change what they value. Specifically, the agents aim to take incremental steps to learn to value things they did not previously value (Callard 2018, 5–6). Discussing Callard's model here will allow me to re-emphasize how ubiquitous gradual transformative experience processes are.

Consider Callard's example of a person who aims to become a classical music enthusiast. At the present moment, this person does not value classical music for its own sake. At some point in the future, however, the person will have transformed their values such that they appreciate classical music for its own sake (Callard 2018, 68–70).

On Callard's view, a decision-based model of transformation, such as the one offered by Paul, cannot make sense of value-changing processes such as becoming a classical music enthusiast. According to Callard, aiming to transform one's values does not usually involve a single decision. Instead, aiming to become a classical music enthusiast is a process which requires a longer-term activity (Callard 2018, 54–58). For example, you might attend music classes, do research or attend operas. In addition, Callard argues that the decision-based model has trouble explaining how a value-changing process can be a rational form of behaviour (Callard 2018, 41). After all, as Paul has argued, how can you decide to value something if you do not yet know what experiencing this value will be like for you and how your preferences will evolve after your decision? Nor, on Callard's view, can we explain value-changing processes by simply appealing to

external forces which change us into a new person. After all, aiming to become a classical music enthusiast is often a form of conscious, intentional activity (Callard 2018, 57).

The challenge for Callard, then, is to explain how aiming to acquire a new value which one currently does not fully grasp can be a rational form of agency without appealing to a decision-based model nor a model on which transformations are merely the result of external influences.

In response to this challenge, Callard offers what she calls the 'aspiration' model of transformation. Aspiration starts with the ambition to change what one values. For instance, my aim now might be to be able to appreciate classical music for its own sake in the future. Notice that my ambition to appreciate classical music is not about merely fulfilling a desire that I already have. After all, it is not the case that I currently have a desire for this type of music (Callard 2018, 71–72). Instead, on Callard's view, I wish to develop this desire for classical music. So, the question, then, is what reasons could drive my new value-acquisition, given that it cannot be a current desire I already possess?

According to Callard, when we aspire to change our values, we start with an incomplete, inchoate or vague grasp of the value we wish to acquire later on (Callard 2018, 72). This incomplete grasp provides us with what Callard calls 'proleptic reasons' for action (Callard 2018, 72). Unlike other reasons for action, proleptic reasons are based on a currently incomplete value grasp and they anticipate the end state in which we have a full grasp of the value we wish to realise (Callard 2018, 72). For example, although my grasp of classical music's value is currently incomplete, this incomplete grasp is sufficient to motivate and rationalise my desire to learn to value classical music for its own sake.

In practice, Callard argues, our proleptic reasons are often reflected in the admiration or inspiration we have for mentors or other inspirational figures in the relevant value domain (Callard 2018, 80). For example, in the music case, we might be inspired by a music teacher who acts as a mentor and who has already completed the value shift we also seek to complete. In addition, Callard suggests that people who aspire to acquire new values often feel conflicted about their

value-acquisition process (Callard 2018, 117; 122-123). While they have a vague sense of, say, music's intrinsic value, they are also aware that they do not yet appreciate music for its own sake. They might therefore feel like an 'imposter' or someone who pretends to appreciate something that they, in fact, do not appreciate.

So, on Callard's view, we need to stop thinking of transformation as always involving a *single decision* in which an individual relies on a current desire to bring about some future state of affairs. Instead, according to Callard, we should also recognise cases in which transformation is a *learning process* motivated and rationalised by proleptic reasons which anticipate a future in which our grasp of a new value will be complete (Callard 2018, 237).

Like Carel and Kidd's account, Callard's aspirational model of transformation has allowed me to emphasize how common it is to undergo gradual transformational processes, whether voluntarily or involuntarily. As I mentioned above, the rest of this chapter will focus precisely on cases of gradual transformation of these kinds.

2. The 'Reflection Principle' in Epistemology

I have just begun to lay the ground for my puzzle by explaining Paul's concept of transformative experience. In this section, I continue to lay the ground, by explaining the 'Reflection Principle' in epistemology, by arguing that this principle applies to moral commitments, and by explaining some implications for cases of transformative experience.

2.1. *Reflection*

Consider the following example.

> *Driving*: Imagine that you are a teenager who currently believes that driving a car is a risky and dangerous activity. However, you also know, based on your observation of other drivers and your studies for the driving test, that it is highly likely that you will no longer

believe driving to be risky and dangerous once you have regularly driven your own car in the future.

One question about *Driving* is whether it is rational for you to stick to your current belief that driving a car is risky and dangerous even though you also believe that in the future you will not believe driving to be a risky and dangerous activity.

According to a popular view in epistemology, defended by, amongst others, C. van Fraassen (1984, 1995), you should change your current belief in cases such as *Driving*. On van Frassen's view, you should now believe what you know you will come to believe in the future, i.e. that driving is not a risky and dangerous activity (Van Fraassen 1984, 244). His view relies on a principle known as

> *Reflection*: 'Given that I now believe I will believe some proposition P at some future time, I should now believe P' (Beardman 2013, 2982).[70]

On a traditional interpretation, the main rationale behind *Reflection* is a form of 'diachronic coherence' (Foley 1994, 750) — that is, coherence between our current and future beliefs. Often, philosophers have taken van Fraassen to prescribe this coherence because he thinks the mere fact that we will have a certain future belief gives us a reason to adopt this belief now (Christensen 1996, 473).

Read in the traditional way, however, *Reflection* is not a plausible epistemic principle. The main reason for this is that, on the traditional interpretation, the principle relies on an idealised

---

[70] Van Fraassen actually refines his principle by arguing that you might now only know the range within which your future belief will fall and that you should therefore now adopt a belief which falls within this range (Van Fraassen 1995, 16).

view of epistemic agents. It does not take into account cases in which we now know that there will be a strong reason not to adopt our future belief. For example, we might now know that our future belief will be formed after a night out during which we will have had a lot to drink. In this case, contrary to *Reflection*, we should not adopt the belief that we know we will form after the night out. So, it seems that, on the traditional interpretation at least, *Reflection* suffers from problematic counter-examples.

However, there is another more plausible version of *Reflection* discussed by Ray Briggs (2009). According to Briggs (2009), the main problem with traditional *Reflection* is that it forces us to blindly trust our future self without any regard for how this future self has formed its beliefs (Briggs 2009, 68). On Briggs' (2009) view, we therefore need to qualify *Reflection* roughly as follows:

> *Qualified Reflection*: Given that I now believe I will believe some proposition P at some future time, and I am now almost certain that a) I will have updated my beliefs rationally and b) I will have updated my beliefs on the basis of veridical evidence, I should now believe P (Briggs 2009, 69-70).

On *Qualified Reflection*, we avoid the counter-example discussed above. If you will have formed your future belief after becoming drunk, you are highly likely to have violated either condition a) or condition b) of *Qualified Reflection*. For example, it might be that you have not updated your beliefs rationally due to being inebriated, or that you have not updated your beliefs on the basis of veridical evidence due to drunken hallucinations. In any case, given *Qualified Reflection*'s in-built safety conditions, we ought not to adopt our future belief now when it was formed during a night out when we had a lot to drink.

Before moving on, it is worth emphasizing just how plausible *Qualified Reflection* is as an epistemic principle. Suppose you know that you will believe P in the future and that *Qualified Reflection*'s in-built safety conditions are satisfied. In that case, the principle tells you to defer to

your better informed and reliable future self, by coming to believe P now. Following the principle thus seems, in relevant respects, like deferring to a better informed and reliable peer. And this, I think, should strike us as the rational thing to do.

If we accept *Qualified Reflection*, we should also accept some natural generalisations of this principle. Specifically, I think we should accept the following extension.

> *Extended Qualified Reflection*: Given that I now believe it is highly unlikely that I will believe some proposition P at some future time, and I am now almost certain that a) I will have updated my beliefs rationally and b) I will have updated my beliefs on the basis of veridical evidence, I should *not* believe P now.

We can support this principle by considering our judgements about cases such as the following. Suppose you consult a very reliable oracle. The oracle tells you that, in the future, you will no longer believe P, and that your no longer believing P will be the result of a rational response to veridical evidence. Should you abandon your current belief in P? It seems very intuitive to answer that you should, as implied by *Extended Qualified Reflection*.

It is worth noting that 'abandoning' your belief in P can mean either that you come to believe not-P or that you suspend your belief in P.[71] I will assume that in the cases I consider the

---

[71] In the literature, suspension of belief (also called suspension of judgement) is usually taken to be different from mere absence of belief/disbelief. Instead, most philosophers think of suspension of belief as a neutral attitude or 'committed neutrality' (Sturgeon 2010, 136) which is open to further evidence about the relevant question. For different elaborations of this 'neutral attitude' see, for instance, Scott Sturgeon (2010), 'Confidence and Coarse-Grained Attitudes' in Tamar Szabó Gendler and John Hawthorne (eds.) *Oxford Studies in Epistemology Volume 3* (Oxford: Oxford University Press), 126-149; Jane Friedman (2017), 'Why Suspend Judging?', *Noûs* 51 (2), 302-326; C.J. Atkinson (2021), 'The Aim of Belief and Suspended Belief', *Philosophical Psychology* 34 (4), 581-606. For an exception to this line

right response is to suspend your belief in P, although my arguments would also hold even if the right response were instead to believe not-P.

For the remainder of this paper, when I refer to *Qualified Reflection*, I will also have in mind the natural extension of this principle encapsulated in *Extended Qualified Reflection*.

2.2. *Qualified Reflection* and Moral Commitments

I will now argue that *Qualified Reflection* applies to moral commitments.

For the sake of the argument, I will assume that moral commitments are a form of belief. This is a controversial assumption because some moral non-cognitivists deny it.[72] I will return to this controversy later.

In addition, I will narrow the scope of my discussion of moral commitments to beliefs about *non-obvious* moral truths. A belief such as 'torturing a person for fun is morally wrong' is an obvious moral truth. It is a belief that, in a large society of reasonable individuals, most people would adopt. In contrast, the belief that we should increase taxes for the wealthy is a belief in a non-obvious moral truth. It is a belief that, in a large society of reasonable individuals, not everyone would adopt.[73] The reason I am narrowing my discussion to these kinds of commitments is that they are more likely to be affected by transformative experiences. I will discuss how these commitments are affected by transformative experiences in more detail below.

---

of reasoning see Michael R. DePaul (2004), 'Truth Consequentialism, Withholding and Proportioning Belief to the Evidence', *Philosophical Issues* 14, 91-112.

[72] For some classic discussions of non-cognitivism see, for example, R.M. Hare (1952), *The Language of Morals* (Oxford: Clarendon); Horgan T. and Timmons M. (1992), 'Troubles For New Wave Moral Semantics: The 'Open Question Argument' Revived' *Philosophical Papers* (21): 153–175 and Simon Blackburn (1993), *Essays in Quasi-Realism* (New York: Oxford University Press).

[73] For a more detailed discussion of the idea of 'reasonable individuals' and 'reasonable disagreement' see, for example, John Rawls (1993) *Political Liberalism* (New York: Columbia University Press), xvii-xviii, 48-54.

Now, if we assume that moral commitments are a form of belief, it seems natural to think that *Qualified Reflection* also extends to moral commitments. Suppose you currently believe that after your animal ethics class you will believe that it is morally wrong, or at least problematic, to eat non-human animals for pleasure. If *Qualified Reflection* is true, it seems that you should now adopt your future belief about the morally problematic nature of eating non-human animals for pleasure.

Of course, you might expect that your beliefs about eating animals will change only as a result of rhetoric or social pressure, in which case *Qualified Reflection* would not apply. But we can easily imagine the case differently, such that the conditions of *Qualified Reflection* are met. For example, you might be confident that your beliefs will be updated rationally as a result of rational persuasion by your peers and close examination of the relevant literature. In addition, your beliefs might be updated on the basis of veridical evidence given that animal ethics classes tend to avoid obviously false information about relevant topics.

In the following sub-section, I apply *Qualified Reflection* to two examples of transformative experience. I argue that an interesting upshot of applying the principle to these experiences is that, if we expect to live through particular transformative experiences and we know how our moral commitments will change as a result, we are rationally required to adopt some of our future moral commitments now.

2.3. Moral Commitments and Transformative Experiences

Consider two examples of potentially transformative experiences.

> *Political Family*: As a teenager, you have recently become more interested in politics. After attending some political meetings, you feel quite attracted to radical left-wing ideas. You even consider joining the local branch of the Communist Party. At some point, your parents notice your new enthusiasm for left-wing ideas and want to talk to you about it. Given that you consider your parents to be typical conservatives, you expect them to try

to talk you out of joining the Communist Party. Surprisingly, this is not what your parents do. Instead, they tell you a story of how they (as well as some of your grandparents) were also attracted to left-wing ideas during their teenage years. However, your parents also tell you that after they had children and developed a more settled life, their left-wing enthusiasm faded and was replaced by more moderate political ideas.

Military Supporter: As a teenager you find it difficult to be disciplined and organised and you rarely feel a sense of community towards other people. Given these difficulties, you dread the day on which you will start your compulsory military service. However, you know that it is highly likely that your attitudes and feelings will start to change over time. Many of your older acquaintances have told you that the military taught them how to be disciplined and organised. In addition, many of them have experienced a real sense of comradery with other soldiers. These people now consider themselves supporters of the military, or at least, of the values the military is meant to represent.

It is natural to assume that you rationally can (and perhaps rationally should) stick to your fundamental moral commitments, such as your commitment to progressive political ideas and policies and your desire to never join the army. However, if *Qualified Reflection* is correct, and if moral commitments are a kind of belief, then there is a problem with this intuitive view. For in these cases it seems that you should expect your future self to have different moral commitments, and the other conditions of *Qualified Reflection* appear to be satisfied. So, it seems that you should abandon your commitment to radical left-wing political ideas and your reluctance to join the army, and you should instead shift towards more centrist or conservative political ideas and become more open to joining the army.

Now, as with the animal-ethics example, we might be inclined to challenge the claim that *Qualified Reflection*'s in-built safety conditions are satisfied in these two cases. For example, it might

be that your parents' testimony in *Political Family* is unreliable. Your parents' wealth as well as their desire to secure the best possible future for you might have biased their judgement about what will happen to your political/moral commitments in the future. So, on this view, it would be irrational for you to believe that it is highly likely that you will undergo the same transformation as your parents. Similarly, we might wonder whether the strict military discipline imposed on new soldiers might have biased your informants in *Military Supporter*. If this were the case, *Qualified Reflection* would not apply as its in-built safety conditions would be violated. And, if the principle does not apply, there is no reason for you to give up your current moral commitments in these cases.

Let me make three points in response to this challenge.

First, there seems to be something potentially epistemically arrogant about assuming that your informants in these cases are biased. It seems reasonable to hold that people with more experience should be trusted more, other things being equal. More importantly, we tend to think that we should not definitively judge someone's commitments if we have not been through what they have been through. So, while this is not meant as a straightforward reply to the worry, I want to emphasize the need for epistemic modesty in our assessment of cases like *Political Family* and *Military Supporter*.

Second, it is not clear to me why we should believe that transformative experiences are necessarily epistemically problematic or impoverishing. If anything, transformative experiences seem to equip the people who live through them with more and a unique kind of evidence. Recall that an experience is considered epistemically transformative if you cannot know what the experience is like before going through it. So, people living through epistemically transformative experiences seem to acquire new evidence and knowledge that is not available to people who have not lived through the same type of experience.

Of course, there seem to be cases similar to *Political Family* in which controlling parents do bias their children's judgements. However, it also seems plausible to assume that, in many cases,

deferring to reliable and often well-intentioned elders might lead to more mature and informed commitments.

A similar point applies in *Military Supporter*. The soldier in this case is not necessarily losing evidence or becoming biased. Instead, the soldier is plausibly gaining further phenomenological knowledge about what it is like to internalise the army's values of discipline, organisation, etc. This further phenomenological knowledge then informs their decision to start supporting the army. Now, this does not mean that there cannot be cases in which joining the military biases the people going through these experiences. However, this should not make us forget the unique phenomenological insights that many of these experiences can offer.

Third, suppose the challenge accurately applies to some transformative experiences, and as such *Qualified Reflection* does not apply in these cases. Still, this gives us no reason to believe that all transformative experiences involving moral commitments lead to violations of the safety conditions specified in *Qualified Reflection*. Specifically, it seems that there are many transformative experiences that do not involve any epistemically problematic belief formation, and so there are likely to be at least some instances of transformative experiences to which *Qualified Reflection* does apply.

Before considering the general upshot I take to follow from my discussion, I want to briefly consider two natural questions about the cases I discussed in this section.

First, what exactly is it that makes cases such as *Political Family* and *Military Supporter* transformative? At first sight, there does not seem to be any major decision which upsets the lives of the relevant people. For example, in *Political Family*, there is no particular decision or event which can be easily identified as a transformative trigger.

As I see it, the cases I discuss in this chapter follow the extended model of transformative experiences developed by Carel and Kidd (2020). Recall that, on their view, transformative experiences need not necessarily be thought of as resulting from major life decisions. In fact, Carel and Kidd argue that transformative experiences do not need any decision-making context at all

(Carel and Kidd 2020, 210). And this also seems to be the case in examples such as *Political Family*. There is no particular moment at which you decide to change your political values towards a more centrist or conservative point of view. Instead, as emphasized by Carel and Kidd, your political transformation can be thought of as a 'cumulative' process in which individual and even sometimes seemingly inconsequential experiences combine into a larger transformation (Carel and Kidd 2020, 209).

In addition, cases such as *Political Family* and *Military Supporter* have elements in common with Callard's aspiration model of transformation. As noted in subsection 1.2.2, Callard does not believe that all forms of transformation are the result of single decisions. Instead, on her view, transformations are the result of aspiration or new value-acquisition processes (Callard 2018, 7). Similarly, *Political Family* and *Military Supporter* are cases involving long-term transformational processes. We could even go as far as *converting* these cases into aspiration cases like the ones discussed by Callard. That is, we could consider a version of *Political Family* in which a young person aspires to acquire the more centrist or conservative political values of their parents. This person has proleptic reasons (that is, reasons which anticipate the new value acquisition) for the actions they take towards their goal (Callard 2018, 72). On a practical level, it is also easy to imagine this person struggling with their value change, just as Callard predicted (Callard 2018, 122–123). The person might feel torn between the values they currently hold and the new values they aspire to hold in the future. In addition, as Callard suggests, this person might have a particular role model in their family who has already completed the value shift and who acts as a source of motivation (Callard 2018, 80).

Second, if cases such as *Political Family* and *Military Supporter* are really transformative, then how can we know what belief we will come to hold in the future? After all, transformative experiences are usually characterised precisely as experiences for which we do not know how our beliefs will change in the future. I want to offer a three-part reply to this question.

The first point to note is that, for many transformative experiences, we can easily predict how at least some of our beliefs will change in the future. For example, choosing to become a parent is a transformative experience, and we can easily predict that, as a result of going through this experience, you will come to believe that you have a child. The same lesson seems to hold for our moral commitments. For example, observing people's value changes allows us to predict relatively easily that many people will move from progressive to more centrist or conservative commitments over time.

The second point to remember is that what makes an experience transformative is not only determined by whether or not we know what our beliefs will be after going through the experience. Instead, what makes an experience transformative can be the phenomenological insight or evidence that the experience gives us access to and the fact that we could not have had access to this insight or evidence without going through the experience (Paul 2014, 15). This is the case regardless of whether we know what our belief will be after the experience. In other words, accepting that, in cases such as *Military Supporter* and *Political Family*, we know how (some of) our beliefs will change does not force us to deny the transformative nature of these cases. While we might know what our beliefs will be after these experiences, it is the otherwise inaccessible phenomenological evidence that these experiences provide that makes them transformative.

The third point to note is that, even if, in most real-world cases, we cannot reasonably predict what beliefs will follow from our transformative experiences, this does not undermine the work that examples such as *Political Family* are doing in this chapter. The point of these examples is a theoretical one, namely to illustrate how *Qualified Reflection* would, *in theory*, apply to transformative experiences for which we know what kind of beliefs will result. The point of these examples is not to illustrate how transformative experiences actually work in practice. Below, in section 3, I will discuss more realistic transformative scenarios in which we do not know how our beliefs will change in the future. I hope this later discussion will be clearer in light of my discussion of how *Qualified Reflection* applies to these perhaps less-realistic cases.

2.4. The Upshot

The general issue raised by cases such as *Political Family* and *Military Supporter* can be developed as follows. Many people's moral commitments extend into the future. But we know that some future events are personally transformative and also that many of us will live through (some of) these personally transformative events. After living through these personally transformative events, it is highly likely that some of our deepest moral commitments will be different than before living through these events. And if we know that it is highly likely that we will have different moral commitments after living through these events, and if *Qualified Reflection* is true, and if moral commitments are a form of belief, then, in at least some cases, it will be irrational for us to stick to our present moral commitments. Instead, we seem rationally required to adopt our future moral commitments now.

I take this to be a surprising upshot of *Qualified Reflection*, because it seems to me intuitive that we can and should rationally pursue our *current* long-term moral commitments.

I will now apply *Qualified Reflection* to a new set of transformative experiences — transformative experiences in which, though we know that our moral commitments will be changed, we do not know *how* they will change. This, I will argue, gives rise to the *I Can't Commit* puzzle.

3. The 'I Can't Commit' Puzzle

So far, I have discussed transformative experiences in which we know both that we will come to hold a particular future moral commitment which conflicts with our current moral commitments and that this future belief was formed rationally and on the basis of veridical evidence. However, there seem to be many cases involving transformative experiences in which, while we know which of our current moral commitments is likely to change, we do not know what our new future moral commitment will be.

Consider first an example which does not involve any transformative experience nor any moral commitments.

Chemistry: You are about to start your university studies in chemistry. You know that studying at university-level will significantly increase your knowledge about chemistry. In addition, you are able to project that it is highly likely that beliefs you have about particular areas of chemistry now will likely be proven wrong during your studies. These beliefs are highly likely to be replaced by new beliefs about these areas of chemistry, though you are not quite sure yet what the content of these new beliefs will be.

If *Qualified Reflection* is true, it seems that you should suspend the current beliefs about chemistry that you think are likely to change in the future. This is true even though you do not yet know what the content of your new future beliefs about chemistry will be. This is because, for any of the existing beliefs P that you think are likely to change, although you do not know what belief will replace P, you do expect that, in the future, you will not believe P. And, based on the description of *Chemistry*, we have no reason to doubt that your future beliefs about chemistry will be formed rationally and on the basis of veridical evidence.

Consider now a transformative experience which might change your moral commitments in an indefinite way. The example is a generalised version of the earlier *Political Family* case.

Politics: Many young people find themselves attracted to progressive or even strongly left-wing political ideas. However, it also seems that many of these young people undergo a political transformation as they get older. Some will defend moderate political ideas in the future while others might become stern conservatives. While it is not clear how their political views will change, it does seem clear that for many of these people their views *will* change.

There is no doubt that many people undergo a political transformation over the course of their life as described in *Politics*. What makes this process transformative is the experiences it can include, such as having children, establishing one's career, losing loved ones, etc. These experiences provide you with insights about political priorities/values which you could not have had as a young person. As I mentioned in section one, this transformation can be a lengthy process and might even take several years.

For our purposes, it is important to note that we can point relatively accurately to the kinds of current political (and hence often also moral) commitments that are likely to change as a result of this political transformation. These commitments might include beliefs in what constitutes a just distribution of resources, how much weight should be attached to law and order, etc. However, it also seems plausible to assume that, even though we know that this political transformation is highly likely to change these moral commitments, we often do not know *how* these experiences will change these commitments. In other words, we do not know which future commitments will replace our current commitments.

As in *Chemistry*, it seems that, if *Qualified Reflection* is true, we ought to suspend the current moral commitments of which we know that they are highly likely to change as a result of our life experiences. After all, we know that we are highly likely to have new moral commitments after the transformative experience and there is no reason to think that these will have been developed irrationally or on the basis of non-veridical evidence.

It is worth emphasising how my conclusion in *Politics* differs from my earlier discussion of *Military Supporter* and *Political Family*. In the latter two cases, I concluded that we are rationally required to adopt our future moral commitments now. This is not true in *Politics*. In this case, we ought to suspend our current moral commitment(s) and not replace them with any future moral commitments. The reason for this is that we know that our current belief is likely to change but

186

not *how* it will change. As a result, *Qualified Reflection* seems to instruct us only to abandon our present moral commitment but not to replace it with a future commitment.

This conclusion is further strengthened by the fact that there is always the possibility that we will experience further transformative experiences which render our future commitments even more uncertain. In other words, there is radical uncertainty about what our future moral commitments will be, and *Qualified Reflection* therefore tells us that we should suspend our current moral commitments that are highly likely to change. This is a puzzling result given that we intuitively feel that it is rational for us to stick to our current long-term moral commitments.

We can develop this puzzling conclusion into a more general puzzle. As I mentioned above, the people who undergo transformative experiences typically have moral commitments that extend into the future. However, we also know that living through (one or more) transformative experiences is highly likely to change their current long-term moral commitments. Importantly, for many of these transformative experiences, we cannot tell in advance *how* they will affect the relevant people's current moral commitments. As a result, if we follow *Qualified Reflection*, it seems that, under these conditions, these people are rationally required to suspend their current long-term moral commitments. I call this puzzle *I Can't Commit*, because it seems to force us to adopt a form of 'moral short-termism' — the view that it is only rational to pursue short-term moral commitments that are unlikely to be affected by future transformative experiences.[74]

Note that my puzzle is more limited in scope than it might initially seem. Above, I mentioned that my discussion would concern only non-obvious moral truths — that is, moral commitments which not all reasonable members of society are likely to agree on. This restriction

---

[74] Note that I am not using this term as an antonym to 'moral longtermism', a concept recently developed and defended by some effective altruists. For a detailed exposition of 'moral longtermism' see, for instance, William MacAskill (2022) *What We Owe The Future: A Million-Year View* (London: Oneworld Publications), pp.17-31; Toby Ord (2021) *The Precipice: Existential Risk and the Future of Humanity* (London: Bloomsbury Publishing).

is needed because, if a moral commitment is obvious, it is unlikely that we will abandon it in the future, even after having undergone transformative experiences. In addition, I mentioned that my puzzle is about transformative experiences in which we know which of our current moral commitments are likely to change but not how they will change. So, *I Can't Commit* applies only in cases of commitments to non-obvious moral truths and transformative experiences in which we do not know what our future belief will be.

I have used this section to show that the interesting upshot of *Qualified Reflection* identified in the previous section of this chapter can be extended into a puzzle about the interaction between our moral commitments and a certain type of transformative experience. If *Qualified Reflection* is true, it seems that we are forced to adopt a form of moral short-termism, at least if we expect to undergo transformative experiences of the kind that I have been discussing.

4. Three Objections

In this section, I address three potential objections to *I Can't Commit*.

4.1. Retaining Moral Commitments until their Replacement?

I want to consider an objection which presents an alternative to my use of *Qualified Reflection*. Recall that, on this principle, if we now know or are relatively certain that I will believe P in the future and the two safety conditions are met, then I should believe P now. Recall also my argument that, if we accept *Qualified Reflection*, it is natural for us to also accept *Extended Qualified Reflection*. This principle holds that, given that I now believe it is highly unlikely that I will believe P in the future and the two safety conditions are met, I should not believe P now.

On the objection that I am here considering, *Extended Qualified Reflection* offers the wrong verdict. Suppose I believe it is highly likely that I will not believe P in the future, and I also do not know what future belief will replace P. On this objection, it could be rational to hold onto P or, at least, act as if P were true until I know what future belief will replace P. In other words, unlike

Extended Qualified Reflection, this objection suggests that we should hold onto P and not suspend our current belief in P until we know what future belief replaces P. On this objection, the same point applies to moral commitments. We might know that our moral commitments will change as a result of future transformative experiences, but, at least until we know what our new future moral commitments will be, we should hold onto our current moral commitments.

In response, I want to return to the oracle case that I used in subsection 2.1 to support *Extended Qualified Reflection*. Recall that, in the oracle case, you consult a highly reliable oracle which tells you that you will not believe P in the future. I suggested that a natural and rational response to the oracle's prediction is to stop believing P now. If this is true, then it also seems natural and rational not to hold onto P or to act as if P were true, even if you do not yet know what future belief will replace P. Suppose that, after you consult the oracle, someone offers you a bet with odds 1/3 on whether P is true or not. Should you accept or decline this bet? It seems to me that you should decline. After all, you know that you will rationally believe not-P in the future. So, it seems natural for you to stop believing in P now, even though you do not yet know what future belief will replace P.

My response also applies in the case of moral commitments and transformative experiences. We often know that our moral commitments will be replaced through future transformative experiences. And sometimes we know that going through these transformative experiences will equip us with new phenomenological evidence and we therefore have no reason to think that these experiences lead to irrational beliefs. So, even if we do not know what our future moral commitments will be, it is natural for us to suspend our current commitments now.

4.2. Can We Simply Avoid Transformative Experiences?

We might wonder whether there is a way to stick to our current moral commitments without facing the problem of transformative experiences. For instance, is it not possible to rationally stick to our current moral commitments by trying to avoid future transformative experiences? By avoiding

these experiences, we make it more likely that our core commitments will stay the same over time. If we are serious in our commitments, we might even be rationally or morally required to avoid experiences that are likely to result in us giving them up.

In reply, it is worth considering the following point. It does not always seem possible to avoid transformative experiences in the way suggested by this objection. Consider again the *Military Supporter* case discussed above. In this particular case, given the mandatory nature of being drafted into the army, it seems difficult to avoid the transformative experience of becoming a soldier. And, even if it were possible to avoid the transformative character of becoming a soldier in this particular case, there seem to be other cases where it is even clearer that you cannot try to avoid transformative experiences. This is particularly true for natural catastrophes, societal events such as violent conflicts or displacements, the deaths of loved ones, falling seriously ill, or falling in love. So, some people, and possibly many, will be affected by transformative experiences and thus faced with the puzzle discussed above.

4.3. Non-Cognitivism's Advantage

I now want to return to an assumption that I made earlier. At the beginning of section 2.2., I assumed that moral commitments are a form of belief. I also pointed out that this assumption is controversial because some non-cognitivists might deny it.

Non-cognitivism is a family of metaethical views that share some common convictions about how we should think of moral beliefs. Usually, we assume that a belief is 'descriptive' — that is, a belief is about something in the world. According to Mark Schroeder (2010), we can think of descriptive belief as having a 'mind to world direction of fit', i.e. a belief tries to map the world as it is (Schroeder 2010, 11). Many non-cognitivists accept this description of regular descriptive belief. However, they argue that we should not confuse these regular descriptive 'beliefs' with moral 'beliefs'. On this view, moral 'beliefs' are not the same kind of mental state as regular

descriptive beliefs.[75] On a non-cognitivist view, moral beliefs have a world to mind direction of fit, i.e. they are mental states that try to make the world match themselves (Schroeder 2010, 11).

If non-cognitivism is right and moral beliefs are not the same kind of mental state as regular descriptive beliefs, we might wonder what kind of mental state these moral 'beliefs' in fact are. Different strands within non-cognitivism offer different answers to this question. Very roughly speaking, different non-cognitivist theories might argue that moral beliefs/commitments primarily express some form of non-cognitive mental state such as emotion or approval or prescription, etc.

If these non-cognitivists are right that moral commitments are not a form of regular descriptive belief, *I Can't Commit* does not seem to arise. After all, the puzzle relies on *Qualified Reflection*, which is a principle about descriptive beliefs. And if moral commitments are not a form of descriptive belief, the principle does not apply to them. If the principle does not apply to them, it remains silent on whether we need to give up our long-term moral commitments.

My response to the non-cognitivist has two parts, one critical and one sympathetic.

The critical point is to note that non-cognitivism itself remains a highly controversial thesis in metaethics. Many philosophers working on metaethics reject it. Indeed, many of these philosophers subscribe to alternative metaethical theories on which moral commitments are a form of descriptive belief. And, if these rival views assume that moral commitments are a form of descriptive belief, their proponents seem to face *I Can't Commit*.

The more sympathetic point is that the non-cognitivist might be able to use *I Can't Commit* as an argument to their advantage in two ways.

---

[75] One famous motivation for denying that moral beliefs are of the same kind as descriptive beliefs is the so-called 'motivation problem'. On this problem, it seems hard to explain how moral beliefs could be regular descriptive beliefs while also motivating people to act in ways in which regular descriptive beliefs do not. For a formulation of this problem see, for example, Russ Shafer-Landau (2003), *Moral Realism: A Defence* (Oxford: Clarendon Press), 121.

First, *I Can't Commit* is counterintuitive. If non-cognitivism avoids this puzzle, while cognitivist views give rise to it, that gives non-cognitivism an intuitive advantage over its rivals.

Second, *I Can't Commit* can contribute to the so-called 'license for optimism' arguments for non-cognitivism. Usually, these arguments do not try to solve problems facing non-cognitivism directly. Instead, the goal is to point to some phenomena that fit well with the non-cognitivist view of morality and claim that the existence of these phenomena should make us optimistic about non-cognitivism's ability to eventually solve its problems (Schroeder 2010, 189). Now, *I Can't Commit* does not give us evidence that non-cognitivism can eventually solve its problems. However, I suggest that we might be able to point to *I Can't Commit* as further reason to be optimistic of non-cognitivism's overall prospects vis-à-vis its cognitive rivals.

This novel double way of supporting non-cognitivism has not been explored in the literature, but it seems to me worthy of attention.

5. Extending *I Can't Commit*

There is something particularly counter-intuitive about *I Can't Commit*'s implication that some people ought to suspend certain of their long-term *moral* commitments. After all, moral commitments are commitments that are central to people's lives, how they think about themselves and their relations with others. So, suspending and not replacing them seems particularly troublesome.

But *I Can't Commit* might also have broader implications, beyond moral commitments. To see this, consider the following example.

Music: You are currently twelve years old and enjoy listening to the music that most people in your age cohort enjoy. However, from observing your older friends and family, you are quite confident that you will no longer be committed to this kind of music when you are twenty-five.

There are two possible interpretations of this case.

First, we can interpret *Music* in the same spirit as *Politics*. On this interpretation, we can think of your enjoyment of a particular type of music as a belief in the (aesthetic, political, etc.) value of this type of music. Furthermore, on this view, your change in music taste can be conceived as part of the broader transformative process of getting older. For example, we might think that changes in music taste are part of broader changes towards more mature and informed life priorities.

On this view, *Qualified Reflection* seems to instruct the twelve-year old teenager to stop listening to, or at least stop valuing, the particular type of music they currently enjoy. After all, they are quite confident now that they will no longer believe in the value of this type of music in the future and there seems to be no reason to think that this new belief will be developed irrationally or on the basis of misleading evidence.

Second, we might think that enjoying a type of music is not necessarily about believing in the value of this music. As a result, enjoying this type of music is very different from holding moral commitments (such as the ones in *Politics*). On this view, *Music* is much closer to the following case.

> *Olive:* As a twelve-year-old, you do not enjoy eating olives. However, from observing your older friends and relatives, you are quite confident that you will enjoy eating olives when you are twenty-five.

Many people would find it very counter-intuitive to apply *Qualified Reflection* in this case and instruct you to start eating olives now. The reason for this is that, in this case, it is not plausible to talk about a 'commitment' to or a belief in the value of olives. Instead, *Olive* involves a kind of *preference*.

Now, if we think that *Music* is similar to *Olive*, in involving a preference rather than a belief, then it would be equally inappropriate to apply *Qualified Reflection* in *Music*. So, on this interpretation, there is no rational pressure to stop listening to your music.

Personally, I think that both interpretations of *Music* are reasonable. There seem to be cases where you listen to a certain type of music because you feel strongly committed to its value, while there are other cases where you simply enjoy the entertainment it provides.

If it is reasonable to interpret *Music* as an example to which *Qualified Reflection* applies, then *I Can't Commit* has broader implications, beyond moral commitments. Many people believe not only in moral value but also in other types of value. These other types of value include aesthetic value, prudential value, etc. We also know that some of these people are likely to go through transformative experiences. As a result, many of the values they currently believe in are likely to change in the future. So, if we accept *Qualified Reflection*, it seems that, under some circumstances, these people have to suspend their current values and not replace them with other values. *I Can't Commit* thus becomes a puzzle not only about *moral* short-termism but about *value* short-termism more broadly.

6. Potential Solutions

Before concluding, I will discuss five potential solutions or responses to *I Can't Commit*.

6.1. Solution One: Should We Reject 'Transformative Experiences'?

I Can't Commit assumes that transformative experiences exist. One question we might ask, then, is: Can we avoid the puzzle by simply denying the existence of transformative experiences? Maybe, on this first, somewhat flat-footed objection, there are no events that transform us as radically as the concept of 'transformative experiences' assumes. Maybe, on this solution, we have overestimated the degree to which 'transformative' experiences change the moral commitments of the relevant people. For example, in a case such as *Military Supporter*, it might be that the young

person in question did not have any fully-fledged moral commitments before joining the army and it is only through the army that this person has acquired strong moral commitments. So, rather than radically changing their commitments, joining the army has created new commitments that did not exist before.

On this view, then, our experiences change us but not in a radical and fundamental way. We can be confident, on this view, that our core values, desires, and commitments stay relatively constant over time. And, if this line of reasoning is correct, and there are no such transformative experiences, *I Can't Commit* cannot be sustained. There would be no reason, on this view, to doubt that our long-term moral commitments would stay relatively constant over time.

6.1.1. Problems with Solution One

In reply, I want to make two points.

First, we could deny the particular concept of 'transformative experiences' as it is presented by Paul (2014). However, this should not make us think that there are no cases of transformation in the first place. In this chapter, I have discussed several obvious cases of transformation in which the affected people's core desires, values, and commitments are radically transformed. For example, it seems very hard to deny that the quite radical transformation from a progressive to a conservative political supporter in *Political Family* is a common phenomenon for many people. And, if there are obvious cases of transformation, then *I Can't Commit* remains.

Second, instead of simply rejecting the concept of 'transformative experience', it is worth noting the constructive role it can play. The concept acts as a *lens* which can help us to bring out the puzzle driving *I Can't Commit*. By focusing on the idea of 'personally transformative experiences' we can give a name to the radical transformation caused by many of our life experiences. We use the vocabulary developed by Paul (2014) and others to highlight the fragile nature of many of our core moral commitments. Ultimately, then, the concept of 'transformative experience' should not be seen as the unique way of writing about personal transformation. Rather,

I think that we should see it is a useful lens for bringing out the particularly delicate relationship between our selves now and our future moral commitments.

6.2. Solution Two: Should We Reject *Qualified Reflection*?

The second potential solution takes *I Can't Commit* as evidence that *Qualified Reflection* is not a good epistemic principle. If the principle really leads us into this puzzle, it might be natural to wonder whether it was a good principle to adopt in the first place.

The first alternative way of proceeding is to reject *Qualified Reflection* entirely both in cases in which we know how our future commitments will develop and in cases in which we do not know how our future commitments will develop. By rejecting the principle, we are no longer forced to give up our long-term moral commitments in the relevant transformative experience cases without knowing what they will be replaced by. Instead, we can hold onto our commitments up until the point where we have developed new commitments.

Another, perhaps less radical, way of proceeding is to hold onto *Qualified Reflection* in cases in which we know how our commitments will evolve in the future and reject the principle in transformative cases in which we do not know how our commitments will evolve in the future. Given the uncertainty surrounding our future moral commitments in certain transformative cases, it might be best, on this view, to hold onto our commitments in these cases until we are sure what our new commitments will be.

6.2.1. Problems with Solution Two

The main problem with this potential solution is to find a way of rejecting *Qualified Reflection* which is neither too radical nor arbitrary. On the first version of this solution, it seems too radical to reject *Qualified Reflection* entirely. After all, the principle strikes many people as a plausible principle for ordinary cases in which we know how our beliefs/commitments will change in the future. On the second version of this solution, it seems arbitrary to hold onto *Qualified Reflection* in cases in

which we know how our commitments will evolve and reject the principle in transformative cases in which we do not know how our commitments will evolve. As it stands, there is no good justification for applying the principle to one kind of case but not the other.

As I just noted, simply rejecting *Qualified Reflection* seems either too radical or too arbitrary. Having said that, this second solution to the puzzle does allow us to bring out an interesting dilemma that seems to underlie *I Can't Commit*.

Qualified Reflection pushes us towards giving up our long-term moral commitments. It exercises an *epistemic* pressure on us to give up commitments of which we know that they are likely to change radically in the future. Not giving up our commitments would be irrational on *Qualified Reflection*. In addition, however, there might be a *moral* pressure on us not to give up our long-term moral commitments. Usually, we give up our moral commitments only if we know *why* they are wrong. For example, most people would probably only want to give up their deeply held moral views about what constitutes a just distribution of resources if they knew why these views turn out to be wrong. In transformative experience cases, however, this is not the case. Our projected change of belief does give us some evidence that our current belief is wrong. After all, we take this projected change to be a rational response to the phenomenological evidence we will gain during the transformative experience. However, before undergoing the transformative experience we do not know *why* our current moral commitments are wrong.

These conflicting pressures leave us with the following dilemma: Either we give in to the epistemic pressure and reject our moral commitments, thereby going against our moral obligation of only giving up commitments of which we know why they are wrong. Or we give in to the moral pressure and stick to our moral commitments, thereby risking acting irrationally.[76]

---

[76] The conflict between epistemic and moral norms has also become a topic in the recent literature. For example, Boris Babic and Zoë Johnson King (2019, 2023) argue that in many moral encroachment cases (i.e. cases in which moral norms bear on our doxastic attitudes) we can find formal solutions which permit epistemic behaviour that is

One might think that the most constructive and least problematic way forward is to embrace the possibility of this epistemic/moral dilemma. On this view, *I Can't Commit* shows us that there are pro tanto moral and epistemic reasons to stick to or reject our moral commitments in transformative experience cases. Maybe, on this view, we can only decide what to do by making an all things-considered decision which weighs the relevant epistemic and moral reasons on a case-by-case basis.

There might be some cases where the moral pressure to hold onto our commitments is stronger than the epistemic pressure to give them up. For example, if you have deep moral commitments about how to treat your loved ones and you know that they are likely to change through a transformative experience, maybe the moral reason to stick to such fundamental commitments is stronger than the epistemic reason to give them up. Alternatively, there might be cases where the epistemic pressure is stronger than the moral pressure. For instance, the less central to your life moral commitments are, the more plausible it seems to give the epistemic reason priority over the moral reason. Finally, sometimes there might also be other reasons, such as prudential reasons, which can help us decide whether we should prioritise epistemic or moral reasons.

6.3. Solution Three: Should We Be Moral Short-Termists?

The third potential solution bites the bullet on *I Can't Commit*. The puzzle seems to force us to be moral short-termists in certain cases. But is this really such a bad outcome after all? Are there any reasons to think that being a moral short-termist in these cases is an appropriate moral attitude?

We might think that being moral short-termists in these cases reflects an openness and flexibility that moral agents should have. Rather than sticking dogmatically to our current beliefs,

---

also morally good. In fact, in the cases they consider, they argue that there is no conflict between epistemology and morality after all (Babic and King 2019, 8-15; Babic and King 2023, 12-19).

we prefer being open-minded about how our commitments evolve in the future. So, rather than seeing moral short-termism as a constraint, we should see it as a natural attempt to remain flexible and open-minded in an uncertain moral environment.

6.3.1. Problems with Solution Three

I can see two main problems with moral short-termism, in addition to its potentially counter-intuitive nature. The two problems stem from the roles that long-term moral commitments fulfil in our individual and social lives.

First, it is important to remember that long-term moral commitments, such as political or moral commitments, constitute an important part of who we are and how we see ourselves. Many people live their lives according to, or at least organise their lives around, some of their long-term moral commitments. The commitments provide important reference points that provide meaning to many of these people's actions.

Second, long-term moral commitments are often shared between large groups of people. For example, many people share the same progressive or conservative moral/political commitments. Given this shared nature, it seems plausible to assume that these commitments contribute to a feeling of social cohesion and togetherness. Long-term moral commitments are part of the fabric that holds society, or parts thereof, together.

Moral short-termism risks undermining the important individual and social roles that long-term moral commitments play. By forcing us to give up some of our long-term moral commitments, moral short-termism risks making our individual lives less fulfilling and our social lives less social.

6.4. Solution Four: Should We Draw on Other People's Experience(s)?

I Can't Commit seems to assume that transformative experiences are individual and isolated experiences for the people undergoing them. The puzzle tells us that there is *one* person going

through a transformative experience and, given that *this* person knows that their moral commitments will change but not in what way, they have to become a moral short-termist. On this fourth solution, this assumption is mistaken.

Rather than thinking of people undergoing transformative experiences as isolated individuals, we should think of them as situated in a social context in which many other people are going or have gone through similar transformative experiences. On this solution, individuals who know that they will undergo transformative experiences should draw on other people's experiences to see what moral commitments they should adopt now. This is true even in cases of transformative experience in which you do not know *how* your moral commitments will change.

For example, if you know that most people who have undergone experiences similar to *Politics* have adopted more moderate political commitments, then perhaps it is rational for you to adopt similar commitments now. On this solution, this is the case even if you are not sure yet what your actual future commitments will be. While you, as an individual, might not know what your future commitments will be, there is a lot of social/collective evidence to suggest what your commitments will be.

This fourth solution solves the puzzle by emphasizing the social context of transformative experiences. On this solution, looking at *other* people's transformations and their resulting commitments often gives us enough evidence to make rational decisions about what *our* moral commitments should be now.

The solution bears an interesting resemblance to some of the papers I discussed in the literature review. Like Srinivasan (2015), this solution suggests that we and hence also our (moral) commitments might not be as unique as we sometimes think. Instead, on this fourth solution, by recognising that our commitments are often similar to those of the people around us, we can draw on this social/collective evidence to make a rational decision. Similarly, in his decision-theoretic reply to Paul (2014), Pettigrew (2015, 2019) also suggests that we have enough collective evidence

(e.g. statistics) about other people's future commitment changes to be relatively confident about how our own preferences or commitments are likely to change in the future.

6.4.1. Problems with Solution Four

The main problem with the fourth solution is its limited scope. The solution cannot be applied to all transformative experiences in which *I Can't Commit* arises.

First, there seem to be cases of transformative experiences in which you do not have enough collective/social evidence to take a rational decision about your future moral commitments. These experiences might include rare transformative experiences or new transformative experiences. For example, we can imagine that the rapid pace of technological progress will lead to completely new immersive transformative experiences. Now, suppose you happen to be one of the first people to go through this kind of experience. You know that it will change your moral commitments but not in what way. Given that you are one of the first people to go through this kind of experience, there are no other people whose experiences you can rely on to know what your moral commitments should be now.

Second, there also seem to be cases of transformative experiences in which it is not clear what the collective/social evidence tells you to do. We can imagine a case where around fifty percent of people going through this experience developed certain moral commitments while the remaining fifty percent developed the opposite moral commitments. In this kind of case, in which the evidence is split, the fourth solution cannot offer us any way forward as to what is the rational thing to do.

6.5. Solution Five: Are Moral Commitments Really 'Unbreakable Vows'?

The fifth potential solution aims to show that there is nothing particularly puzzling about *I Can't Commit* because our moral commitments already accommodate the possibility of transformative experiences.

Recall that, on *I Can't Commit*, you have to give up your moral commitments in certain cases because you know that they will change but not in what way. The puzzle relies on there being something puzzling in the fact that you have to give up your commitments *against your will*. The reason *I Can't Commit* seems puzzling in cases such as *Politics* is that you want to hold onto your, say, left-wing ideas but it appears to be rational for you to abandon them nonetheless.

On this fifth solution, the puzzle seems puzzling only if we assume a particular view of moral commitments. Specifically, *I Can't Commit* has to assume a view of moral commitments as 'unbreakable vows'. On this view of commitments, people think of their moral commitments as irreversible and definitive commitments. Once we adopt this view of commitments, we can see why *I Can't Commit* might strike us as puzzling. Given that we think of our commitments as irreversible and definitive on this view, we find it unacceptable to give up our commitments in the cases of transformative experience discussed in this chapter.

However, on this fifth solution, there is an alternative view of moral commitments. Instead of thinking of moral commitments as unbreakable vows, we should think of them as containing a clause accommodating the possibility of later changes in commitment. For example, in *Military Supporter*, it seems implausible to think that the person prior to undergoing military service conceived their aversion to the military as irreversible and definitive. Rather, they had a strong aversion which they consciously or unconsciously thought might change in the future.

If we think of moral commitments in this alternative way, we can see why *I Can't Commit* seems much less puzzling than initially assumed. If we know that our moral commitments are likely to change but not in what way, giving up our commitments is less problematic on this view. We always knew that there was a possibility of change in the future and we thought of our moral commitments in this way. In other words, when committing to a cause or idea, we always leave open a door to change our mind if necessary.

6.5.1. Problems with Solution Five

The main problem with the fifth potential solution to *I Can't Commit* is also one of scope. Many people do seem to experience their moral commitments as unbreakable bonds, at least in many cases. Political (or moral) commitments seem to be exactly the kind of commitments that many people think of as definitive. For example, there is something strange in thinking that, in *Political Family*, your commitment to left-wing ideas contains a 'clause' about potential changes in the future. If there really was such a clause, we would probably not regard you as a *genuine* left-wing political activist. Instead, we would maybe perceive you as someone who is undecided or not ready to fully commit to their political cause.

So, while the fifth solution's view of moral commitments might seem plausible in some cases, it does not seem appropriate for moral or political commitments.

6.6. What to Do Now

I have presented and briefly reviewed five potential solutions to *I Can't Commit* (in addition to the non-cognitivist strategy of denying that moral commitments are a form of belief). All of these solutions have problems. On a personal level, I prefer solutions four and five — that is, drawing on the experience of others or adopting a new conception of commitments. It is true that both of these solutions do not have the scope to address all cases in which *I Can't Commit* might arise. Having said that, in the cases in which they do apply, the solutions offer genuine ways forward which take *I Can't Commit* and its underlying assumptions seriously.

Conclusion

We are sometimes quite confident that some of our future beliefs conflict with beliefs we currently hold. If we know that our future belief will be formed rationally and on the basis of veridical evidence, *Qualified Reflection* instructs us to abandon our present belief in favour of our future belief.

And even if we do not yet know what the content of our future belief will be, a generalisation of *Qualified Reflection* implies that we should abandon our present belief.

In the first part of this chapter, I suggested that *Qualified Reflection* applies to moral commitments, conceived as a form of belief. I argued that, if *Qualified Reflection* is true, then, under the right conditions, it is rational for us to adopt our future moral commitments now. This is a surprising result, as we tend to think that it is rational to stick to our current long-term moral commitments.

In the second part of the chapter, I extended my discussion of *Qualified Reflection* to cases of transformative experience in which we do not know *how* the transformative experience will affect our future moral commitments. I argued that, if we follow *Qualified Reflection*, it seems that, under certain conditions, it is rational to abandon our present moral commitments. This result is counterintuitive, and hence we have a puzzle. I argued that this puzzle seems to push us towards a form of moral short-termism.

In the third part, I discussed a way of extending *I Can't Commit* to other value commitments. I argued that, on one interpretation of cases such as *Music*, *Qualified Reflection* instructs us, under certain circumstances, to abandon our current value commitments and not replace them with new value commitments. So, although it is particularly counter-intuitive that we have to suspend our moral commitments in certain cases, *I Can't Commit* extends beyond these moral cases.

Finally, I reviewed five potential solutions to *I Can't Commit*. The solutions ranged from radical strategies such as rejecting transformative experiences or *Qualified Reflection* or being moral short-termists to more nuanced strategies of taking all things considered decisions, redefining what a moral commitment is or drawing on other people's experiences. I concluded this section by suggesting that the less radical solutions seem to me the most promising.

Chapter Five:
Conclusion

Introduction

Transformative experiences — experiences that radically transform us epistemically and/or personally — raise a number of moral issues. If we accept the existence of these experiences, some of our common moral attitudes and practices, including regret, moral evaluation, and the possession of long-term moral commitments, become problematic. I discussed these moral issues in detail in the previous four chapters.

The main purpose of this final chapter is to find out whether these specific moral issues can reveal any broader lessons about the relationship between transformative experiences and morality or rationality. I will argue that transformative experiences lead to ethical puzzles because they make it difficult for us to identify who the primary subject of our moral concern should be.

I proceed as follows.

First, I will take stock and consider where exactly the value of my thesis lies. This will also allow me to put this final chapter into its broader context. Second, I will discuss the idea that living through transformative experiences is comparable to death, or, at least, to going out of existence for the relevant self. Third, I argue that, given that transformative experiences are in many ways akin to death, they make it difficult for us to identify the proper subject of moral concern. This problem — the problem of identifying the proper subject of moral concern — I suggest, is at the heart of the three puzzles previously presented. While this idea has been discussed in the literature before[77], my puzzles will allow me to explain how important this idea is. Finally, I compare my approach in this final chapter to a discussion by Derek Parfit (1984) in *Reasons and Persons*, regarding

---

[77] See, for instance, Farbod Akhlaghi (2023), 'Transformative Experience and the Right to Revelatory Autonomy', *Analysis* 20 (20): 3-12.

personal identity and morality and rationality. This comparison will allow me to bring out two further general insights about the relationship between transformative experiences and morality.

1. Taking Stock

I want to start this final chapter by taking a step back and thinking again about the overall project of this thesis.

Traditionally, transformative experiences have been discussed mainly in decision theory. L.A. Paul's (2014) initial presentation of transformative experiences is meant to challenge the traditional decision-theoretic model according to which rational decisions are expected-utility-maximising decisions. Since the publication of Paul's book, as I noted throughout this thesis, several philosophers have responded to this challenge by offering different types of decision-theoretic solutions.

As I showed in Chapter One, there has been little discussion of *ethical* issues relating to transformative experiences. While there are a few papers dealing explicitly with the ethics of transformative experience, there has been no systematic discussion of transformative experiences in the ethics literature yet.

The overall goal of this thesis is to present a 'launchpad' for engaging with new issues in the ethics of transformative experience. I expand the ethics of transformative experience in two main ways.

First, my three puzzles offer new evidence for the claim that transformative experiences lead to ethical issues because they create a 'gap' between our old and our new selves. As mentioned before, this idea has been discussed in the literature. For instance, Farbod Akhlaghi (2023) builds much of his discussion of transformative experiences around the idea that these experiences create an ethically significant gap between our current and future selves. Akhlaghi challenges us to think about whether it is another person's current or future self which grounds our duty not to interfere with their transformative choices (Akhlaghi 2023, 5, 7). Ultimately, on his view, it is the person's

current self which grounds their right to revelatory autonomy and our duty of non-interference (Akhlaghi 2023, 7). However, Akhlaghi's use of the idea that transformative experiences create an ethically significant 'gap' between our current and future selves differs somewhat from mine. While Akhlaghi uses the idea to raise new challenges for our interpersonal relations, I use the idea to challenge some of our common moral attitudes and practices, including regret, moral evaluation and moral commitment.

Second, I emphasize that the ethical issues arising in the context of transformative experiences are more diverse than has previously been recognised. My puzzles highlight that transformative experiences lead to ethical issues for our *backward-looking* attitudes, *present* attitudes and *forward-looking* attitudes. And the puzzles' strong counterintuitiveness as well as the difficulty involved in solving them makes the ethics of transformative experience an exciting and potentially fruitful place for further work.

I started this broader project in Chapter One by providing the first comprehensive literature review for the ethics of transformative experience. I then tested the consequences of transformative experiences for ethics by providing three new puzzles which have not been discussed in the literature so far. Each of the three puzzles has a particular temporal focus. In Chapter Two, I discussed a puzzle about our *backward-looking* notion of moral regret. In Chapter Three, I presented a puzzle about how we can, *in the present*, morally evaluate a transformative experience that we have not been through. In Chapter Four, I discussed a puzzle about our *future moral commitments* and when it is rational to adopt them in the present.

In this final chapter, my goal is to pull the strings of the different chapters of the thesis together by exploring in more detail what it is about transformative experiences that leads to the three puzzles. The main focus of this final chapter will be on drawing out general insights that can help us to think about the overall relationship between transformative experiences and ethics.

I will suggest that the main lesson that we can draw from the literature review and the three puzzles is that transformative experiences raise ethical issues because their character makes it hard

to define who *the subject of our moral concern* should be. Again, while this lesson has been discussed in the literature before (e.g. Akhlaghi 2023), I hope to have demonstrated just how important it is, by showing the variety of kinds of puzzle to which it gives rise. In addition, later in this concluding chapter, I will offer a new comparison of this idea with some related ideas in Parfit's (1984) work on personal identity. The comparison will allow me to highlight how Parfit's argument and the puzzle *I Can't Commit* both suggest that there should be a narrowing of our rational concern towards the present. In addition, both some of Parfit's arguments and some of my arguments make it more natural for us to focus on the quality of relevant experiences, in an impersonal way, rather than on the subject of these experiences.

2. Who is the Moral Subject?

2.1. Going Out of Existence

The main characteristic that distinguishes transformative experiences from other types of experiences is the *radical transformation* of people that the former usually lead to. As Paul shows in her classic presentation of transformative experiences, these experiences can be radically transformative epistemically and/or personally (Paul 2014, 3, 16-18, 47). In Chapter One, I also argued that one of the main lessons we can draw from existing contributions to the ethics of transformative experience literature is the gap between former and current selves that transformative experiences create. It seems useful therefore to ask ourselves whether going through a transformative experience, particularly a personally transformative experience, is in some ways akin to death, or at least, to going out of existence for the self undergoing the experience.

In Chapter Two, I defined a self as essentially made up of core values, desires and commitments. The self before the transformative experience is composed of a particular set of core values, commitments and desires that are replaced by a new set of core values, commitments and desires after the experience. Any replacement of a sufficient number of these core constituents also seems to imply the replacement of the self. The idea that transformative experiences are akin

to going out of existence for the self is also confirmed by people who underwent transformative experiences. Primo Levi ([1958], 2000), known for his book *If This is a Man* in which he recounts his experience in a Holocaust concentration camp, describes the transformational impact of his experience as follows:

> 'Strictly speaking, I do not and cannot know *what I would be today if I had not been in the camp* [...] this would be, precisely, a case of describing *a future that never took place*. (emphasis mine)' (Levi [1958], 2000, 478)

The idea that transformative experiences can replace values, desires, and commitments and hence selves is also confirmed by scholars who study transformative experiences in disciplines other than philosophy. For instance, Jan Burzlaff, a historian who studied hundreds of Holocaust survivor testimonies at Yale University's Fortunoff Video Archive, writes about Holocaust survivors that they experienced 'trauma that shattered their prewar personalities and left them with *two separate selves*, each supported by irreconcilable memories' (Burzlaff 2021, 2). Burzlaff writes further that 'Holocaust survivors often mention a trigger moment, such as the loss of a loved one, that seems to have *divided* their personality — the personality that existed before the event and the war, and the one that existed afterward. There is a *sense of rupture*, the experience of a radical physical and emotional disorientation. (emphasis mine)' (Burzlaff 2021, 2).

2.2. Identifying the Moral Subject

If transformative experiences are like death, or at least like going out of existence, for the relevant selves, what does this mean for the relationship between ethics and transformative experiences more generally?

Many of our moral practices and attitudes rely on the clear identification of a single subject of moral concern. I understand 'subject of moral concern' broadly as someone who is the author

or subject of various moral attitudes and practices such as moral evaluation, blaming, regretting, etc. For example, if we want to blame someone for some action, we try to identify a blameworthy subject. The same seems true for the other moral attitudes discussed in this thesis. For example, if we morally regret a historical event, we often try to identify a subject for whom we can regret the event. Similarly, if we have moral commitments extending into the future, we want to be sure that these commitments will be *our* commitments.

However, transformative experiences can make it difficult, if not impossible, for us to identify an appropriate or single subject of moral concern. Once a person has undergone a transformative experience, there are at least two selves who are potential subjects of moral concern.[78] There is the self before and the self after the transformative experience. The idea of at least two selves competing for our moral attention has been discussed in the ethics of transformative experience literature before, and I am in agreement with the philosophers (e.g. Akhlaghi 2023) who have raised this idea. Having said that, my three puzzles offer new ways of bringing out this idea, thereby demonstrating just how important it is.

In Chapter Two I argued that there can be three moral subjects for which we wish to regret a historical tragedy: the victims killed by a historical tragedy, the surviving victims' old selves and the surviving victims' new selves. I argued further that we seem unable to regret the tragedy on behalf of the victims' new selves because these new selves possess lives worth living and only exist because of the tragedy. Without the tragedy, these new selves would not even exist. So, it seems that one of the lessons of Chapter Two's puzzle is that our moral attitudes — here in the form of moral regret — do not always have the scope to cover all relevant subjects of moral concern.

In *Only We Know*, the idea of competing units of moral concern manifests slightly differently. Recall that the main issue is that there seem to be some people who can morally

---

[78] Sometimes transformative experiences also make it hard to know when the old self has been replaced by the new self and hence who exactly should be the subject of moral concern.

evaluate certain transformative experiences because they have been through them while other people who have not been through these experiences cannot morally evaluate them. If we reframe the puzzle in a slightly more intrapersonal way, we can see that the problem of identifying a subject of moral concern also plays a role in this puzzle.

Suppose that you have not been through the transformative experience of, say, pregnancy, and you are trying to morally evaluate someone else's pregnancy experience. On this puzzle, we are left with a choice. First, we could morally evaluate the other person's experience based on the desires and preferences of their self *before* their pregnancy. However, if we take this approach, we fail to take into account the person's lived pregnancy experience. As I noted in Chapter Three, people's lived experience strongly shapes their desires and preferences such that we cannot deduce the relative importance and strength of these desires and preferences from how they were *before* the transformative experience. I also argued that there is something morally paternalistic or 'detached' about moral evaluations which ignore people's lived experiences.

Second, we could morally evaluate the person's experience based on the desires and preferences of their self *after* their pregnancy. However, if we take this approach, our own phenomenological knowledge is not rich enough to understand what this person went through and hence to morally evaluate their experience. So, if reframed in this more intrapersonal way, one of the main questions raised by *Only We Know* is which self — the one before the transformation or the one after the transformation — should be the focus of our moral evaluation of these experiences. And this question seems particularly intractable because both answers seem problematic in somewhat different ways.

In *I Can't Commit*, the selves competing for our moral attention belong to the same person. As I discussed in Chapter Four, the main problem raised by the puzzle is that we seem rationally unable to stick to our current moral commitments if we are relatively certain that we will live through a transformative experience that will transform our self in an unpredictable manner. The reason for this is that the transformative experience radically changes our self and hence also our

moral commitments. In other words, we are left with two selves and their respective commitments and we do not know which self should be the subject of our rational/moral concern.

3. A Comparison with Parfit on Personal Identity

My approach in this chapter might remind some readers of a famous discussion by Derek Parfit in Chapters Fourteen and Fifteen of *Reasons and Persons* (1984). In his discussion, Parfit explains the lessons about rationality and morality that he takes to follow from his views about personal identity (Parfit 1984, 307-350).

I will use the following subsections to explain relevant parts of Parfit's views in more detail and, above all, to examine whether there are any interesting similarities and differences between Parfit's view and my own approach in this chapter. I will highlight how both Parfit's argument and my own approach have a similar starting point — identity is sometimes fluid and uncertain — and how they both arrive at similar views about our moral commitments, our relationship to our future self and the prospects for an impersonal form of morality.

3.1. Parfit on Personal Identity

In his discussion of personal identity, Parfit focuses on two main questions:

a) What is a person?

b) What makes it the case that a person at two different times is the same person? What features are necessary for this to be the case? (Parfit 1984, 202)

Parfit's response to question a) is that a person is nothing above and beyond interrelated physical and mental events. This is a form of 'reductionism' about persons (Parfit 1984, 210). In order to

unpack Parfit's reductionism about persons a bit more, it is useful to start by considering his comparison of persons and nations.[79]

Here is one way of thinking about nations that Parfit does not find plausible. When we think about nations, on this view, there is a permanent entity called a 'nation', which exists separately from the association of relevant people that live in this nation (Parfit 1984, 211). For example, on this view, the nation of 'Brazil' is an entity called 'Brazil' that exists separately from the relevant association of people. This is a form of non-reductionism about nations because it does not reduce the nation to the people that compose it.

In contrast, on an alternative way of thinking, nations are composed of nothing more than an association of relevant people (Parfit 1984, 211). On this alternative view, there is no separately existing entity called, say, 'Brazil'. All there is to nations is the 'existence of a number of associated people' (Parfit 1984, 340). This alternative view is a form of reductionism about nations because it reduces the concept of 'nation' to the people that compose it. Parfit claims that most people find this second way of thinking about nations more plausible (Parfit 1984, 211-212).

According to Parfit, just as in the case of nations, we can be non-reductionists or reductionists about people. On a non-reductionist view of people, a person is essentially a soul or an 'I' that exists separately from the physical and mental events making up the life of that person. In addition, this soul or 'I' is permanent and unaffected by changes in the physical and mental events that make up the person's life (Parfit 1984, 210-211).

As in the case of nations, Parfit also believes that we should reject non-reductionism about people and be reductionists instead. More concretely, according to Parfit, a person is nothing above and beyond the existence of a set of interrelated physical and mental events. On this view, there is no permanent 'soul' or 'I' that exists separately from the interrelated physical and mental

---

[79] Parfit's comparison is inspired by a discussion of nations in David Hume ([1739] 2014) *A Treatise on Human Nature* (Oxford: Oxford University Press), p.261.

events (Parfit 1984, 211). One important feature of this view, Parfit argues, is that we can describe a person's life in a completely impersonal way. More precisely, by describing the interrelated mental and physical events that make up a person's life, we have, on Parfit's view, described all there is to this person's life (Parfit 1984, 212-213). (Below, in section 3.2, which deals with the moral consequences of reductionism about persons, I will show why Parfit's emphasis on the possibility of an impersonal description of a person's life matters morally.)

By being a reductionist about persons, Parfit has, in part, told us what he takes a person to be. However, this still leaves open question b) — what makes it the case that a person at two different times is the same person? According to Parfit, there are several answers to this question that are compatible with reductionism (Parfit 1984, 210). One of these views that is often associated with Parfit's work is the

Psychological Continuity View: Person A at time t1 is numerically identical to person B at time t2 only if person A and person B are 'psychologically continuous' — that is, only if there are overlapping chains of psychological connectedness between the relevant mental states of person A and person B (Parfit 1984, 207, 222).

For example, on this view, a person at the present moment t2 is numerically identical to a person one year ago at moment t1 because they are psychologically continuous — that is, they are psychologically connected through mental states such as memories, intentions, beliefs, etc.

Parfit cautions us about the limits and actual importance of the *Psychological Continuity View*. In cases which Parfit calls 'ordinary survival' — cases in which we consider whether person A survives an event as person B — we tend to think that psychological continuity matters because it allows for numerical identity between person A and person B (Parfit 1984, 261).

However, in order to question this typical way of reasoning, Parfit discusses the following case:

> *Fission*: You, person A, are scheduled to undergo brain surgery. Specifically, the surgeon will take one half of your brain out of your current body and place it into person B's surrogate body, and they will remove the other half of your brain and place it into person C's surrogate body. After the procedure, both person B and person C will wake up psychologically continuous with person A (Parfit 1984, 254-255).

The key issue, for Parfit, is to find out how we should think of the relation between person A and person B and person C. At first glance, it might seem like this is a relation of numerical identity. After all, the *Psychological Continuity View* tells us that psychological continuity is what we need for numerical identity between two people. According to Parfit, however, this cannot be true in *Fission*. Numerical identity is always a one-one and not a one-many relation. Put differently, one person — person A — cannot be numerically identical to both person B and C (Parfit 1984, 255-257).

If you, as person A, cannot be numerically identical to both person B and person C, what happens to you during *Fission*? According to Parfit, there are two points to consider in response to this question.

First, he argues that *Fission* allows us to illustrate the more general point that numerical identity is sometimes indeterminate. Put differently, there is no fact of the matter with whom, say, you, as person A, are numerically identical (Parfit 1984, 258). Second, what *Fission* helps us to highlight, Parfit argues, is that numerical identity cannot be the thing that actually matters to us, even in ordinary survival cases. After all, in *Fission*, we have everything that we take to matter in our ordinary survival cases — namely psychological continuity between person A and persons B and C — and yet, there is no one who is numerically identical to person A (Parfit 1984, 260-261).

From this, Parfit concludes that what actually matters in our ordinary survival is not numerical identity but rather only

Relation R: 'Psychological connectedness and/or continuity … with … any cause' (Parfit 1984, 215, 262)[80]

Notice that, unlike personal identity, *Relation R* can be a one-many relation. In *Fission*, for instance, Parfit argues that you survive because *Relation R* holds between you and person B and person C. In addition, according to Parfit, *Relation R* can come in different *degrees* — since you can be more or less psychologically connected or continuous with a past or future self, you can survive to a greater or lesser extent (Parfit 1984, 262, 313).

3.2. Comparing Approaches

Before explaining what Parfit takes to be the consequences of his views about personal identity for morality and rationality, it is worth briefly taking stock of Parfit's approach so far and comparing it to my strategy in this thesis and, specifically, in this final chapter.

Both approaches start by considering how living through certain experiences can make the person/self living through these experiences go out of existence. Parfit, focusing on persons, has offered thought experiments, such as *Fission*, which question our ordinary belief that psychological continuity implies survival as a numerically identical person. From this, Parfit argues that we should focus on selves and *Relation R*. Similarly, I have focussed on selves and how transformative experiences which make selves go out of existence. I argued that the existence of these experiences

---

[80] The caveat of 'with any cause' is meant to capture cases in which the psychological continuity or connectedness is not the result of standard continuity of the necessary body/brain parts (Sauchelli 2020, 51).

challenges the belief that our selves stay constant over time, an assumption behind many of our moral attitudes and practices.

Another interesting similarity is that both approaches highlight cases in which it is, in a sense, indeterminate or uncertain what will happen to the person/self who underwent the relevant experience. Parfit argues that it can sometimes be indeterminate to whom we are numerically identical, while I argue, in *I Can't Commit* for example, that transformative experiences often leave it uncertain which of the possible future selves we will possess after undergoing these experiences.

One important difference to highlight here is that Parfit's starting point is an argument about persons and personal identity. Once he has established the view that personal identity is not what matters, Parfit moves on to talking about 'selves' and *Relation R*. In contrast, I focus on changing selves right away and leave open the question of whether transformative experiences also have an impact on our numerical identity as persons.

3.3. What Follows for Morality and Rationality

In a second step, Parfit tries to work out the *consequences* of his initial arguments for rationality and morality more broadly (Parfit 1984, 307-350). I will here compare two of Parfit's arguments with some of my own remarks about transformative experiences and rationality and morality.

First, recall that, according to Parfit, *Relation R*, the relation that matters morally, can come in degrees. On this view, you can be more or less psychologically connected with your past or future self. Parfit even argues that, in some cases, your degree of connectedness to your earlier self is so weak that your earlier self can almost be seen as a different person than your current self (Parfit 1984, 313-314). For instance, Parfit writes that 'the conversion of a pleasure-seeking Italian youth into St Francis' is an example of a 'great discontinuity' which would warrant treating the Italian youth and St Francis as almost different people (Parfit 1984, 325).

According to Parfit, the fact that *Relation R* can come in degrees has moral implications. One of these implications, on Parfit's view, is that the less psychologically connected you are to

your past self who made a moral commitment, the less bound you are by this commitment now. According to Parfit, you can bind only *yourself* with a moral commitment. However, again, Parfit believes that, if there has been a great discontinuity between you now and your past self, you now can be treated as a different person than your past self. So, your past self's moral commitment cannot bind you now because you are a different person now (Parfit 1984, 326-329).

Parfit argues that another implication concerns how we ought to relate to our future self. According to Parfit, many people are inclined to believe that it is irrational to do what they think will make them worse off. He calls this view the 'Self-Interest Theory'. On this view, Parfit argues, it seems that you should be concerned as strongly about your near future as you should be about your distant future, provided that both are equally likely to occur. After all, if you believe that it is irrational to do something that makes you worse off, you cannot neglect your distant future because things you do now can strongly influence how well off you will be in that distant future (Parfit 1984, 307).

Parfit argues that the Self-Interest Theory is wrong in its claim that we should be equally as concerned about our distant future as we are about our near future. Recall that, on Parfit's view, the thing that matters to us in ordinary survival as well as in special cases such as *Fission* is *Relation R*. As noted above Parfit thinks psychological connectedness can come in different degrees — that is, it can hold between you now at time t1 and you at time t2 to a lesser or greater degree. He argues further that it is reasonable to assume that psychological connectedness usually holds to a lesser degree between you now and you in the distant future. Given that psychological connectedness is what matters and given that it holds to a lesser degree over long periods of time, Parfit argues that we should be less rationally concerned with our distant future than with our near future (Parfit 1984, 312-318).

Parfit's arguments about commitments and how we ought to relate to our future self bear some resemblance to my own arguments about the rational and moral significance of transformative experiences. In broad terms, both Parfit's argument and the puzzle *I Can't Commit*

reach a similar conclusion. They both conclude that there should be a *narrowing* of our rational concern towards the *present*. Having said that, the details of the two arguments differ.

The two positions differ slightly in the actual content of their conclusion. Parfit offers a more limited conclusion than *I Can't Commit*. He does not argue that, in certain cases, we should reject our long-term moral commitments and only adopt short-term moral commitments. Instead, Parfit argues that the more temporally distant our future self and its commitments are, the less rationally concerned we should be with them (Parfit 1984, 313-314, 326-328). The puzzle *I Can't Commit* offers a slightly stronger conclusion. It concludes that, in some cases, it is rational for us to suspend our long-term moral commitments and become moral short-termists.

The two positions also differ slightly in the way in which they reach their conclusions. Parfit arrives at the conclusion that we should be less concerned about our distant future with a more metaphysical argument. On his view, our concern for our future self (as well as the strength of our moral commitments across time) should mirror the strength of the psychological connectedness between our selves. On this view, it is a metaphysical fact (the degree of connectedness between selves) which determines the rational attitude to adopt (Parfit 1984, 313-314).

In *I Can't Commit*, I argued that, in some cases, we ought to be moral short-termists not only because transformative experiences will replace our current self but also because, on an epistemic level, the *Reflection Principle* seems to make it rationally obligatory to give up our long-term commitments in these cases. In contrast to Parfit, then, it is a metaphysical fact (transformative experiences replace our selves) *combined with* an epistemological principle which determines the rational attitude to adopt.

Despite their differences, both Parfit's argument and *I Can't Commit* leave us with a surprising and somewhat counterintuitive result. Usually, we feel very attached to our own future and our long-term moral commitments. However, the lesson to draw from this comparison with Parfit is that, if the metaphysical relationship between our different selves is precarious (and we

adopt certain epistemological principles), our rational concern should focus more on the present, i.e. on our current self and its short-term commitments.

Second, Parfit suggests that his reductionism about personal identity offers some support to utilitarianism. According to Parfit, reductionism about personal identity makes distributive moral principles — principles that assume that it matters morally to whom benefits and burdens are distributed — less plausible. In addition, he argues that by making these distributive principles less plausible, reductionism about personal identity supports utilitarianism (Parfit 1984, 333-342, 346).[81] Parfit offers several arguments for this conclusion, one of which I will now consider in more detail.

Parfit considers the

Principle of Equality: Benefits should be distributed equally between equally deserving people (Parfit 1984, 339).

According to Parfit, many people give this distributive principle at least some weight in their moral reasoning. He then points out that, in contrast, utilitarians do not endorse this principle (Parfit 1984, 339-340). One argument that utilitarians can give, according to Parfit, is the following: Many people believe that the mere fact of *when* something happens is not morally relevant as long as it does not affect the nature of the thing that happens (Parfit 1984, 340). If it is true that *when* something happens is not morally relevant, then, plausibly, it should also not be morally relevant *to whom* something happens, as long as this does not affect the nature of what happens. On this argument, 'when' and 'to whom' are merely differences in position that should not matter morally.

---

[81] According to Parfit, his views about reductionism do not make utilitarianism the correct moral view overall. Instead, he claims that they make utilitarianism more plausible than if reductionism about personal identity had not been true (Parfit 1984, 342).

The main problem with the *Principle of Equality* therefore, on this argument, is that it attaches moral importance to whom, i.e. equally deserving people, benefits should be distributed. In fact, however, on this utilitarian argument, as long as utility is maximised, it does not matter to whom or when these benefits are distributed (Parfit 1984, 340).

Parfit does not endorse this argument as such. However, he does argue that, if reductionism about personal identity is true, this utilitarian argument becomes more plausible than if reductionism is false (Parfit 1984, 342). According to Parfit, if reductionism is true 'it becomes more plausible, when thinking morally, to focus less upon the person, the subject of experiences, and instead to focus more upon the experiences themselves' (Parfit 1984, 341). The main reason for this, on Parfit's reductionist view, is that we are nothing more than interrelated physical and mental events. And, as we saw in the case of *Fission*, our personal identity is not what matters. And, if personal identity is not what matters, then, according to the argument that Parfit is considering, it is natural also to focus less on *whose* experiences we are referring to and more on *what* those experiences are like (Parfit 1984, 211, 340-341).

There is again an interesting comparison to be made between Parfit's proposed argument for utilitarianism and my own argument about the moral consequences of the three puzzles. Both arguments start from the premise that identity cannot be the thing that matters. According to Parfit, personal identity is not what matters, as shown through cases such as *Fission* (Parfit 1984, 260-261). Similarly, I have argued that transformative experiences make it difficult to identify the correct subject of our moral concern. After all, transformative experiences tend to make old selves go out of existence and new selves 'appear'. For example, as I showed in Chapter One, it is highly unlikely that you will possess the same self before and after living through a historical tragedy.

From this similar starting point, one might be tempted to offer a similar argument. Parfit's emphasis on the quality of the relevant experiences rather than their subject is echoed in my discussion of transformative experiences. Given the radical and often unpredictable changes caused by transformative experiences, we often cannot identify the unique self that is morally

relevant for evaluating any given transformative experience. As a result, just as in Parfit's discussion, it might become more natural to focus on the quality of the experiences, in an impersonal way, rather than on the subject of these experiences.

For example, as I showed in Chapter One, we are unable to identify the unique self for whom we could regret the relevant tragedy. However, given that we do not wish to give up entirely on regretting historical tragedies, it might be natural to approach the tragedies in a more impersonal way. On this view, we regret the impersonal negative quality of historical tragedies rather than the way in which they affected particular selves. Similarly, in *Only We Know*, the epistemic nature of transformative experiences (such as pregnancy) makes it difficult to give a personal moral evaluation of the relevant issues. However, a more impersonal evaluation of the relevant moral issues would avoid the epistemic limitations caused by transformative experiences.

Having said that, Parfit and I both express reservations about how strongly the arguments about reductionism/transformative experiences support a more impersonal approach. Parfit does not take his reductionism to show that utilitarianism is the correct overall moral view. Instead, he thinks that utilitarianism is more plausible if reductionism is true than if reductionism is false (Parfit 1984, 342). Similarly, I have expressed serious doubts throughout my three puzzles about the impersonal approach's ability to successfully solve the puzzles. But, nevertheless, these puzzles may make an impersonal moral view more tempting than it otherwise would be.[82]

So, what this point of comparison seems to show is that the more uncertainty there is about our metaphysical identity and the identification of the correct moral unit, the more we might be *tempted* to adopt an impersonal morality. However, as Parfit and I both emphasise, we should be cautious about how much we can conclude from these arguments.

---

[82] Parfit is also known for developing arguments, such as the so-called 'Repugnant Conclusion', which speak strongly against utilitarianism (Parfit 1984, 388). I have also mentioned these and other arguments which problematise an impersonal conception of morality throughout my discussion of my three puzzles.

Conclusion

The overall project of this thesis consists in offering a 'launchpad' for engaging with new issues in *the ethics of transformative experience*. My three puzzles provide fresh evidence that transformative experiences create an ethically significant gap between our current and future selves. In addition, I show how these puzzles offer new and previously unrecognised possibilities for work in the ethics of transformative experience.

In this final chapter, I brought the different strands of the thesis together by analysing what general insights the puzzles could provide about the relationship between ethics and transformative experiences. I started by arguing that going through a transformative experience is often akin to death or, at least, to going out of existence for the relevant selves. Transformative experiences leave us with at least two different selves, the self before and the self after the experiences. I suggested that the main lesson to learn from this is that transformative experiences can make it difficult to identify who our primary subject of moral concern should be. While this lesson has been discussed before, my three puzzles have shown how important it is and how varied the ethical issues are that arise from it.

In addition, I compared my discussion of transformative experiences with a discussion by Derek Parfit about personal identity and its implications for rationality and morality. Both discussions start by considering how a person's/self's metaphysical identity can be radically reshaped by the experiences they go through. In some cases, going through a certain experience can be like death for the concerned self. This has consequences for morality and rationality. I discussed two of these consequences in more detail. First, some of our experiences reduce the degree of connectedness between our past and current selves to such an extent that our commitments also reduce in strength or it even becomes irrational for us to have long-term moral commitments. Second, I discussed the possibility that the metaphysical uncertainty about our subject of moral concern supports an impersonal form of morality. The latter argument seems to

me particularly interesting, though the well-known problems with impersonal views mean that it should be treated with caution.

Bibliography

- Abizadeh, Arash (2008), 'Democratic Theory and Border Coercion: No Right to Unilaterally Control Your Own Borders', *Political Theory* 36 (1): 37-65.

- https://www.age-simulation-suit.com. Accessed 15.03.2023.

- Akhlaghi, Farbod (2023), 'Transformative Experience and the Right to Revelatory Autonomy', *Analysis* 20 (20): 3-12.

- Akhlaghi, Farbod (2024), 'The Moral Dangers of Drifting' at *The Practical, The Political, and the Ethical Seminar Series* (The Institute of Philosophy, School of Advanced Studies, University of London), 1-4. 02.05.2024.

- Atkinson, C.J. (2021), 'The Aim of Belief and Suspended Belief', *Philosophical Psychology* 34 (4): 581-606.

- Babic, Boris and Johnson King, Zoë (2019), 'Moral Obligation and Epistemic Risk' in Mark Timmons (ed.), *Oxford Studies in Normative Ethics: Volume 10* (Oxford: Oxford University Press), 81-105.

- Babic, Boris and Johnson King, Zoë (2023), 'Moral Encroachment Under Moral Uncertainty', *Philosopher's Imprint* 23 (12): 1-28.

- Bagnoli, Carla (2000), 'Value in the Guise of Regret', *Philosophical Explorations* 3 (2): 169-187.

- Barnes, Elizabeth (2009), 'Disability, Minority, and Difference' (2009), *Journal of Applied Philosophy* 26 (4): 337-355.

- Barnes. Elizabeth (2014), 'Valuing Disability, Causing Disability' (2014), *Ethics* 125 (1): 88-113.

- Barnes, Elizabeth (2015), 'Social Identities and Transformative Experience', *Res Philosophica* 92 (2): 171-187.

- Barnes, Elizabeth (2016), *The Minority Body: A Theory of Disability* (Oxford and New York: Oxford University Press).

- Beardman, Stephanie (2013), 'A Non-Factualist Defense of the Reflection Principle', *Synthese* 190: 2981-2999.

- Bentham, Jeremy (1970 [1789]), *The Collected Works of Jeremy Bentham: An Introduction to the Principles of Morals and Legislation* (Oxford: Oxford University Press).

- Billingham, Paul and Taylor, Anthony (2022), 'A framework for analyzing public reason theories', *European Journal of Political Theory*, 21(4): 671-691.

- Blackburn, Simon (1993), *Essays in Quasi-Realism* (New York: Oxford University Press).

- Boonin, David (2014), *The Non-Identity Problem and the Ethics of Future People* (Oxford and New York: Oxford University Press).

- Bradley, Richard (2012) 'Decision Theory: A Formal Philosophical Introduction' in Vincent F. Hendriks and Sven Ove Hansson (eds), *Handbook of Formal Philosophy* (New York: Springer), 615-655.

- Briggs, Ray (2009), 'Distorted Reflection', *Philosophical Review* 118 (1): 59-85.

- Briggs, Ray (2015), 'Transformative Experience and Interpersonal Utility Comparisons', *Res Philosophica* 92 (2): 189-216.

- Burzlaff, Jan (2021), 'The Holocaust Personality', *Fortunoff Video Archive for Holocaust Testimonies*,

  https://fortunoff.library.yale.edu/wp-content/uploads/2022/05/BurzlaffReflection.pdf. Accessed on 15.02.2024.

- Callard, Agnes (2018), *Aspiration: The Agency of Becoming* (New York: Oxford University Press).

- Carel, Havi and Kidd, Ian James (2020), 'Expanding Transformative Experience', *European Journal of Philosophy* 28: 199-213.

- Chan, Rebecca (2019), 'Transformed by Faith', *Faith and Philosophy* 36 (1): 4-32.

- Chan, Rebecca (2023), 'Transformative Experience' T*he Stanford Encyclopaedia of Philosophy* (Spring 2024 Edition), Edward N. Zalta & Uri Nodelman (eds.), forthcoming URL = <https://plato.stanford.edu/archives/spr2024/entries/transformative-experience/>.

- Chappell, Richard Yetter (2015), 'Value Receptacles', *Noûs* 49 (2): 322-332.

- Christensen, David (1996), Dutch-book arguments depragmatized: Epistemic consistency for partial believers, *The Journal of Philosophy* 93 (9): 450–479.

- Das, Nilanjan and Paul L.A. (2020), 'Transformative choice and the non-identity problem', in Sauchelli, Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (London and New York: Routledge), 187-208.

- DePaul, Michael R. (2004), 'Truth Consequentialism, Withholding and Proportioning Belief to the Evidence', *Philosophical Issues* 14: 91-112.

- https://www.empathyaction.org/workshops. Accessed on 15.03.2023.

- Enoch, David (2013), 'The Disorder of Public Reason', *Ethics*, 123(1): 141–176.

- Enoch, David (2014), 'Why I am an Objectivist about Ethics (And Why You Are, Too)', in Russ Shafer Landau (ed.), *The Ethical Life: Fundamental Readings in Ethics and Moral Problems* (New York: Oxford University Press), 192-205.

- Enoch, David (2015), 'Against Public Reason', in David Sobel, Peter Vallentyne, and Steven Wall (eds.), *Oxford Studies in Political Philosophy* 1 (Oxford: Oxford University Press), 112–142.

- Foley, Richard (1994), 'How should future opinion affect current opinion?' *Philosophy and Phenomenological Research* 54 (4): 747–766.

- Frick, Johann (2020), 'Conditional Reasons and the Procreation Asymmetry', *Philosophical Perspectives* 34 (1): 53-87.

- Friedman, Jane (2017), 'Why Suspend Judging?', *Noûs* 51 (2): 302-326.

- Golub, Camil (2019), 'Personal Value, Biographical Identity, and Retrospective Attitudes', *Australasian Journal of Philosophy* 97 (1): 72-85.

- Hare, R.M. (1952), *The Language of Morals* (Oxford: Clarendon).

- Harman, Elizabeth (2009), '"I'll Be Glad I Did It" Reasoning And The Significance Of Future Desires', *Ethics* 23: 177-199.

- Harman, Elizabeth (2015), 'Transformative Experiences and Reliance on Moral Testimony', *Res Philosophica*, 92 (2): 323-339.

- Heathwood, Chris (2016), 'Desire-Fulfilment Theory.' in Fletcher G. (ed.) *The Routledge Handbook of the Philosophy of Well- Being* (Routledge), 135-147.

- Heathwood, Chris (2020), 'An Opiniated Guide to "What Makes Someone's Life Go Best"' in Sauchelli, Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (London and New York: Routledge), 94-113.

- Hoefer, Carl (2016), 'Causal Determinism', *The Stanford Encyclopedia of Philosophy* (Spring 2016 Edition), Zalta, Edward N. (ed.), URL = <https://plato.stanford.edu/archives/spr2016/entries/determinism-causal/>.

- Horgan, T. and Timmons, M. (1992), 'Troubles For New Wave Moral Semantics: The 'Open Question Argument' Revived' *Philosophical Papers* (21): 153–175.

- Howard, Dana Sarah (2015), 'Transforming Others: On the Limits of "You'll Be Glad I Did It Reasoning"', *Res Philosophica* 92 (2): 341-370.

- Hume, David ([1739] 2014), *A Treatise on Human Nature* (Oxford: Oxford University Press).

- Isaacs, Yoaav (2020), 'The Problems of Transformative Experience', *Philosophical Studies* 177: 1065-1084.

- Jackson, Frank (1982), 'Epiphenomenal Qualia', *The Philosophical Quarterly* 32 (127): 127-136.

- Jemisin, N.K. (2010), *The Hundred Thousand Kingdoms* (London: Orbit).

- Kahane, Guy (2019), 'History and Persons', *Philosophy and Phenomenological Research* 99 (1): 162-187.

- Kahane, Guy (2021), 'The Significance of the Past', *Journal of the American Philosophical Association* 7 (4): 582-600.

- Krishnamurthy, Meena (2015), 'We Can Make Rational Decisions to Have a Child: On the Grounds for Rejecting L.A. Paul's Arguments', in Hannan, Sarah; Brennan, Samantha and Vernon, Richard (eds.), *Permissible Progeny?* (New York: Oxford University Press), 170-183.

- Kukathas, Chandran (2004), 'The Case for Open Immigration' in Andrew I. Cohen and Christopher Heath Wellman (eds.), *Contemporary Debates in Applied Ethics*. (Malden, MA; Oxford: Wiley-Blackwell), 376-390.

- Kumar, Rahul (2018), 'Risking Future Generations', *Ethical Theory and Practice* 21: 245-257.

- Levi, Primo ([1958], 2000), *If This Is a Man; The Truce* (London: Everyman's Publisher).

- MacAskill, William (2022), *What We Owe The Future: A Million-Year View* (London: Oneworld Publications).

- MacMahan, Jeff (1981), 'Problems of Population Theory', *Ethics* 92(1): 96-127.

- Medina, José (2013), *The Epistemology of Resistance: Gender and Racial Oppression, Epistemic Injustice, and Resistant Imaginations* (New York and Oxford: Oxford University Press).

- Mill, John Stuart (1987 [1863], *Utilitarianism* (Buffalo: Prometheus Books).

- Miller, David (2005), 'Immigration: The Case for Limits' in Andrew I. Cohen and Christopher Heath Wellman (eds.), *Contemporary Debates in Applied Ethics*. (Malden, MA; Oxford: Wiley-Blackwell), 193-206.

- Nagel, Thomas (1974), 'What Is It Like to Be a Bat?', *The Philosophical Review* 83 (4): 435-450.

- Narveson, Jan (1973), 'Moral Problems of Population', *The Monist* 57: 62-86.

- Nishimwe, Consolee (2012), *Tested to the Limit: A Genocide Survivor's Story of Pain, Resilience and Hope* (Bloomington, Indiana: Balboa Press).

- Nozick, Robert (1974), *Anarchy, State and Utopia* (Malden, MA; Oxford: Blackwell Publishing).

- Olson, E.T. (1997), *The Human Animal: Personal Identity Without Psychology* (New York: Oxford University Press).

- Ord, Toby (2021), *The Precipice: Existential Risk and the Future of Humanity* (London: Bloomsbury Publishing).

- Parfit, Derek (1984), *Reasons and Persons* (Oxford and New York: Oxford University Press).

- Parfit, Derek (2011), *On What Matters. Volumes I and II* (Oxford: Oxford University Press).

- Parfit, Derek (2012), 'We Are Not Human Beings', *Philosophy* 87 (1): 5-28.

- Pascal, Blaise ([1670], 2000), *Pensées* (Paris: Le Livre de Poche).

- Paul, L.A. (2014), *Transformative Experience* (Oxford and New York: Oxford University Press).

- Paul, L.A. (2015), '*Transformative Experience:* Replies to Pettigrew, Barnes and Campbell', *Philosophy and Phenomenological Research* 91(3): 794-813.

- Pettigrew, Richard (2015), 'Transformative Experience and Decision Theory', *Philosophy and Phenomenological Research* 91 (3): 766-774.

- Pettigrew, Richard (2019), *Choosing for Changing Selves* (Oxford and New York: Oxford University Press).

- Podgorski, Abelard (2021), 'Partiality, Identity, and Procreation', *Philosophy and Public Affairs* 49 (1): 51-77.

- Quong, Jonathan (2011), *Liberalism Without Perfection* (Oxford: Oxford University Press).

- Quong, Jonathan (2013), 'On The Idea of Public Reason' in Jon Mandle and David Reidy (eds.), *The Blackwell Companion to Rawls* (Oxford: Wiley-Blackwell), 265–280.

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

- Rawls, John (1993), *Political Liberalism* (New York: Columbia University Press).

- Raz, Joseph (1998), 'Disagreement in Politics', *The American Journal of Jurisprudence*, 43: 25–52.

- Sartre, Jean-Paul ([1943], 2017), *L'être et le néant: Essai d'ontologie phénoménologique* (Paris: Gallimard).

- Sauchelli, Andrea (2020), 'Introduction to Part Three: Personal Identity', in Sauchelli Andrea (ed.), *Derek Parfit's Reasons and Persons: An Introduction and Critical Inquiry* (Oxon and New York: Routledge), 48-67.

- Schroeder, Mark (2010), *Noncognitivism in Ethics* (London: Routledge).

- Setiya, Kieran (2016), 'Retrospection', *Philosophers' Imprint* 16 (15): 1-15.

- Shafer-Landau, Russ (2003), *Moral Realism: A Defence* (Oxford: Clarendon Press).

- Smilansky, Saul (2020), 'The Moral Evaluation of Past Tragedies: A New Puzzle', *Journal of Moral Philosophy* 17: 188-201.

- Srinivasan, Amia (2015), 'All the Same', *The Times Literary Supplement*, no. 5854, 8.

- Sturgeon, Scott (2010), 'Confidence and Coarse-Grained Attitudes' in Tamar Szabó Gendler and John Hawthorne (eds.) *Oxford Studies in Epistemology Volume 3* (Oxford: Oxford University Press), 126-149.

- Thomson, Judith Jarvis (1971), 'A Defense of Abortion', *Philosophy&Public Affairs* 1 (1): 47-66.

- Tilton, Emily and Toole, Briana (forthcoming), 'Standpoint Epistemology and the Epistemology of Deference' in Kurt Sylvan, Matthias Steup, Ernest Sosa and Jonathan Dancy (eds.) *Blackwell Companion to Epistemology* (Malden, MA: Wiley-Blackwell).

- Ullmann-Margalit, Edna (2006), 'Big Decisions: Opting, Converting, Drifting', *Royal Institute of Philosophy Supplement* 58: 157–172.

- United Nations, Africa Renewal: 'I was tested to the limit – Rwanda genocide survivor'. Accessed on 13.10.2022 at https://www.un.org/africarenewal/web-features/i-was-tested-limit-—-rwanda-genocide-survivor.

- USA Today News, 'Vietnam War: 6 personal essays describe the sting of a tragic conflict'. Accessed on 13.10.2022 at https://eu.usatoday.com/story/news/world/2017/09/11/vietnam-war-voices/105499474/.

- Van Fraassen, C. (1984), 'Belief and the Will', *The Journal of Philosophy* 81 (5): 235-256.

- Van Fraassen, C. (1995), 'Belief and the Problem of Ulysses and the Sirens', *Philosophical Studies* 77 (1): 7-37.

- Velleman, J. David (1996), 'Self to Self', *The Philosophical Review* 105 (1): 39-76.

- Velleman, J. David (2015), *Beyond Price: Essays on Birth and Death* (Cambridge: Open Book Publishers).

- Villiger, Daniel (2024), 'Transformative Experience', *Philosophy Compass* 19 (6): 1-12.

- Wallace, R. Jay (2013), *The View From Here: On Affirmation, Attachment, and the Limits of Regret* (Oxford and New York: Oxford University Press).

- Williams, Bernard and Smart, J.J.C. (1973), *Utilitarianism: For and Against* (Cambridge: Cambridge University Press).

- Woollard, Fiona (2021), 'Mother Knows Best: Pregnancy, Applied Ethics, and Epistemically Transformative Experiences', *Journal of Applied Philosophy* 38 (1): 155-171.

- Woollard, Fiona (2022), 'Your Mother Should Know: Pregnancy, the Ethics of Abortion, and Knowledge through Acquaintance of Moral Value', *Pacific Philosophical Quarterly* 103 (3): 471-492.

- Wylie, Alison (2003), 'Why Standpoint Matters', in Robert Figueroa and Sanda G. Harding (eds.) *Science and other Cultures: Diversity in the Philosophy of Science and Technology* (New York and London: Routledge), 26-48.