

**From the holdings of the Department of Special Collections , Princeton University Library,
Princeton, NJ**

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyright law. We do not charge any permission or usage fees for the publication of images of material in our collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff. For more information regarding copyright, credit, and citation please review the information on the Copyright, Credit and Citation Guidelines page of our website.

<https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines>. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. <https://library.princeton.edu/special-collections/ask-us>

Doctoral Dissertations, 2011-2026

Philosophy

2023

Past and Possibility: Essays on Value, Justice, Time, and Wasted Potential

Masny, Michal

Princeton, NJ : Princeton University

<http://arks.princeton.edu/ark:/88435/dsp011z40kx08p>

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. <https://www.copyright.gov/title17/title17.pdf>

PAST AND POSSIBILITY

*Essays on Value, Justice,
Time, and Wasted Potential*

Michal Masny

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY
THE DEPARTMENT OF PHILOSOPHY

Advisers: Johann Frick, Thomas Kelly,
Jacob Nebel, and Gideon Rosen

May 2023

© Copyright by Michal Masny, 2023.

All rights reserved.

Abstract

This dissertation consists of five papers which examine relationships between value, justice, time, and wasted potential.

In Chapter 1, “Wasted Potential”, I challenge the widely accepted view that how good a life is for a person depends only on what actually happens within it, such as its pleasures, achievements, and loving relationships. Instead, I argue that it also depends on the individual’s potential: what experiences, achievements, and relationships they could have had but perhaps didn’t.

In Chapter 2, “The Shape of History”, I consider whether it is better if the history of humanity features a pattern of improvement rather than deterioration, holding other things equal. I argue that it typically is, use a broadly conservative view about value to explain that, and consider the implications of my view for matters related to the future of humanity.

In Chapter 3, “Conservatism about Prudential Value”, I argue that we have a distinctive moral reason to preserve certain prudential goods, such as loving relationships and important personal projects, even when a superior replacement is available. I then show how this view illuminates several old and new puzzles about navigating childhood, adulthood, older age, and death.

In Chapter 4, “Extension and Replacement”, I develop a novel conservative account of when and why it is better to extend the length of a happy life rather than to create a new happy life, even if the total welfare is the same in both cases. I then show how this account also applies to the choice between extension and replacement with respect to non-human animals and humanity as a whole.

In Chapter 5, “Healthspan Extension, Completeness of Life, and Justice”, I argue that justice requires that we provide people with sufficient opportunities to have a ‘complete life’, that many people currently lack such opportunities, and that unconditional access to anti-ageing technology would substantially improve the status quo.

Acknowledgements

I arrived in Princeton, one large suitcase in hand, on a typically hot and humid afternoon in August of 2018. It was my first time in New Jersey, the United States even. Unlike most of my peers, I did not visit the university as a prospective student. On paper, it was an almost unreservedly better place to continue my education than Oxford. I did not schedule any virtual meetings either, sceptical of the value of inevitably brief and potentially awkward conversations in that setting. That was before the pandemic.

Only a few years later did I learn that the philosophy department at Princeton is notorious for making hiring decisions in the same fashion. No interviews, no flyouts; paper doing all the talking. While unusual, this approach seems to work. It certainly did for me. The last five years have been intellectually and personally transformative, and I leave Princeton with a deep sense of gratitude.

My greatest debt is to my primary advisor, Johann Frick, for his unwavering support, intellectual guidance, and contagious enthusiasm for the craft. From day one, Johann has been my favourite interlocutor, the first and the last person I wanted to talk to about each of my projects. The quality and generosity of his philosophical feedback are unmatched. He has also taught me that every aspect of life is ripe with interesting ethical questions and that getting things right is only half of the task. I see and write differently now.

I am also thankful to my other advisors. Tom Kelly commented on my writings with his trademark assertiveness and imaginative quality, and I have spent many hours thinking about the issues and examples raised in our conversations. Jake Nebel first read my work long before he became my advisor, when as a master's student at Oxford I reached out to him on the encouragement of Jeff McMahan. He was able to identify some of the most piercing objections to my arguments and draw my attention to controversial presuppositions that I, and others, have overlooked. Gideon Rosen read the materials I sent to him with one eye on the big picture, and his input has been critical in identifying connections and tensions within my work, as well as new avenues for inquiry.

As I reflect on the feedback that I received from my advisors over the years, I am struck by how coordinated their individual efforts would appear to an outsider. Despite coming from four different philosophical generations, they did not contradict each other, instead presenting a clear path forward. There was no redundancy either: hardly any comment or question was repeated, as they naturally honed in on distinct aspects of my work.

Halfway through my fourth semester, the academic world came to a grinding halt. The COVID-19 pandemic forced us to retreat home, embrace new methods of communication, and establish new routines. During this time, I was fortunate to

have a friend and interlocutor in Erik Zhang. I learned a lot from our evening conversations. I was also supported by members of my extended philosophical family, many of whom I met through or left behind at Oxford. For many hours of discussions and pages of rich feedback over the years, I am grateful to Aleksander Domoslawski, Tomi Francis, Todd Karhu, Kacper Kowalczyk, and Aidan Penn. I hope we will all get to work under one roof again one day.

Thanks also to other participants of the virtual ethics discussion group we started during this time: James Evershed, Jessica Fischer, Petra Kosonen, Ben Lange, Adriano Mannino, Joseph Moore, Milan Mosse, Korbinian Rueger, Rhys Southan, Margaret Shea, Bastian Steuwer, Leora Sung, Elliott Thornley, Alice Van't Hoff, Nikhil Venkatesh, and others. Our sessions helped me make strides in my early-stage projects and kept my spirits high.

Around the same time, Johann Frick entered the academic transfer market, contributing to the uncertainty of that period. With lucrative offers from all over the country, and after months of deliberation, he eventually settled for the temperate climate and lush campus of UC Berkeley. I followed him there without hesitation as a visiting student, founding a nest in a small foodie co-op in North Berkeley which at one point boasted a personal assistant to Alice Waters and a classically trained French chef with multiple Michelin-starred restaurants under his belt. I swapped early-morning weekday runs along the Delaware & Raritan Canal northbound for golden-hour runs down Virginia St and by the Bay; and weekend long runs along the same canal southbound for outdoor adventures in Tilden, Half Moon Bay, Santa Cruz, Marin Headlands, Point Reyes, Mount Diablo, Tahoe, and Yosemite. During this time, I also met several people whose lust for life has influenced me profoundly: especially Amanda, Beba, and Natalia.

Invigorated by these experiences, I found myself entering new philosophical territories. The ideas about the ethical issues concerning the future of humanity and emerging technologies that make up the last three sections of this dissertation were all conceived during my first year in California. I am grateful for the feedback I received while developing these ideas from Jay Wallace, Niko Kolodny, and graduate students at UC Berkeley. This chapter of my life is set to continue for another couple of years. In the fall, I will join UC Berkeley as a postdoctoral fellow at the Kavli Center for Ethics, Science, and the Public.

Contents

Introduction.....	7
PART I: VALUE AND WASTED POTENTIAL	
Chapter 1: Wasted Potential: The Value of a Life and the Significance of What Could Have Been	16
PART II: VALUE AND TIME	
Chapter 2: The Shape of History	59
Chapter 3: Conservatism about Prudential Value.....	98
Chapter 4: Extension and Replacement	140
PART III: JUSTICE AND TIME	
Chapter 5: Healthspan Extension, Completeness of Life, and Justice	179
Bibliography.....	204

Introduction

In this dissertation, I aim to advance our understanding of two central topics in ethical theory: *value* and *justice*. One aspect of my discussion is critical. I claim that philosophers have largely overlooked the significance of certain temporal and modal features, such as the order in which things happen and what else could have happened, for theorising about these topics. Another aspect of my discussion is constructive. I defend a set of principles that account for the significance of these temporal and modal features, and explain how to incorporate them into complete theories of value and justice.

The dissertation is divided into three parts and five chapters. The first part concerns the significance of modal features for theorising about value. To get a grip on the central idea, consider the case of Sophie Germain, a French mathematician of the early nineteenth century. She was born to a wealthy family and enjoyed a life rich in meaningful relationships, sophisticated pleasures, and significant achievements. However, much of her exceptional academic talent was wasted because of the obstacles she faced as a woman. Early on, her parents tried to hinder her fascination with mathematics. Later, she was barred from attending the Ecole Polytechnique and the meetings of the Paris Academy of Sciences, and her contemporaries regularly ignored her manuscripts and published work.

Germain's case evokes an attitude of *evaluative ambivalence*. On the one hand, her story is uplifting: she experienced much of what makes life valuable. On the other hand, there is a sense of tragedy that we just cannot shrug off: she could have received more appreciation, avoided many frustrations, and achieved much more, but she didn't.

Situations which have this structure suggest something important for value theory. According to the orthodox view, how good a life is for a person depends only on what actually happens within it, such as its pleasures, achievements, and loving relationships. In "Wasted Potential" (Chapter 1), I challenge this view and argue that it also depends on the individual's potential: what experiences, achievements, and relationships they could have had, but perhaps didn't.

The second part of the dissertation concerns the significance of temporal features for theorising about value. In particular, I examine how the fact that our lives are situated between the past and the future affects what is valuable and what we have reasons to do.

To get a feel for the core idea, take the example of the Golden Gate Bridge, which connects the bustling city of San Francisco with the idyllic landscapes of Marin Headlands. Suppose that we could destroy it and build an even more impressive structure in its place. Like many people, I have a strong sentiment that we shouldn't do it. Crucially, that's not just because this project would be costly or inconvenient, but more importantly because the Golden Gate Bridge already exists

and has a special kind of value. I think the same is true of many other valuable things: we have a distinctive moral reason to preserve them, even if they could be replaced with something better. G. A. Cohen (2012) dubbed this view ‘conservatism about value’. In this part of the dissertation, I discuss both the theoretical foundations and the practical implications of this broadly conservative view about value.

In “The Shape of History” (Chapter 2), I discuss a new puzzle about value and time. Consider two ways in which the condition of humanity could change over time. In the first scenario, our history starts with hardship and ends with flourishing, so there is a discernible pattern of improvement. In the second, the opposite happens. These scenarios are permutations of each other—in the sense that for every period of flourishing in one course of history, there is exactly one such period in the other, and so on—which, on one kind of view, suggests that they are equally good. However, many people believe that an improving history is better than a deteriorating one. So, for example, if a god was deciding which of these two worlds to create, they should create the improving one. In this chapter, I argue that we cannot provide a compelling rationale for this judgment by extending any of the popular answers to the structurally similar puzzle about the shape of a life. I then develop a novel account rooted in conservatism about value.

In “Conservatism about Prudential Value” (Chapter 3), I propose expanding the scope of conservatism about value. Discussions of this view—by G. A. Cohen,

Samuel Scheffler, Johann Frick, and Jacob Nebel—have focused on things which are *good simpliciter*, such as traditions and artworks. Drawing on real-life and fictional cases, I argue that we also have a distinctive moral reason to preserve certain *prudential goods*, such as loving relationships and important personal projects, even when a superior replacement is available. I then show that this view illuminates several old and new puzzles about navigating youth, adulthood, older age, and death.

In “Extension and Replacement” (Chapter 4), I draw on insights from the preceding two chapters to address issues concerning preserving entire lives and civilisations. Many people believe that it would be better to extend the length of a happy life than to create a new happy life, even if the total welfare would be the same in both cases. Likewise, many people believe that it would be better to extend humanity’s tenure rather than to let it be replaced by another intelligent and sentient species, even if the total welfare would be the same in both cases. Despite the popularity of these views, however, one would be hard-pressed to find a fully compelling justification for them in the literature. In this chapter, I rectify this situation and provide a unified account of when extension is better than replacement that appeals to conservatism about value.

The third part of the dissertation concerns the significance of temporal features for theorising about justice. In particular, I consider whether the length of life should be a matter of distributive justice, in the same way that wealth and health are. My discussion of this issue is animated and framed by recent progress in

geroscience which holds the promise of significantly slowing down—or even reversing—the ageing process and age-related diseases, such as cancer, heart disease, and dementia. Crucially, these anti-ageing technologies have the potential to extend not just our lifespans but also healthspans. That is, in addition to increasing the number of years that we are alive, they could also increase the number of years we remain in a state of physical, cognitive, and reproductive health.

In “Healthspan Extension, Completeness of Life, and Justice” (Chapter 5), I argue that insofar as possible, we have a justice-based duty to develop and widely distribute such technology. This is because justice requires that each person be provided with sufficient opportunities to have a ‘complete life’ (i.e., a life that involves, in adequate quantity and quality, *all* fundamental ingredients of well-being), that many people currently lack such opportunities, and that we would substantially improve the status quo by giving them unconditional access to anti-ageing technologies.

Overall, my dissertation has several noteworthy theoretical and practical upshots. For example, the discussion of value and wasted potential in Part I implies that an adequate theory of the goodness of a life must take a dual form: how good a life is for a person is determined jointly by their well-being and the degree to which they realise their potential, where the latter is understood in terms of how close one is to their maximal possible well-being.

Moreover, the discussion of value and time in Part II implies that because our existence is temporally situated, we cannot start a new life or society whenever we please. Existing cultural artefacts, loving relationships, and important personal projects warrant a kind of loyalty, which gives us a reason to conserve them. Moreover, that discussion provides an insight into the normative status of our loving relationships and important personal projects. We should be concerned with the survival and condition of these things above and beyond the contribution that they make to our well-being, and thus some of our reasons to preserve them are moral and agent-neutral, just like our reasons to preserve the Grand Canyon or the Mona Lisa.

Finally, the discussion of justice and time in Part III demonstrates that new possibilities afforded by emerging technologies can significantly influence the demands that justice places on us. In particular, the prospect of direct and substantial healthspan extension makes salient the fact that justice should be concerned with the length of life, and it might also force us to revise the policies through which our society pursues the ideal of justice.

In closing, I would like to briefly comment on one recurrent issue in this dissertation that has animated much of my philosophical thinking in the last few years: a puzzle about the temporal shape of value. A careful reader will notice that in Chapters 1 and 3, I propose what appear to be competing solutions to a long-standing puzzle about the shape of a life. In Chapter 1, I claim that we can explain

when and why an improving life is better than a deteriorating life by appealing to the evaluative significance of wasted potential. Then, in Chapter 3, I suggest that we can explain the same issue by appealing to the evaluative significance of conserving certain prudential goods over time.

But there need not be any tension here: these accounts might prove to overlap. In particular, it may turn out that the superiority of an improving life over a corresponding deteriorating life is explained just as well by a difference in terms of the realisation of one's potential and by a difference in terms of the conservation of certain prudential goods. To illustrate this, consider the case of a deteriorating life. At least on the face of it, when a person fails to conserve their loving relationships and important personal projects, and lets them be replaced by something worse, they also fail to make the most of their potential. Alternatively, a more careful inquiry might reveal that these accounts complement each other. There might be edge cases in which only one of these accounts identifies an evaluatively significant difference between an improving and a deteriorating life, whereas the other remains silent about that comparison. Both of these possibilities, overlap and complementarity, raise further interesting questions about the relationship between conservation and realising one's potential, but I leave them for another occasion.

NB: Two chapters of this dissertation have already been published. Chapter 1 appeared as “Wasted Potential: The Value of a Life and the Significance of What Could Have Been” in *Philosophy & Public Affairs*, 51, no. 1 (2023): 6-32. Chapter 5 appeared as “Healthspan Extension, Completeness of Life, and Justice” in *Bioethics*, 37 (2023): 239-245.

Part I: Value and Wasted Potential

Chapter 1:

Wasted Potential: The Value of a Life and the Significance of What Could Have Been

1. Introduction

It is a striking fact about us that we care deeply about what could have happened, but didn't. In this paper, I explore the significance of this concern for theorising about the goodness of a life. To get a feel for the issues that we will be examining, consider the case of Sophie Germain, a French mathematician of the early 19th century. She was born to a wealthy Parisian family and enjoyed a life rich in meaningful relationships, sophisticated pleasures, and important achievements. However, much of her exceptional academic talent was wasted because of the obstacles she faced as a woman. Early on, her parents tried to hinder her youthful fascination with mathematics. Later, she was barred from attending the École Polytechnique and the meetings of the Paris Academy of Sciences, and both her manuscripts and published work were regularly ignored by her contemporaries.

Germain's case evokes an attitude of *evaluative ambivalence*. On the one hand, her story is uplifting: she experienced a lot of what makes life valuable. On the other, there is a sense of tragedy that we just cannot shrug off: she could have achieved much more, but didn't, and could have been more appreciated and spared many frustrations, but wasn't.

Situations which have this general structure suggest something important for value theory. According to the orthodox view, how good a life is for its subject depends exclusively on the things that actually happened within it, such as its pleasures and pains, the satisfaction of its subject's preferences, or the presence of various objective goods and bads. In this paper, I challenge this orthodox view and argue that the goodness of a life also depends on what could have happened, but didn't.

My discussion has three parts. In the first part (§2-3), I present the initial motivation for *the Dual Theory*, which holds that how good a life is for someone is determined jointly by their level of well-being and the degree to which they realise their potential. In the second part (§4-5), I examine different ways in which this theory can be spelled out. In particular, I consider what counts as an individual's potential, and how well-being and the realisation of one's potential combine to determine the overall goodness of a life. In the third part (§6-8), I provide additional support for the Dual Theory by highlighting its explanatory power. Specifically, I argue that it helps us resolve three ethical puzzles concerning the standards for a life worth living for non-human animals, the significance of a life's shape, and the badness of death.

Three preliminary remarks will make it easier for the reader to appreciate the scope of this discussion and its position in the broader theoretical landscape.

First, a few words about the key concepts. This paper is about the goodness of a life for its subject or, equivalently, the prudential value of a life. Many philosophers assume that there is only one dimension of the good life, well-being, which leads them to use these two terms as synonyms. For the present purposes, however, we want to use the ‘the goodness of a life’ as a more encompassing notion than ‘well-being’. This terminological convention has an influential precedent: for example, Susan Wolf (2010), Antti Kauppinen (2012), and Aaron Smuts (2017) hold that how good a life is for a person is determined not just by their well-being, but also by how meaningful it is. This paper, too, contends that there are at least two dimensions along which we can assess the goodness of a life: well-being and the realisation of one’s potential.¹

Second, some context. Although the idea sketched above has not yet been discussed in contemporary value theory, there are some thinkers who have come within a stone’s throw of appreciating it. Jeff McMahan (1996) and Peter Vallentyne (2005), for instance, consider the idea that facts about an individual’s potential for well-being bear on whether we have egalitarian reasons to redistribute limited resources to that individual.² Moreover, Johann Frick and Adam Lerner (ms) suggest that similar considerations influence whether it is permissible to bring an individual

¹ I do not take a stance here on whether meaningfulness also affects the goodness of a life.

² For a recent discussion of this issue, see Kagan (2019).

into existence. Finally, R. Jay Wallace (2013) claims that what we have reasons to regret in our lives depends in part on what else could have happened.³

If the arguments of this paper are sound, however, the significance of what could have been is more pervasive than these philosophers have contended. For one thing, this significance is not restricted to the first-personal, retrospective standpoint from which we tend to assess the aptness of attitudes like regret. For another, it is already present at the level of prudential value, and its bearing on deontic matters—such as those concerning the demands of justice and the reasons that apply to the creation of individuals—is at most derivative.

Third, a word on methodology and aims. This paper does not offer anything like a formal proof of the Dual Theory. In value theory, and at a certain level of generality, a careful discussion of cases and application of the method of reflective equilibrium are sometimes the best tools we have. It is in this spirit that I present Sophie Germain's case and several other scenarios in which, it seems to me, facts about an individual's wasted potential bear on the overall goodness of their life. I am conscious that some of these intuitive judgments are not shared by everyone, but I think that they are common enough to call for a careful philosophical inquiry. I am also aware that this paper alone may not suffice to convince the reader that the Dual Theory is true. The orthodox view, on which the goodness of a life depends

³ See also Harman (2009) and Setiya (2016) for discussions of this issue.

exclusively on the things that actually happened within it, is so deeply entrenched in ethical theory that uprooting it is a monumental task. This paper is just the first step in that project and it will have succeeded if it persuades the reader that, despite its highly controversial nature, the Dual Theory *may* be true and merits further consideration.

But let's start at the beginning, with a closer look at the phenomenon of evaluative ambivalence, which we have identified earlier.

2. Evaluative ambivalence

Consider the case of Sophie Germain again. By the lights of any standard theory of well-being, she had an excellent life, even once we factor in the disappointments and the struggles. And yet it is clear that things could have gone much better for her. This makes her life seem at once heartening and tragic. Crucially, this evaluative ambivalence persists even as we continue to think long and hard about this case. Germain's life does not suddenly become just heartening or just tragic. A full and accurate description of the quality of her life seems to preclude merely summing the two evaluations, positive and negative. Instead, it requires that we maintain both judgments at once.

The enduring character of evaluative ambivalence in such cases leaves us with two options. The first is to resign ourselves to the idea that the judgments in question are inconsistent: a whole life cannot simultaneously be both very good and

very bad. The second, more attractive option is to consider the possibility that these judgments are not in conflict but have different objects. This is the line of inquiry pursued in this paper.

The Dual Theory, I believe, gives an elegant explanation of our evaluative ambivalence in Germain's case. One of our judgments has as its object her well-being, whereas the other concerns the realisation of her potential, and both of these are relevant to the overall goodness of her life.

To get a better grip on the strength of this explanation, let's consider some of its competitors. According to the first set of rival explanations, only one of the judgments comprising our evaluative ambivalence has to do with the prudential value of a life, whereas the other is about its *impersonal value*. There are two versions of this thought.

First, some thinkers maintain that certain things are valuable even if they are not good for anyone in particular: existence of beautiful artworks and landscapes, equality between members of the moral community, or the accumulation of knowledge. That is, in addition to prudential value, we should recognise aesthetic, egalitarian, or epistemic value.⁴ Now, one might suggest that when we judge Germain's life to be in one respect tragic, we are making a judgment about one of

⁴ This view has been popularised by Moore (1903).

these impersonal values. Perhaps a life which falls short of its potential makes for a worse story, and thus has low *aesthetic value*.

Second, problems of population ethics raised by Derek Parfit (1984) have led some authors to distinguish between the prudential value of a life and its *contributive value*. According to the critical level view, for example, a person's life contributes positively to the value of a state of affairs just in case its prudential value is sufficiently high, above a certain positive critical level.⁵ Consequently, a life which is 'worth living' might turn out to be not 'worth creating'. In a similar vein, one might suggest that facts about what could have happened bear just on the contributive value of a life. Perhaps we find Germain's life to be in one respect tragic because, although her life is worth living in virtue of all the good things that happened to her, the fact that she fell short of her potential somehow makes her life not worth creating.

None of these alternative explanations seems compelling to me. When I attend to the tragedy of Germain missing out on greater intellectual achievements and appreciation, her life does not strike me as particularly aesthetically disvaluable and I am not directly concerned with whether her life was worth creating. Instead, when I reflect on her life, I cannot help but think that it is, in one respect, going well

⁵ See Parfit (1984), Blackorby, Bossert, and Donaldson (2005), and Broome (2004).

for her and, in another respect, poorly *for her*. Thus, both of my judgments comprising evaluative ambivalence seem to concern the prudential value of a life.

The third rival explanation accepts this conclusion but insists that the goodness of a life is determined solely by one's well-being and that the established theories of well-being can wholly account for the phenomenon of evaluative ambivalence.

Admittedly, the established theories of well-being can explain *some* forms of evaluative ambivalence. Hedonism, for example, can make sense of a divided outlook on a life which involves an abundance of joy and substantial suffering. Preference views can account for evaluative ambivalence directed towards a life in which numerous preferences are satisfied while many others are frustrated. And the objective list view could explain an apparent conflict of judgments concerning a life which is deprived of close relationships but features extraordinary intellectual achievements.⁶

But all of these are instances of what we can call *shallow* evaluative ambivalence. The phenomenon that we have discovered earlier, *deep* evaluative ambivalence, is importantly different. For one thing, it seems that Germain's life was rich in all the recognised constituents of well-being. For another, her case engenders a distinctive phenomenology. Our evaluative ambivalence has to do with

⁶ Pummer (2017) calls such lives 'lopsided'. For concise overviews of these theories of well-being, see Gregory (2016), Heathwood (2016), and Fletcher (2016).

judgments about the entirety of her life. We don't feel the same way when a life is good at the beginning but bad at the end, or when it contains many pleasant experiences but also a considerable amount of suffering. Of course, calculating the net amount of well-being in such cases is not straightforward and this may be reflected in an assessment that is initially divided. But once this difficulty is behind us, there is no emotional and evaluative residue like that witnessed in Germain's case.

The last rival explanation is subtler. It holds that the judgments comprising evaluative ambivalence track not two different kinds of prudential value, but rather two different kinds of facts about the same prudential value: well-being. In particular, one judgment tracks a *non-comparative fact* about Germain's well-being (it was high), whereas the other tracks a *comparative fact* about her well-being (it was lower than in another possible scenario). This latter, comparative fact is what makes Germain's life in one respect tragic. But, crucially, it does not make her life worse.

I do not find this explanation satisfactory, for two reasons. To begin with, it cannot adequately capture our patterns of intuitions in cases which have a dynamic structure, such as the following.

Suppose that you are in a café and overhear a conversation at the nearby table. "How is your son doing these days?", one woman asks. "Same old", the other responds, "Billy doesn't leave the house much. He hardly ever talks to anyone and has no interest in reading—or anything, really." The uncomfortable silence that

follows is eventually interrupted. “But lately he seems to be developing some attachment to his brother. I’m curious to see where it goes.” If you are like me, you would probably think that Billy’s life is mediocre. But now suppose that the conversation progresses and you come to understand something important. Billy is not some rebellious teenager going through a difficult period, but rather a boy with a serious cognitive impairment. That newly established bond with his brother is probably as good as it could possibly be.⁷ Once this is in view, I believe, Billy’s life doesn’t seem quite as bad for him as it initially did.

Now, according to the presently considered rival explanation, our inclination to revise the initial assessment of Billy’s life is a mistake. After all, the second part of the conversation between the two women has not revealed any new non-comparative facts about his well-being. But that goes against our considered judgment. In cases like this one, whether a person realises their potential does seem to affect how good their life is for them.

The second shortcoming of this rival explanation is that it is fundamentally ill-equipped to help us make progress on the ethical puzzles addressed in the third part of this paper. As we will see, to solve these puzzles, we need something with axiological implications, whereas this rival explanation has none.

⁷ This and other disability-related examples in this paper concern severe *cognitive* disabilities. For a compelling argument that negative impacts of many *physical* disabilities on one’s well-being are greatly overstated, see Barnes (2016).

In sum, the Dual Theory appears to provide the best explanation of the phenomenon of evaluative ambivalence that we have encountered in the case of Sophie Germain and other similar scenarios.

3. The crib test

When philosophers seek to spell out the concept of prudential value and identify the ingredients of a good life, they often appeal to a thought experiment known as *the crib test*, which works as follows.⁸ Imagine that you are a proud parent of a newborn. As you look into the crib where the baby is peacefully sleeping, your heart fills with parental concern. You think to yourself, “I hope my baby gets to live a very good life”. The kinds of things that you, a rational and benevolent caretaker, wish for in this situation, for the baby’s sake, seem to be the ones which make a life worth living.

So what exactly do you wish for? Loving relationships, presumably. Achievements of various kinds. Pleasant experiences. These are the standard answers that people give and it might be tempting to stop there. But as you continue to think about that baby peacefully sleeping in the crib, I believe that you will eventually come to appreciate something important. What you care about is not just the number of relationships or achievements your child will have. You also care about the extent to which the child will realise its potential. To put it differently, the

⁸ See Adams (1999), Darwall (2002), Feldman (2006), and Bradley (2009).

realisation of one's potential appears to be a *distinctive object of prudential concern*. These kinds of things, according to the crib test, are among the basic determinants of the goodness of a life. This is one way in which the crib test supports the Dual Theory.

But there is another lesson as well: no quantity of achievements, loving relationships, and pleasant experiences seems sufficient for an individual to have an excellent life. It depends, in part, on one's potential. If one's potential is high, more of these goods are needed. If it's low, fewer will suffice. In other words, the goodness of a life seems *subject-relative*: a given life, in terms of what actually happens within it, can be better for one subject than for another.⁹

To illustrate this feature, return to the case of Sophie Germain. Suppose that she and her sister Angelique had lives which, on the whole, contained comparable pleasant experiences, loving relationships, achievements, and disappointments, and so had the same levels of well-being. There was, however, one salient difference between them: Angelique had no extraordinary talent for mathematics. Thus, the gap between how high Sophie's well-being was and how high it could have been was larger than in Angelique's case. In virtue of this feature, Sophie's life seems to me overall worse than Angelique's life, even though they had equal levels of well-being.

⁹ Subject-relativity should not be confused with the more widely discussed property of *agent-relativity*. A normative ethical view is agent-relative if, at least in some situations, different agents have different moral reasons for action despite having the same set of options. For canonical discussions, see Nagel (1970), Scheffler (1982), and Parfit (1984).

The Dual Theory predicts this kind of subject-relativity. Because this view recognises two basic determinants of the goodness of a life, well-being and the realisation of one's potential, it implies that individuals who are tied in terms of well-being can differ in terms of the overall value of a life.

Notably, the Dual Theory also implies that it would be better for a person to have lower potential, holding fixed the way their life actually goes. Some people may initially find this corollary implausible. But I think that we should accept it upon reflection.

To begin with, let me offer a debunking explanation for the contrary intuition. There are two main mistakes that people tend to make when thinking about this issue.

The first mistake is to assume that a person could be benefited by, say, having their fabulous musical gifts removed after they are born, if they would never exercise them anyway. This would indeed be a strange implication. But the Dual Theory does not say that. On this view, it would be better for a person to be born without fabulous musical gifts, if they would never exercise them anyway. But having one's gifts removed is not like never having them in the first place. Instead, it's like being banned from playing music or suffering a career-ending accident: it does not reduce one's potential, but merely determines which of the many possible paths their life will actually take. This is how we need to think about the case of Sophie Germain as well: the sexist actions of her contemporaries did not alter what

she was truly capable of, but merely determined which of the many possible paths her life took, and in that way affected her well-being.

But even once this is clarified, it's easy to make a second mistake: namely, to fail to hold fixed how one's life actually goes across the compared cases. There is, after all, a close connection between the potential and the actual. For instance, if someone were to lack the kind of emotional sensitivity needed to engage in loving relationships, that would typically impact not just what bonds they could have with other people, but also what bonds they would actually have. The same goes for talents. There are normally plenty of occasions for a person to put their athletic, musical, or scientific gifts to use—even if that person does not end up pursuing a career in any of these fields—which would often make a positive contribution to their well-being. We may, therefore, implicitly associate lower potential with lower well-being even when we are explicitly asked to assume that away.

In addition to these debunking explanations, there is a positive reason to embrace the claim that it can be in one respect better for a person to have lower potential. In particular, this claim provides a natural explanation for a range of attitudes that we express towards certain kinds of news. Take the case of Michael Jordan, who retired from professional basketball in 1998 after winning a third consecutive, and sixth overall, NBA title with the Chicago Bulls. In the recent television documentary, *The Last Dance*, Jordan reflects on his spectacular career and the controversial circumstances of his decision to leave the sport. About the latter,

he remarks: “It’s maddening because I felt like we could have won seven”.¹⁰ In light of this, it seems to me that Jordan would be *glad* to learn that he was mistaken and, in fact, could have never won the seventh NBA championship, perhaps because his body could not have handled another demanding season. In a similar way, many people are *relieved* to hear that their loved one, who had been missing for years, has been discovered to be dead. These attitudes seem fitting, and a natural explanation for why they are is precisely that having lower potential is, in one respect, better for us. After all, a key feature which these cases share is that something that had seemed attainable—an extraordinary achievement or being reunited with a loved one—turns out not to have been in the space of possibilities after all.

With these remarks in place, let’s return to the main line of the argument. The previous section provided initial support for the view that there are at least two basic prudential values. The upshot of this section—that the value of a life is subject-relative—strengthens that case. For if two lives can have the same level of well-being and yet differ in terms of how they are for the people living them, then there must be a further determinant of the life’s goodness. And that determinant appears to be the degree to which an individual realises their potential.

¹⁰ *The Last Dance*, Episode 10.

What remains to be established is what counts as one's potential, and how well-being and the realisation of one's potential combine to determine the overall goodness of a life. This is what the next two sections set out to do.

4. What counts as an individual's potential?

According to the Dual Theory, the goodness of a life is determined in part by the degree to which one realises their potential. But what exactly counts as one's potential? There are several theoretic choice points concerning this issue.

The first thing that we should consider is whether one's potential is determined by facts about them as an individual, or rather by facts about some reference class to which one belongs, such as their species. In the latter vein, one might propose that one's potential is the maximal level of well-being that members of their species can attain.

However, this proposal is not plausible. If there was just one evaluatively significant level of potential for all humans, then many of the judgments we made earlier would not make sense. For example, there would be nothing particularly heartening about Billy's rudimentary relationship with his brother, and no room for the claim that his life is 'fortunate' either. Likewise, in Michael Jordan's case, we would have to accept either that there was nothing tragic about him missing out on that seventh NBA title, or else that it is tragic for people like you and me that we

have failed to win that many accolades. These considerations suggest that we should instead think of one's potential as determined by facts about that specific individual.

This brings us to the second choice point: which facts about the individual are relevant? Is it the level of well-being they could have easily had otherwise? Or perhaps their maximal possible well-being?

The first of these alternatives might seem attractive. While the significance of modal facts has been largely overlooked in the literature on the goodness of a life, it plays an important role in epistemology. Many authors think that the concept of knowledge includes an anti-luck condition.¹¹ Specifically, an individual fails to know that some proposition is true if they could have easily believed otherwise. Something similar could be said in the present context: the extent to which one realises their potential is determined by the relation between their actual well-being and the level of well-being they could have easily had otherwise.

However, as the case of Sophie Germain illustrates, this cannot be right. The prejudice against women, and female scientists in particular, was so entrenched in Germain's society that she couldn't have easily achieved more, received greater recognition, and avoided upsetting experiences borne of the obstacles she faced. Despite this, we intuitively judge her potential well-being to be *much* higher than her

¹¹ See, for example, Sosa (1999), Williamson (2000), and Pritchard (2005).

actual well-being. In light of this, and in the absence of other natural candidates, we should tentatively accept that one's maximal possible well-being is what matters.

Turn now to the third question: is the relevant concept of potential subjective or objective? Michael Jordan certainly believes that he could have won that seventh NBA title. And many of us have experienced that one romantic relationship that we had thought would last forever but which did not work out. Such instances of unfulfilled potential sting the most. It might therefore seem that our evaluations should be sensitive to one's subjective potential—in particular, the maximal possible well-being one believes to have.

Sophie Germain's case once again provides a counterexample. While Germain probably believed that her scientific achievements could have been greater, it is unlikely that she appreciated just how much larger they could have been. Through no fault of her own, of course: she is just one of the millions of women in history who have been told that they are not capable of excellence. It is preposterous to think that these are not genuine stories of wasted potential. Thus, the relevant concept of maximal possible well-being must be the objective one.

Our last question concerns the very notion of *possibility*. This question is as important as it is difficult to answer. To see that, return to the case of Billy. On the one hand, because Billy is severely cognitively disabled, it seems that his life *could not have been* much better than it actually is. On the other hand, insofar as Billy could

have been born without that disability, it seems that his life *could have been* much better.

To separate such cases, we may refer to the distinction between *intrinsic* and *extrinsic* potential. The former is the potential whose attainment does not require any direct or significant alteration of the individual's constitution, whereas the latter notion is more encompassing.¹² Thus understood, Billy's intrinsic potential is not much higher than his actual well-being, whereas his extrinsic potential is considerably higher. Of these two notions, intrinsic potential appears to me to be closer to what matters. After all, there *is* something heartening about Billy's life. I will assume this view for the remainder of the discussion.¹³

Let's stop here for the time being. The preceding discussion suggests that one's potential is determined by facts about one's individual (rather than species), objective (rather than subjective), and intrinsic (rather than extrinsic) maximal possible well-being. There are, of course, further questions that we could ask about this concept, but the absence of a fully-fledged theory should not deter us from

¹² For a discussion of the limitations of this distinction, see McMahan (2002) and Vallentyne (2005).

¹³ Alternatively, we could lean into the difficulties concerning what counts as one's potential and propose that this concept is context-sensitive. On this alternative view, whether Billy's intrinsic or extrinsic potential is relevant depends on the context in which we are assessing the goodness of his life. This would be an instance of a broader phenomenon of the context-sensitivity of counterfactuals. (See Lewis 1986 for a canonical discussion.). This alternative view is not pursued here any further, but it is worth noting one of its implications. If what counts as one's potential were context-sensitive, and the Dual Theory were true, then the overall goodness of a life would also be context-sensitive—and that would have significant upshots for normative ethics.

pressing on with our inquiry. The choice of what exactly counts as one's potential will sometimes affect the assessment of how good a particular life is, but it should not affect the substance of the claims defended in the remainder of this paper, as long we hold it fixed across the relevant comparisons.

5. Combining values

The Dual Theory recognises that there are two basic prudential values: well-being and the realisation of one's potential. But how exactly do they combine to determine the overall value of a life? I would like to restrict the present discussion to two simple models which represent two different ways of conceptualising the significance of wasted potential.

According to *the Addition Model*, to determine the overall value of a life, we need to add the value of realised potential to the contribution made by well-being, where the former is represented by a non-negative number. This is a 'glass half-full' view: one gets bonus points for realising one's potential, and more points for realising it more fully.

By contrast, *the Subtraction Model* determines the overall value of a life by subtracting the disvalue of unrealised potential from the value of well-being, where the former is represented by a non-negative number. This is a 'glass half-empty' view: one loses points for not realising one's potential fully, and that penalty is greater the bigger the potential gap.

Considering that addition and subtraction are inverse mathematical operations, these two views will have the same implications in a wide range of cases. However, in virtue of how they conceptualise the significance of realising one's potential, these models can diverge in their assessment of lives with barely positive levels of well-being and lives with barely negative levels of well-being.

To see that, first consider a life which contains no positive ingredients of well-being (pleasures, loving relationships, achievements, etc.) and a small amount of negative ingredients of well-being (intermittent pain, for example). The Addition Model allows for such a life to be overall worth living, as long as one's maximal possible well-being is not much higher, and so one's potential is realised to a large extent. This strikes me as deeply counterintuitive. By contrast, according to the Subtraction Model, a life with a negative well-being level is never overall worth living.

On the other hand, the Subtraction Model implies that unless we make *enough* of our talents and opportunities, our lives will fall short of being overall worth living. Thus, a life with a positive level of well-being could be not worth living overall. I find this reverse implication more plausible.

To begin with, consider a human living the life of a domesticated mouse which, let's assume, is short, monotonous, and deprived of sophisticated pleasures, achievements, and deep relationships. Intuitively, although this life would have a

positive level of well-being, it does not seem to be worth living for a typical human whose potential is high.¹⁴

Moreover, unless we accept something like the Subtraction Model, we cannot make sense of our evaluative attitudes towards smaller-scale events, such as the following.

Marathon: Eve is capable of qualifying for the Olympic Games in the marathon but, instead of attending the Olympic Trials, she participates in a low-key local race and wins it by being only slightly faster than the amateur second-place finisher.

Whatever contribution this accomplishment makes to Eve's well-being, it is overshadowed by the badness of failing to realise her athletic potential in that part of her life. Crucially, that is not to say that it would have been better for Eve not to participate in that amateur race at all. After all, the gap between no achievement and a great achievement is larger than the gap between a small achievement and a great achievement. But, it seems to me, such barely good performances are not enough

¹⁴ McMahan (1996, p. 9) expresses a similar sentiment in the process of arguing that 'fortune', not well-being, should be the focus of principles of justice: "A dog, for example, has a relatively low level of well-being. But while a normal human adult with a comparable level of well-being would be very badly off, the dog may well be flourishing".

to render a life with extraordinary potential overall worth living. These considerations lead me to accept the Subtraction Model.

Admittedly, some people do not share my intuitions about these two cases. They find too extreme the idea that a life with a positive level of well-being can be overall not worth living, even if it realises its potential to a very low extent. To avoid this implication, they could opt for a more sophisticated, asymptotic version of the Subtraction Model, on which unrealised potential always makes a life go worse, but it never makes it overall not worth living.¹⁵ This alternative version of the Dual Theory would be in one way stronger: it would be easier to accept for most people. In another way, however, it would be weaker: some of its explanatory power would be lost. In particular, the Dual Theory would no longer deliver a satisfactory solution to the puzzle discussed in the next section, concerning standards for a life worth living in humans and animals.¹⁶ When weighing these theoretical virtues, I find myself partial to greater explanatory power and thus prefer the simple Subtraction Model. But the reader is free to choose otherwise. All other claims would still follow.

6. Humans, animals, and lives worth living

The preceding discussion provided the initial motivation for the Dual Theory and examined some of the ways in which it can be spelled out. What I want to do in the

¹⁵ For example: Value of a Life = Well-being²/Potential.

¹⁶ I explain these differences in the next section, in footnote 21.

remainder of the paper is to put the Dual Theory to the test and show that it illuminates three ethical puzzles.

In this section, I argue that the orthodox view which holds that well-being is the sole determinant of the goodness of a life faces a *trilemma* regarding standards for a life worth living for humans and non-human animals.¹⁷ We can avoid this problem, however, if we accept the Dual Theory.

The trilemma arises when those who accept the orthodox view are confronted with the following, simple question: What counts as a life worth living for humans and animals?

The first option available to them is to postulate that there is a single threshold for a life worth living for humans and animals, and to set this threshold at a ‘relatively low’ level of well-being. This view captures the intuitive thought that at least some animals have good lives: your friend’s Labrador, perhaps, or the cows grazing peacefully in the Swiss Alps. Notice, however, that this view also implies that the life of a domesticated mouse—short, monotonous, and deprived of sophisticated pleasures, significant achievements, and deep relationships—would be

¹⁷ A background assumption needed to generate this trilemma is *welfare invariabilism*, which is the view that the same theory of well-being is true of every subject capable of well-being. That is, the same list of basic goods (such as pleasures, achievements, friendships) applies to all humans and non-human animals who are capable of welfare, even if some of these goods are in fact inaccessible to them. For a compelling defence of this view, see Lin (2018).

worth living for a typical human whose potential is high. As noted earlier, I find this difficult to accept.¹⁸

The second option is to uphold the commitment to a single threshold but to set it at a ‘relatively high’ level of well-being instead. This avoids the implication that a mouse-like life would be worth living for a typical human. However, this obviously comes at the expense of capturing the first intuition that some animals have good lives. This may not seem like a high cost at first. But if we also grant the claim that there is a moral reason to prevent individuals who would have lives not worth living from coming into existence, then it follows that we have a moral reason to bring about the extinction of many, if not all, animal species.¹⁹ This, too, is deeply counterintuitive.

The remaining alternative is to propose that there are two standards for a life worth living: one for humans, at a ‘relatively high’ level of well-being, and another for animals, at a ‘relatively low’ level of well-being.

But this option is also unsatisfactory, for at least a couple of reasons. For one thing, it is speciesist: it discriminates among individuals on the basis of their

¹⁸ Two clarifications. First, a mouse-like life need not literally be the life of a mouse. It could be a life that is equivalent to that of a mouse in terms of the level of well-being. Second, such a life does not obviously seem not worth living for an atypical human whose potential is low (perhaps in virtue of a severe cognitive disability.) The argument presupposes only that it is not worth living for a typical human whose potential is high.

¹⁹ This claim is one part of the so-called *Procreation Asymmetry*. See McMahan (1981) and Frick (2020).

species membership *alone*.²⁰ Two individuals who have the same level of well-being should not differ with respect to whether they have lives worth living just because they belong to different species.

For another, this last answer has some unpalatable implications for distributive ethics, at least if we think that animals are within the scope of egalitarian concern. For the sake of precision, let's say that well-being can be represented on a ratio scale, and that the threshold for a life worth living is at 0 units of well-being for humans, and at -10 units of well-being for animals. Suppose that there is a human, Ada, whose well-being is at -1 and that there is a dog, Fido, whose well-being is at -9. In this case, we would have an egalitarian reason to redistribute well-being from Ada to Fido. But this means that we would have a reason to redistribute well-being from an individual who has a life not worth living to an individual who has a life worth living. That's not plausible.

We can avoid this trilemma if we recognise that equating the goodness of a life with the level of well-being is a mistake. A mouse might have the same level of well-being as a typical human living a mouse-like life, but these two individuals are likely to differ in terms of the degree to which they realise their potential. The mouse's life could not have been much better than it is, whereas the human's life could have been. This allows us to explain how, despite having the same levels of

²⁰ This term was popularised by Singer (1975).

well-being, a mouse could have a life worth living, and a human could have a life not worth living. Thus, the Dual Theory delivers a compelling response to our trilemma. It allows us to have a single threshold for a life worth living for all beings. At the same time, it avoids implying that no animals have lives worth living or that a mouse-like life would be worth living for a typical human.²¹

In closing, let's briefly address two issues. First, why should we think that animals typically realise their potential to a higher degree than humans? Humans are generally equipped with greater and more numerous capacities than non-human animals, and therefore can access a wider range of goods. For instance, humans are capable of valuing their bonds with others and experiencing a broader array of emotions regarding these bonds, which allows them to participate in highly intimate relationships. Likewise, humans seem to be uniquely equipped with the awareness of what is worth pursuing and the grit to succeed in doing so, which are essential ingredients for the greatest achievements.

Thanks to these capacities, most humans can have truly extraordinary lives. We can get an idea of what is possible by reflecting on our lives' best moments and

²¹ When paired with the asymptotic version of the Subtraction Model, the Dual Theory would not avoid the presented trilemma. It would imply that a given life is worth living for a member of some species, regardless of their potential, if and only if it is worth living for members of all species. Thus, if we determined that a particular life is worth living for an ant or a mouse, we would have to accept that it's worth living for a human as well. And if we determined that a particular life is not worth living for a human, then we would have to accept that it is not worth living for an ant or a mouse either, and thus that we have reasons to prevent such animals from coming into existence.

imagining that our entire lives were like that. Our maximal possible well-being is indeed very high. But, as the experiments in living of billions of our ancestors and contemporaries teach us, it is extremely difficult to flourish in all respects. Thus, human lives which brush against their maximal possible well-being are exceedingly rare. By contrast, for non-human animals, whose capacities are much more limited, the task of reaching one's potential is easier. We should therefore expect animals to realise their potential to a greater degree than humans.²²

Second, isn't there a rival non-orthodox solution to the trilemma identified above? On the Dual Theory, the realisation of one's potential and well-being jointly determine the goodness of a life. But one could propose a more radical view, on which the realisation of one's potential is the sole determinant of the life's value. Both of these views avoid the trilemma identified above.

²² *Objection:* The Dual Theory implies that many people have very unequal lives because some are undiscovered geniuses whereas others are just ordinary, and such hidden inequalities seem bizarre. *Response:* First, this theory does not imply that. For one thing, it is plausible to assume that people's levels of potential follow a normal distribution, so geniuses are rare. For another, as I explain in Section 3, people with higher potential will generally have higher well-being and those with lower potential will have lower well-being, so people generally realise their potential to a similar extent. Thus, the Dual Theory will not have very different distributive implications compared to the orthodox view on which well-being is the sole determinant of a life's value. Second, even if the Dual Theory had this implication, it would not be bizarre. Any theory of well-being similarly holds that its rivals are making a systematic error in assessing people's lives. If it were true that most people experience similar amounts of pleasure and pain, but differ significantly in terms of achievements, relationships, and appreciation, then hedonism would imply that people have mostly equal lives whereas the objective list theory would predict significant inequalities.

However, the radical view does that at a high price. First, in virtue of its monistic character, the radical view cannot account for the phenomenon of evaluative ambivalence identified earlier. Second, this view judges animal lives to be implausibly good: for instance, it implies that the life of a fortunate mouse (low well-being, low potential) is just as good as the life of a fortunate human (high well-being, high potential). Third, the radical view has some unpalatable implications for distributive matters: for example, it implies that we should benefit Aisha rather than Bertrand if Aisha has a much higher level of well-being than Bertrand, but realises her potential only to a slightly lesser degree than him. The Dual Theory does not run into these problems.

7. The significance of a life's shape

The Dual Theory also helps us resolve a puzzle concerning the following question: What is the relation between how well one's life is going at any particular time and how well it goes overall?

The following, simple answer might initially seem attractive. To determine how well a life goes overall, we can simply assess how well it is going in each period and then put these values together. Moreover, the temporal features of a life, such as the order of events, do not affect the prudential value of a life. Of course, a vigorous game of tennis followed by an indulgent lunch makes for a more pleasant

day than the converse. But, the thought goes, it makes no difference which days are happy and which are miserable, as long as their number is the same.

However, this simple view is challenged by the following pair of cases.

Improvement. Your life begins in the depths and takes an upward trend: misery in the early years, mixed fortune in midlife, followed by flourishing in old age.

Deterioration. Your life begins at the heights and takes a downward trend: an early period of flourishing, mixed fortune in midlife, followed by misery in old age.

These lives are permutations of each other—in the sense that, for every temporal segment with some amount of well-being in one life, there is exactly one such segment in the other life, and so on—so the simple view regards them as equally good. However, many people have a strong intuition that the improving life is better than the deteriorating life.²³ The challenge is to find a compelling theoretical rationale for this judgment.

²³ This intuition is shared by Slote (1983), Bigelow et al. (1990), Velleman (1991), Kamm (2003), Portmore (2007), Temkin (2012), Kauppinen (2012), Glasgow (2013), Dorsey (2015), and Hirose (2015).

Although this puzzle has generated a substantial literature, it will be instructive to start with my positive proposal and only then turn to the shortcomings of the existing views. In essence, I argue that an improving life typically features a greater degree of potential-realisation than a deteriorating life, and in virtue of this, it is overall better even if the total amounts of well-being are the same. This explains why we are inclined to judge that the improving life is better in the original, abstract case.

To illustrate this, consider two ways in which a philosophical career could turn out.

Early Breakthrough: Adam writes his most original and insightful article shortly after finishing his graduate studies. His next publications continue to make valuable contributions to philosophical debates, but their quality gradually diminishes.

Late Breakthrough: Adam begins his career with a run-of-the-mill article, but the originality and insight of his work steadily improve.

These two lives, let's assume, are otherwise identical.²⁴ The second life seems better than the first. Why?

Take Early Breakthrough. There is just something tragic about a situation in which the greatest achievements take place early on in Adam's life and are not replicated further down the road. A natural explanation of this tragedy draws on the same kinds of considerations that we have been discussing throughout: the significance of realising one's potential. The fact that Adam writes a terrific essay early on in his life reveals that he is already capable of producing great works of philosophy at that stage. His *momentary potential*, as it were, is already high. When his next paper is a lesser achievement, this amounts to more than just a deterioration in terms of *momentary well-being*. In addition, the gap between what Adam could accomplish and what he actually accomplishes grows, and with it grows the gap between his well-being and his potential.

Consider Late Breakthrough now. The fact that Adam's career starts with a minor paper suggests that he does not yet have what it takes to write a brilliant article. Consequently, the gap between what he achieves and what he could achieve at that time is small. Moreover, it is natural to assume that the subsequent

²⁴ Of course, this involves a high degree of idealization. Discovering that your work is getting worse would be a source of distress for many people. But we need to assume that away or suppose that Adam just does not give any thought to this matter.

improvement in the quality of Adam's work mirrors the increase in his momentary potential.

Thus, although these two lives have the same levels of *lifetime well-being*, they differ in terms of the level of *lifetime realisation of one's potential*. It is greater in Late Breakthrough than in Early Breakthrough. This explains why the former life seems better than the latter.

An analogous gloss can be given for a comparison between two lives which differ in terms of when one's most prized relationships occur.

Lost Love. In his youth, Alex has a number of highly fulfilling relationships. Over the years, however, these attachments crumble and are replaced by other, more superficial connections.

Found Love. Alex's youthful relationships are relatively superficial. Over time, however, he makes new, highly fulfilling connections.

In Lost Love, the fact that Alex has highly fulfilling relationships early in his life reveals that he already has the kinds of emotional sensitivity, sense of respect, and ability to trust another that are needed to engage in and sustain the most valuable relationships. When his next attachments are of lesser quality, the gap between how high his well-being is at any given time and how high it could be at that time starts

to grow. By contrast, in Found Love, the relationships that Alex actually has at any given time presumably correspond to the relationships that he could have at that time. Thus, these two lives differ in terms of the level of lifetime realisation of one's potential. This explains why Lost Love strikes us as worse than Found Love.

Crucially, this account does not imply that a life which improves in terms of well-being is *always* better than a life which features a corresponding deterioration. There are at least two kinds of exceptions.

First, note that not all cases of early breakthrough strike us as tragic. Mathematicians, for example, tend to make the most substantial contributions to their field at a relatively young age. Indeed, the highest honour a mathematician can receive, the Fields Medal, has an age limit (the recipient must be under forty), but this is hardly a constraint in selecting the winner. Similarly, most athletic careers end in one's thirties. These life stories do not seem as upsetting as that of a philosopher who fails to make a substantial scholarly contribution in the second half of their life.

Why? Evidently because of the difference in terms of what the subjects are capable of at any given point in their life. Since the abilities of mathematicians and athletes typically peak in their late twenties or early thirties and then deteriorate, they do not fail to realise their potential if they fail to replicate the achievements of their youth later in life. Arguably, philosophers are not like that: insofar as there is a peak, it typically comes later in life, and thus a story of an early breakthrough is bound to be more upsetting.

Second, there are cases in which the contrasting patterns of well-being are attributable to changes in one's pleasant experiences rather than achievements or relationships. Dale Dorsey (2015, p. 139) considers one such scenario.

[C]ompare two different experiments. The first person is hooked up to a pleasure-stimulating computer with software designed to start his life at a neutral level: no pleasure, no pain. Gradually, say, twice a year, the pleasure is increased in a linear fashion. The second is precisely the opposite: the person's machine starts out with quite a lot of pleasure. Gradually, also twice per year, the software decreases the pleasure in a linear fashion, such that both people, over the course of their lives, will generate the same amount of sensory pleasure.

Intuitively, neither of these two lives seems better than the other. My account accommodates this verdict. In the situation imagined by Dorsey, there is no reason to suppose that there is a difference in terms of the subjects' realisation of their potential. Thus, the Dual Theory implies that these lives are equally good.

To summarise, here is how we should think about the puzzle at hand. The shape of a life indicates (but does not guarantee) the presence of something else that affects the goodness of a life: whether one's potential has been realised to a high extent. It is in virtue of differences in this respect that an improving life is *typically*

better than a deteriorating life. This explains why we are inclined to judge that an improving life is better than a deteriorating life in the original case. Call this *the Potential Account*.²⁵

This solution avoids the shortcomings of three popular accounts offered in the literature. According to Michael Slote's (1983) *Temporal Location Account*, an upward-sloping life is better than a downward-sloping life because what happens later in life has greater weight in determining the life's overall value. That is, any given good (say, an achievement or a pleasant experience) contributes more to the goodness of a life if it is present in one's adulthood or old age rather than in one's youth.

The Pattern Account offers a different take. It regards improvements in terms of temporal well-being as in themselves good for a person and corresponding deteriorations as in themselves bad for them. Frances Kamm (1998; 2003), Larry Temkin (2012), Joshua Glasgow (2013), and Iwao Hirose (2015) are all in this camp.

²⁵ What about the following case? Life A features an early breakthrough but then decreases over time in terms of both well-being and potential. Life B features a poor start, a gradual increase in terms of well-being, and constantly high potential. Assuming that the total amount of well-being is the same in both lives, the Dual Theory judges A to be better than B because the latter life does a poor job of realising its potential in the early stages. This may seem counterintuitive. On reflection, however, this seems to be the correct verdict. Life A is the paradigmatic story of an athlete who makes the most of their talents and opportunities at every stage of their entire life. Life B, on the other hand, is a story of undiscovered genius and wasted time. It would be, perhaps, the life of Albert Einstein if he had spent not two years but rather two decades as a low-paid clerk poring over patent applications for gravel sorters and typewriters before starting to disseminate his ground-breaking ideas in physics.

Finally, according to *the Narrative Account*, an upward-sloping life is better than a downward-sloping life when and because it features certain narrative relations between discrete events in one's life, such as when one suffers a misfortune and then learns from it, or when an effort is turned into a success rather than a waste. David Velleman (1991) has defended this view, and its variants have been endorsed by Douglas Portmore (2007), Antti Kauppinen (2012), and Dale Dorsey (2015).

An undeniable attraction of the first two views is their simplicity. But it is also their greatest flaw: they *prove too much*. In particular, they predict that an improving life is always better than a deteriorating life. But, as we have seen, this is not the right verdict.

The Narrative Account is closer to the truth in that it recognises that an improving life is only typically better than a deteriorating life. But Velleman's stated account *proves too little*. It cannot capture the relevant intuitions in some central cases, such as the comparison between Found and Lost Love, and between Late and Early Breakthrough. In these scenarios, one of the compared lives seems better than the other, even if we stipulate that they do not differ in terms of whether the agent has learned from their past mistakes or turned their past efforts into an eventual success. Moreover, it is far from obvious how the Narrative Account could be extended to cover such cases because Velleman does not offer a unified account of the significance of the relations between discrete events that he discusses. This is in stark contrast to the Potential Account, which explains our intuitions about the relevant

cases in a theoretically unified way. On this view, an improving life is better than a deteriorating life when and because these alternatives differ in terms of the degree to which one's potential is realised.

8. Badness of death

The Dual Theory can also illuminate a long-standing ethical puzzle concerning the badness of death. Many people share the belief that death is typically bad for its subject, and one of the things we want to know is why and to what extent this is so.

Nagel (1979) and others who have addressed this question quickly rejected the implausible views that death is bad for us simply because being dead or dying are bad for us. This left them with the idea that death is bad because it *deprives* us of future goods. One popular version of this view is known as *the Life Comparative Account*.²⁶ It says that to determine the extent to which death is bad for someone who dies in an accident, we should compare the amount of well-being they actually had with the amount of well-being they would have had, were they not to die in the accident.

One thing that we can notice straight away is that the Dual Theory dovetails with the Life Comparative Account. For instance, both views capture the profound sense of tragedy that accompanies the death of a teenager who 'had a whole life

²⁶ This label is due to McMahan (2002, p. 105). See also Feldman (1991), Broome (1993), and Bradley (2009).

ahead of him'. And both views recognise that, in some cases, an earlier death could be a blessing.

However, the relation between these two accounts is deeper than that: one explains the other. According to the Dual Theory, the badness of death is just one instance of a more general phenomenon, the badness of unrealised potential. To put it in a slogan form, being run over by a bus is like failing to catch a bus. It makes us miss out on various opportunities presented to us by life. Of course, the tragedy of a premature death is typically vastly greater than the setback caused by failing to catch a bus. But that's because the former makes us miss out on *all* opportunities, not just some, and thus leaves a greater chunk of our potential unrealised.

9. Concluding remarks

I have proposed that how good a life is for someone is determined jointly by their level of well-being and the degree to which they realise their potential, where the latter is understood in terms of how close one is to their maximal possible well-being. In particular, I have claimed that this Dual Theory provides the best explanation of the phenomena of evaluative ambivalence and subject-relativity, and that it illuminates ethical puzzles concerning standards for a life worth living for animals, the significance of a life's shape, and the badness of death.

Of course, there is more work to be done if the Dual Theory is to supplant the orthodox view on which the goodness of a life depends exclusively on what

actually happened within it. The arguments of this paper notwithstanding, the Dual Theory remains an extraordinary claim, and such claims require extraordinary evidence. Further analysis may also show that some of the details of this view—for example, concerning what counts as one’s potential and how to combine well-being and the realisation of one’s potential—should be spelled out in alternative ways. But I hope that my discussion has provided at least a blueprint for this larger project and motivation for pursuing it.

In closing, I would like to briefly address two worries which are often raised about this account. They both concern how we should direct our lives, given what we know and don’t know about our potential.

The first issue is this. According to the Dual Theory, prudence demands that we make the most of our talents and capabilities. Indeed, as we saw in Section 5, unless we make enough of our talents and capabilities, our lives will fall short of being worth living. But we also know that, for example, athletes and musicians tend to make tremendous sacrifices in their lives: often, if you are ten years old and not already practising something, you have little chance of world-class success. Moreover, if you do not devote dozens of hours every week to your discipline, you are unlikely to get very far. To make matters worse, only a tiny proportion of those who put in the work will make it to the main stage. Now, we certainly cheer on people who choose to take this path, but a theory of prudence which condemns people to it may seem objectionably demanding.

However, the upshot of the Dual Theory is more nuanced. While my discussion has centred on athletic and intellectual achievements, it is important to keep in mind that any plausible theory of well-being recognises a range of other prudential goods: relationships, pleasures, and knowledge, among others. In virtue of this, we have to be sensitive not only to whether an individual realises their potential for achievements, but also to whether they realise their potential for these other goods. At one extreme, if running or doing philosophy are truly the only things that make you happy and you show great promise in these disciplines, prudence may well demand that you go all-in. On the other hand, if you are just above average and these pursuits would eat up all of your time, then it may be better to choose a safer career and live a more balanced life.

One might also worry that as we expand our understanding of the goodness of a life, it becomes less transparent what is in our interest at any given time and how our lives are going as a whole. This is because now we need to know not just facts about the actual world, but also about various possible worlds. And this, in turn, might be taken to have a paralysing effect on the ability of our theory of prudence to inform our lives.

This concern is overstated. Even the orthodox view, on which the goodness of a life is determined solely by the level of well-being, gives us at most rough guidance regarding matters of self-interest. For example, we would all be hard-pressed to report our lifetime hedonic scores as of this morning. Likewise, the

primary ambition of the Dual Theory is not to guide us through every single decision, but rather to help us better understand what makes life good. I believe it is successful in this regard.

.

Part II: Value and Time

Chapter 2: The Shape of History

1. Introduction

Questions about the condition of humanity and its trajectory are deeply rooted in the Western intellectual tradition. Some philosophers believe in progress: they think that our current situation is better than that of our ancestors and that future generations will fare even better. For example, Immanuel Kant asserts in “Theory and Practice” that

[T]he human race is continually advancing in civilization and culture as its natural purpose, so it is continually making progress for the better in relation to the moral end of its existence, and ... this progress although it may be sometimes interrupted, will never be entirely broken off or stopped.¹

Others believe in decline: they argue that the human condition has deteriorated and will continue to do so. Famously, the archaic poet Hesiod describes in *Works and Days* a ‘golden race’ that populated the earth before us:

¹ Similar sentiments are expressed in the writings of Lucretius, Seneca, Augustine, Newton, Boyle, Comte, Hegel, Darwin, Marx, Spencer, and many others. For helpful comparative discussions of these views, see Bury (1921), Arendt (1963), Nisbet (1980), and Lange (2011).

[T]hey lived like gods without sorrow of heart, remote and free from toil and grief: miserable age rested not on them; but with legs and arms never failing they made merry with feasting beyond the reach of all evils. When they died, it was as though they were overcome with sleep, and they had all good things.²

This disagreement concerns a largely *descriptive* question: Does the history of humanity feature a pattern of improvement or deterioration?³ What has been neglected, however, is a purely *normative* question: Is it better if the history of humanity features a pattern of improvement rather than deterioration?

To get a grip on this latter question, consider two ways in which the condition of humanity could change over time.

Progress. Our history begins in the depths but has an upward trend: hardship in the early years, mixed fortune in the middle period, followed by flourishing at the end of our time.

² This 'pessimistic' point of view can also be found in the work of Tocqueville, Burckhardt, Nietzsche, Schopenhauer, and Weber, among others.

³ Any talk of improvement must involve some evaluative standard, so this question is not *entirely* descriptive.

Decline. Our history begins at the heights but has a downward trend: an early period of flourishing, mixed fortune in the middle period, followed by hardship at the end of our time.

It is important to note two things about these scenarios. First, in describing these cases, I have used vague expressions like ‘hardship’ and ‘flourishing’ because the puzzle I am about to describe arises for a wide range of theories of value. For the sake of concreteness, however, I will assume a pluralistic theory on which the goodness of a state of affairs is determined by aggregate well-being and impersonal values such as equality, accumulation of knowledge, the existence of beautiful artworks, and biodiversity, among others.⁴ We should therefore think of an upward trend in *Progress* as corresponding to an improvement in one or more of these values, and conversely for *Decline*.

Second, a few words about time. My description of *Progress* and *Decline* assumes that the history of humanity has the same length in either case and divides history into three periods. The latter assumption is for simplicity only. All that matters is that humanity’s history features a pattern of improvement or deterioration. These changes could be from one generation to the next or from one century to the next.

⁴ This view was popularised by Moore (1903). For an extensive treatment see Temkin (2012).

With these remarks in place, let's now compare these two scenarios. When confronted with this choice, many people have the intuition that *Progress* is better than *Decline*. On that view, for example, if a god was deciding which of these two worlds to create, they should create the first one.

But this intuitive judgment is puzzling. Notice that these two scenarios are permutations of each other—in the sense that for every period of flourishing in one course of history, there is exactly one such period in the other, and so on. In fact, we could even imagine that exactly the same people would exist in both scenarios and that they would have exactly the same levels of well-being, just at different times. This, on one kind of view, strongly suggests that *Progress* and *Decline* are equally good.

My aim in this paper is to resolve this puzzle by developing a theoretical account which vindicates the intuition that *Progress* is better than *Decline*. My discussion is structured as follows. In Section 2, I argue that we cannot solve the new puzzle about the shape of history simply by extending one of the popular answers to a related puzzle about the shape of a life. In Section 3, I develop a novel solution rooted in a broadly conservative view about value. My central claim is that the shape of history matters just in case and because it has something to do with the conservation of certain valuable elements of our culture across time. Then, in Section 4, I defend the proposed account from three objections. Finally, in Section 5, I explain the theoretical and practical significance of this account for matters related to the future of humanity.

2. Three candidate accounts

Although the puzzle about the shape of history has not yet been discussed in the literature, a related puzzle about the significance of a life's shape has received considerable philosophical attention.

This latter puzzle runs as follows. Consider two ways in which your life might go.

Improving Life. Your life begins in the depths but has an upward trend: misery in the early years, mixed fortune in midlife, followed by flourishing in old age.

Deteriorating Life. Your life begins at the heights but has a downward trend: an early period of flourishing, mixed fortune in midlife, and misery in old age.

These lives are permutations of each other—in the sense that for every period of flourishing in one life, there is exactly one such period in the other, and so on—which, on one kind of view, suggests that they are equally good. Still, many people

have the intuition that the former life would be better for you.⁵ The challenge is to find a compelling account of this judgment.

This puzzle about the shape of a life is similar to the puzzle about the shape of history in two important ways. First, it should be easy to see that they are structurally analogous. The obvious difference is that we are not asked to consider the relationship between how well any given generation fares and the value simpliciter of an entire course history, but rather the relationship between how good a person's life is at any given time and how good it is for them overall.

Second, it is common to think about the history of humanity as if it was going through stages characteristic of a single life. For example, in a recent book on existential risk and the future of humanity, Toby Ord (2020, p. 21) writes:

If we think of one million years [that mammalian species typically survive] in terms of a single, eighty-year life, then today humanity would be in its adolescence—sixteen years old, just coming into our power; just old enough to get ourselves into serious trouble.⁶

⁵ In the literature, this intuition is shared by Slote (1983), Bigelow et al. (1990), Velleman (1991), Kamm (1998; 2003), Portmore (2007), Temkin (2012), Kauppinen (2012), Glasgow (2013), Dorsey (2015), and Hirose (2015).

⁶ See also Kavka (1978) and MacAskill (2022). For a critical perspective, see Lenman (2002).

Given these two similarities, it is quite natural to suppose that we can solve the new puzzle about the shape of history simply by extending our best account of the shape of a life. In this section, I argue that this approach does not work. To that end, I discuss three influential accounts of the significance of a life's shape.

The first view I want to consider is *the Temporal Location Account* due to Michael Slote, to whom the contemporary literature on the shape of a life can be traced. Slote (1983, p. 23) writes:

When a personal benefit or good occurs may make a difference to how fortunate someone is ... quite independently of the effects of such timing in producing other good things and of the greater importance we attach to the distinctive goals and interests of certain life periods. And I believe, in particular, that what happens late in life is naturally and automatically invested with greater significance and weight in determining the goodness of lives.

In other words, any given prudential good (e.g., an episode of pleasure, a loving relationship, or an extraordinary achievement)—contributes more to the personal goodness of a life if it occurs in one's adulthood or old age rather than in one's youth. That is why an improving life, in which the best things happen late in life, is better than a deteriorating life.

Proponents of this view could suggest that something similar applies to the shape of history. In particular, they could assert that *Progress* is better than *Decline* because what happens later in history (e.g., that people have happy lives, that relations of equality obtain between members of the moral community, or that people act justly) has greater weight in determining its overall value.

Before we examine the plausibility of this account, let me introduce another popular view: *the Pattern Account*. Frances Kamm (2003, p. 222) characterises it as follows:

I believe that where in a life story some event occurs can be important because the pattern of one's life can be important. ... So it is better to start off badly in life and head toward improvement than to start off well and head toward decline, even when we hold constant all the goods and bads that are distributed in the two different patterns.

In other words, the fact that your life features an improvement is a source of well-being (and that your life features a deterioration is a source of ill-being), quite independently of any attitudes you might have towards these trends and their impact on your well-being. This view has also been endorsed by the likes of Larry Temkin (2012), Joshua Glasgow (2013), and Iwao Hirose (2015).

When confronted with the puzzle about the shape of history, proponents of the Pattern Account could likewise say that a pattern of improvement in terms of any value (e.g., aggregate well-being, equality, epistemic value, and aesthetic value) is in itself good, and that explains why *Progress* is better than *Decline*.⁷

An undeniable attraction of the Temporal Location Account and the Pattern Account is their simplicity. But I believe that this feature comes at the expense of extensional adequacy. To show that, I will present three objections which apply to both accounts. For each objection, we will hold some aspect of the human condition constant and then compare two variants: one involving progress and another involving decline. I will suggest that the Temporal Location Account and the Pattern Account *prove too much* in these cases.

The first objection refers to a scenario discussed by Johann Frick (2017, p. 362):

Imagine a world in which each generation of humans dies and vanishes without trace before the next one is born (perhaps, like mayflies, each generation of humans lays eggs before its death, but disappears before their offspring hatched).

⁷ Franz Brentano (1973) endorses this kind of general principle about the value of improvement and calls it 'bonum progressionis'.

Frick invokes this possibility in his discussion of our reasons to ensure the survival of humanity. He takes it to suggest that (at least some of) our reasons are not reasons to ensure that the *human species* continues into the future, but rather reasons that *humanity*—with its sense of history, cultural traditions, and bodies of knowledge—does.

I believe that this case also tells us something important about the significance of history's shape. The specific feature of this 'mayflies' world that I want to focus on is that the fate of each generation is entirely independent of the actions of the previous generations. With this in mind, consider two possible histories of this world. In the first variant, members of the first generation have mediocre lives, but each successive generation fares better than the previous one. In the second variant, it's the opposite: the first generation flourishes, but each subsequent generation does less well than the preceding one. Now, assume that what's responsible for these differing fates is, for example, the frequency of extreme weather events beyond human control. In the first variant, they are initially quite common but then become less frequent, so things improve. In the second variant, it's the opposite.

What should we think about this pair of cases? The Temporal Location Account and the Pattern Account imply that *Progress* is better than *Decline* in this case. However, it seems to me that these two courses of history are equally good. The fact that in *Progress* the best things happen to occur later in history, or that there

is a pattern of improvement, seems completely irrelevant when successive generations, and their fates, are entirely independent. Call this *the Independence Objection*.

For the second objection, consider a world much more like ours at present: socially integrated and with a rich array of traditions, languages, artworks, and scientific projects. Suppose that these bonds and engagements remain stable over time. What finds alteration, however, is the hedonic aspect of people's lives. In the first course of history, the first generation experiences only a modest amount of pleasure, but each successive generation enjoys a bit more. In the second course of history, it's the opposite. In both cases, the total amount of pleasure experienced by people is the same. Here, we could imagine that these changes are due to something trivial. Perhaps in one course of history, fruit and vegetables that make up our diet are initially quite bland but then gradually become more and more delicious, whereas in the other course of history, the opposite happens.

The Temporal Location Account and the Pattern Account both imply that *Progress* is better than *Decline* in this case. This is because, in *Progress*, the more intense pleasures are experienced later in history, and there is a discernible pattern of improvement in terms of aggregate well-being. But, once again, this verdict does not seem right. When the only variation between *Progress* and *Decline* concerns simple pleasures, these courses of history seem equally good. I will refer to this as *the Simple Pleasures Objection*.

The third objection concerns another pair of scenarios. Imagine *a civilisation of explorers* who pride themselves on travels to uncharted territories, ascents of the highest mountains, and crossings of the largest seas. Since there are only so many places to explore, the most significant achievements of this civilisation take place early in history, and so there is a discernible pattern of decline. Compare this with *a civilisation of inventors* who strive to advance science and technology. Because these kinds of advances tend to build upon each other, the most significant accomplishments of this civilisation happen late in history, and so there is a clear pattern of improvement.

According to the Temporal Location Account and the Pattern Account, *Progress* is better than *Decline* in these circumstances. I believe this is a wrong verdict. If we set aside possible disputes about the relative value of their pursuits, these courses of history seem equally good. After all, in either case, the temporal pattern of value fits with the nature of the civilisation's projects. We can refer to this as *the Fitting Pattern Objection*.

I draw two lessons from these three objections. The first is simply that we should reject the Temporal Location Account and the Pattern Account, at least as they apply to the shape of history. But I think that there is also a more general lesson here. The original intuition—that *Progress* is better than *Decline* regardless of what happens in these two courses of history—turns out to be incorrect. Our refined

intuition is that *Progress* is *typically* better than *Decline*, and this is what we must ultimately explain.

At the methodological level, this is similar to how an inquiry into the badness of death might proceed. We might start with the intuition that death is always bad for the person who dies. But then we examine a range of cases and realise that death is only typically bad. In some cases, it can be neutral or even good for the person who dies, for example, if they stand to experience a lot of pain in the future. The task then shifts to explaining why death is bad when it is.

Turn now to the third popular view about the significance of a life's shape: *the Narrative Account*. It holds that the shape of a life indicates the presence of certain narrative relations between discrete events in one's life, which can affect the overall value of a life. David Velleman (1991), who proposes this view, gives two examples of such narrative relations. The first concerns situations in which a person learns from their past mistakes, and the second situations in which a person turns their struggles into a success rather than letting them go to waste. He writes:

A life in which one suffers a misfortune and then learns from it may find one equally well-off, at each moment, as a life in which one suffers a misfortune and then reads the encyclopedia. But the costs of misfortune are merely offset when the value of the latter life is computed; whereas they are somehow cancelled entirely from the accounts of the former (p. 54).

[E]vents in a person's life can borrow significance from both preceding and succeeding events. A particular success can be either a windfall or well-earned reward, depending on the amount of effort that preceded it; the expenditure of a particular effort can be either a good investment or a waste depending on the degree of success that ensues (p. 57).

Variants of this view have also been endorsed by Jeff McMahan (2002), Douglas Portmore (2007), Antti Kauppinen (2015), and Dale Dorsey (2015).

Extending this position to the significance of a history's shape may not seem as straightforward as in the previous two accounts. At least at first blush, humanity is not an agent in the same sense that a single person is: it does not expend effort, achieve success, or read the encyclopaedia.

That said, we should recognise that we sometimes use expressions such as 'humanity has learned from its past mistakes' or 'our generation is thwarting the efforts of our predecessors'. Insofar as these statements make sense—and that might require us to accept a notion of collective agency that applies to large groups of people both at a time and across time—there might be room for an analogous explanation of the significance of a history's shape. In particular, proponents of the Narrative Account could say that *Progress* is better than *Decline* when and because

these histories differ in terms of whether certain narrative relations connect different generations.

It is important to emphasise that the Narrative Account does not imply that *Progress* is always better than *Decline*, and thus avoids proving too much in the three cases discussed above. In particular, there appears to be no difference in terms of learning from past mistakes or turning past struggles into a success when the fates of successive generations are independent of each other, when the only thing that changes over time is the hedonic aspect of people's lives, or when the patterns of temporal value fit the nature of humanity's projects.

However, I believe that the Narrative Account faces two other problems. The first problem is that, in a range of cases, the Narrative Account *proves too little*. The following pair of cases will illustrate this.

Lost Equality: In the early stages of history, humanity enjoys a high degree of equality between members of the moral community. Over the years, however, these relations crumble, and each subsequent generation is less equal.

Found Equality: In the early stages of history, there is a high degree of inequality, but each subsequent generation enjoys greater equality.

Let's assume that these histories are otherwise similar: for example, people do not realise the world is getting less or more equal, or perhaps they do, but this realisation does not affect them. Even with these assumptions, Found Equality seems better than Lost Equality. There is just something regrettable about the failure to maintain relations of equality between people. The Narrative Account, however, offers no apparent explanation for this. These two scenarios do not—or at least need not—differ in terms of whether later generations learned from the mistakes of the earlier ones or whether the struggles of earlier generations were later turned into a success. We can call this *the Incompleteness Objection*.

This brings us to the second problem for the Narrative Account. Proponents of this view might try to fend off the Incompleteness Objection by postulating some further narrative relation that can explain why Found Equality is better than Lost Equality. But I think that this approach runs the risk of foregoing an important theoretical virtue of explanatory unity. Notice that even the two narrative relations invoked by Velleman appear to lack a normatively significant common ground. To be sure, learning from past mistakes and turning struggles into a success do have something in common: for example, they are aggregates of discrete events in a person's life or a course of history. But not all such aggregates of events have evaluative significance. For example, the badness of the pain I experienced when I stubbed my toe yesterday is not cancelled by the fact that I had a delicious kale salad for lunch today. So there must be some further, more specific

feature that makes learning from past mistakes and turning struggles into a success evaluatively significant. But it is at least unclear from Velleman's discussion what this feature is. And the more items we add to the list of narrative relations, the more difficult it will be to identify it. Or, to put it more harshly, the more *ad hoc* the Narrative Account will look. We can call this *the Explanatory Unity Objection*.

One noteworthy attempt to anticipate this charge comes from Antti Kauppinen (2015), who is attracted to the Narrative Account but acknowledges that Velleman offers "no systematic theory of what makes a life story prudentially good" (p. 201). On Kauppinen's preferred version of this view, narrative significance has to do with pursuing one's goals: "events that comprise the agent's life gain in intrinsic value for the agent when they contribute to merited success in pursuit of valuable goals" (p. 218).

This account might be sufficient to dispel the Explanatory Unity Objection. After all, learning from past mistakes and overcoming struggles will often play an important role in goal-directed pursuits, whereas stubbing one's toe before lunch will not. However, I believe that the Incompleteness Objection remains unanswered. Found Equality seems better than Lost Equality, even if the improvement witnessed in the former case is not a product of humanity's pursuit of egalitarian ideals, but rather (say) a serendipitous by-product of a widespread increase in prosperity. Indeed, what strikes me as salient about this pair of cases is not so much that equality was *found* in one course of history, but rather that it was

lost in the other. I will return to this insight in the next section when I compare the Narrative Account with my positive proposal.

In sum, I believe that although the Temporal Location Account, the Pattern Account, and the Narrative Account all have *coherent* extensions to the shape of history, none of these extensions is *plausible*. The first two accounts prove too much, whereas the third proves too little and potentially misidentifies what matters in the comparison between *Progress* and *Decline*.

3. The Conservative Account

The cases discussed in the previous section lead me to the following conjecture. The shape of history matters just in case and because it has something to do with the transmission and sustenance of certain elements of our culture, such as important traditions, collective endeavours, beautiful works of art, or relations of equality. Therefore, we should be able to explain its significance by appealing to a view that emphasises preserving those things.

The view that strikes me as especially promising in this regard is *Conservatism about Value*, which holds that we have a distinctive moral reason to conserve certain non-instrumentally valuable things even when a superior replacement is available. For example, we seem to have a moral reason to preserve the marvellous Golden Gate Bridge, even if we could build a more impressive structure in its place. By contrast, we have no moral reason to preserve an ordinary \$10 bill when it could be

replaced with another \$10 bill. In the literature, this view has been discussed in contexts such as immigration and cultural change (Scheffler 2007); creation and preservation of art and traditions (Cohen 2012); cognitive enhancement (Nebel 2015); and the possibility of human extinction (Frick 2017; Scheffler 2018). But its relevance for matters related to the shape of history is yet to be appreciated.

Conservatism about Value has several features which are worth explaining at the outset. First, what kinds of things do we have a moral reason to conserve? In my view, there are many non-instrumentally valuable things like that: not only beautiful works of architecture but also unique languages, important traditions, collective endeavours, or relations of equality. That said, not all non-instrumentally valuable things make the list. For example, some works of art (such as a firework display or graffiti) have a transient character, and we should not try to preserve them indefinitely. More generally, to determine whether we have a reason to conserve any particular thing, we must examine whether conserving it is a fitting response to its value and nature.⁸

Second, conservative reasons are *pro tanto* reasons. If a new work of art, tradition, or endeavour would be vastly more valuable (and not just slightly so), then we might have overall reason to create or pursue it at the expense of the old one.

⁸ This view is close to Cohen's (2012) and Frick's (2017). Somewhat different takes on the scope and ground of conservative reasons are offered by Scheffler (2018), who grounds conservative reasons in our valuing attitudes, and Nebel (2022), who grounds them in the concern for the good of the valuable thing itself. Each of these views could in principle account of the shape of history developed here.

After all, the prospective value of the potential replacement also confers a reason on us, which might prove stronger.

Third, conservative reasons are moral reasons and not, for example, prudential reasons. It's not just me who appreciates the beauty of the Golden Gate Bridge that has a reason to prevent its destruction, but everyone—even people who have never seen it and have no interest in doing so.

Fourth, we rarely have a reason to preserve a valuable object 'as is'. Any valuable object is composed of many features which have different valence. Plausibly, we only have a reason to preserve the value-conferring features. For example, we have no reason to preserve a layer of dust that has gathered on the surface of Picasso's *Les Femmes d'Alger*. Moreover, even with respect to those value-conferring features, it can sometimes be appropriate to allow them to change to some degree. For instance, cultures and languages naturally evolve over time, and stopping this process would threaten their very survival.⁹

So much for the gist of Conservatism about Value. This is the first element of my solution to the puzzle about the shape of history.

The second element is a principle that bridges the gap between what we have moral reasons to do and what makes a state of affairs valuable. This principle is needed because Conservatism about Value is cashed out in deontic rather than

⁹ For an insightful discussion of cultural change, see Scheffler (2007).

axiological terms: it says that we have a *reason* to conserve a valuable thing, but not that it would be *better* to do so. Indeed, one might think that conservative reasons stand in opposition to considerations of value: for example, we should preserve an impressive artwork even though replacing it with a different one would make things better.

However, we only need a less restrictive view about what makes a state of affairs valuable. As Derek Parfit (2016, Ch. 57) points out, certain acts seem to be intrinsically bad, in the sense that they are in one way bad even when they have no further effects on what is good or bad. For example, it seems to make a state of affairs in one respect bad when people deceive each other, act unjustly, or break promises.¹⁰ I find this view plausible. But I also want to go one step further than Parfit. To my mind, what he describes is just one instance of a broader issue. Just as it is intrinsically bad when we deceive and coerce people, it is also intrinsically bad when we fail to treat certain non-instrumentally valuable things, such as important cultural artefacts, in a fitting way, for example, by destroying or neglecting them when we could instead preserve them. Thus, on my view, it's not just that we have a moral reason not to destroy the Golden Gate Bridge; it is also intrinsically bad to do so.¹¹ I will refer to this as *the Bad Acts Thesis*.

¹⁰ See also Temkin (2012, p. 205) and Nebel (2015).

¹¹ How bad is it, though? More generally, how are we to assess the badness of a failure to conserve in any particular case? The following principle strikes me as plausible: the badness of a failure to conserve some non-instrumentally valuable thing X is proportional to the

There is, of course, more to say about each of these views, Conservatism about Value and the Bad Acts Thesis. However, my aim in this paper is not to defend them in isolation but rather to demonstrate the explanatory power they have when combined. In particular, I want to show that they provide the best explanation of when and why the shape of history matters.

According to *the Conservative Account* I propose, *Progress* is better than *Decline* just in case and because these courses of history differ in terms of whether certain valuable elements of our culture—such as important traditions, collective endeavours, relations of equality, or beautiful artworks—have been conserved across generations. In cases of *Decline*, the extent of the failure to conserve these valuable elements of our culture is typically greater than in cases of *Progress*, and such failures are intrinsically bad. So, *Progress* is typically better than *Decline*. This is the refined intuition for which we wanted to provide a theoretical rationale.

I believe this Conservative Account offers a better explanation of the significance of a history's shape than the Temporal Location Account, the Pattern Account, and the Narrative Account. To see that, consider how it handles the five objections discussed in the previous section and how it differs from those three accounts.

value of X. Thus, to figure out how bad it is to fail to conserve the Golden Gate Bridge, century-long relations of equality, or one of humanity's prized collective projects, we need to first determine how valuable each of these things is. I cannot hope to do that in this paper but see, for example, Temkin (1993) for a discussion of the value of equality and Bradford (2015) for an account of the value of achievement.

First, take the Independence Objection. In the ‘mayflies’ world, in which each generation disappears without a trace before the next one is born, no valuable elements of humanity’s culture are transmitted to the next generation anyway. So, there is no conservative difference between *Progress* and *Decline*, and this aligns with the intuitive judgment that neither of these courses of history is better than the other.

Second, consider the Simple Pleasures Objection. There, we looked at a pair of cases in which only the hedonic aspect of people’s lives changed over time. Because simple pleasures are not the kinds of things that we have conservative reasons (or even means) to preserve, the Conservative Account avoids the implication that *Progress* is better than *Decline* in this case.

Third, turn to the Fitting Pattern Objection. This objection referred to a comparison between a civilisation of explorers whose most significant achievements were early in history, and a civilisation of inventors whose most significant accomplishments occurred later in history. Although collective projects and traditions are among the kinds of things that we do have reasons to conserve, neither the civilisation of explorers nor the civilisation of inventors fails to do that. The timing of their successes and the resulting pattern of value align with the nature of their projects. So, once again, there is no conservative difference between *Progress* and *Decline*.

We can thus see that—unlike the Temporal Location Account and the Pattern Account—the Conservative Account avoids proving too much in the relevant cases.

Let me now turn to the challenges levelled against the Narrative Account. The Incompleteness Objection referred to the contrast between Lost Equality and Found Equality. The Conservative Account provides a natural explanation for why the former is worse. In Lost Equality, later generations fail to maintain the non-instrumentally valuable relations of equality that obtain between members of the moral community. This is intrinsically bad for them regardless of whether they in fact have egalitarian goals and ideals. Found Equality is different. There, such relations come to obtain only later in history, and the egalitarian improvement we witnessed speaks to the fact they have been maintained to a large extent. Thus, even if these two courses of history feature equivalently long periods of equality and inequality, the Conservative Account correctly implies that Lost Equality is worse than Found Equality. So, the Conservative Account avoids proving too little.

What about the Explanatory Unity Objection? The Conservative Account identifies a single evaluatively significant feature in all relevant comparisons. In particular, *Progress* is better than *Decline* just in case and because certain non-instrumentally valuable elements of our culture have been conserved to a greater extent in the former course of history than in the latter. To be sure, the relevant elements of our culture include things as diverse as beautiful artworks and relations

of equality, but what they have in common is that they are all among the things which are worth preserving even in the face of replacement. So, the Conservative Account is explanatorily unified.

It is worth taking a moment to reflect on the similarities and differences between my Conservative Account and the Narrative Account. The main similarity concerns the structure of these views. They both recognise that the temporal location and order of events in a course of history matters, but not in the kind of simple way that the likes of Michael Slote and Joshua Glasgow suggest. Instead, these views point to some specific feature that typically, but not always, renders *Progress* better than *Decline*.

But these two views differ in other important respects. First, they recognise different features as evaluatively significant. In the case of the Narrative Account, at least according to Kauppinen's characterisation, that feature is the causal contribution that certain events make to the merited success in pursuit of valuable goals. On the Conservative Account, it's the extent to which certain non-instrumentally valuable elements of our culture have been conserved across time.

Second, as we have just seen, these two accounts differ in terms of their explanatory power. The Narrative Account fails to explain our intuitions about certain cases in which the improvement or deterioration of humanity's condition is unrelated to things like learning from past mistakes and overcoming struggles. The contrast between Lost and Found Equality was one example. Still, others are easy

to come by: we just need to substitute equality for some other non-instrumentally valuable cultural artefact that is worth preserving irrespective of any particular person's goals and interests.

Third, these two accounts differ in terms of *where* they locate the relevant evaluatively significant feature. The Narrative Account claims that *Progress* is better than *Decline* when and because there is something good about *Progress*: namely, that certain events in history play a causal role in humanity's merited success in the pursuit of valuable goals. By contrast, the Conservative Account claims that *Progress* is better than *Decline* when and because there is something bad about *Decline*: it involves a failure to conserve certain non-instrumentally valuable elements of our culture.

These three differences, I believe, show that the Conservative Account is significantly different from, and superior to, the Narrative Account.

4. The allure of progress and the tragedy of decline

Having made a positive case for the Conservative Account, I now want to discuss three challenges to this view. While neither challenge ultimately succeeds, this discussion will enhance our understanding of the proposed account and its implications.

The first challenge latches on to the fact that the Conservative Account solves the puzzle about the shape of history in an asymmetric way. As we have just

seen, this view holds that *Progress* is better than *Decline* when and because there is something bad about *Decline*. But one might think that there must also be something good about *Progress*. To make this vivid, we might contrast *Progress* with a new scenario. In *Plateau*, humanity's fortune is mixed at the beginning of history and stays roughly at the same level for the remainder of our time. Now, the objection goes, *Progress* seems better than *Plateau*, whereas the Conservative Account fails to explain that. Call this *the Plateau Challenge*.

This objection rests on a mistake about the content of conservative reasons. In some cases, Conservatism about Value gives us a reason to merely *preserve* the valuable thing in question. For example, in the Grand Canyon case, all we have to do is try not to turn it into a landfill or a motocross park. In other contexts, however, a more active stance might be required. For instance, conserving the Golden Gate Bridge involves performing various maintenance works regularly. Moreover, when it comes to specific collective projects that humanity might undertake, I think we have a conservative reason to *cultivate* them. Take the example of the fight against malaria, a disease that kills over half a million people yearly. A society which sets out to fight malaria has a reason to pursue research programmes into novel forms of treatment, improve access to healthcare facilities in the affected areas, and educate people about available prevention measures, among other things.

Now, I find it plausible that whenever such efforts are undertaken, they will typically improve the human condition. For example, we should expect the fight

against malaria to reduce the incidence or the burden of that disease or at least lay the foundations for future breakthroughs in that domain. Conversely, if we witness no improvement to the human condition over many generations, this suggests that certain valuable elements of our culture—in this case, humanity’s collective endeavours—have not been adequately conserved. In other words, *Progress* is typically better than *Plateau* not because there is something good about progress, but rather because there is often something bad about its absence.¹²

This response can be strengthened by the following, more speculative thought. Many of the aforementioned thinkers who believe that the condition of humanity is getting better view progress not as a series of happy accidents but rather as a natural and virtually inevitable aspect of human activity. In the same spirit, we might think that some valuable elements of our culture improve with time as long as they continue to be instantiated. For example, democratic institutions might be thought to have this profile. If that’s right, this gives us another reason to believe that the absence of improvement to the condition of humanity indicates that certain valuable things have not been adequately conserved.

The second issue that I want to discuss concerns radical progress. There were several moments in history when humanity truly made a leap forward: the Agricultural Revolution around 12,000 years ago, the Industrial Revolution of the

¹² Isn’t *Progress* always better than *Plateau* though? I think it isn’t, for roughly the same reasons that *Progress* is not always better than *Decline*.

18th and 19th centuries, and the Information Revolution that started in the 1990s. When we reflect on these events, it is clear they were occasioned by the destruction or overhaul of many valuable elements of our culture. The Industrial Revolution, for example, transformed virtually every aspect of our society and daily life, ranging from political and economic arrangements to the nature of work and pastime. In virtue of this, it looks like Conservatism about Value which underlies the Conservative Account must regard the Industrial Revolution as something that we should regret. But, to many people, it seems to be a highly welcome episode in the history of humanity. We can refer to this as *the Radical Progress Challenge*.

My response to this objection has three components. First, I want to emphasise that many elements of the pre-industrial world were not non-instrumentally valuable: for example, widespread illiteracy, pervasive malnutrition, and feudalism. From the conservative point of view, there is no reason to regret the disappearance of these things. So, Conservatism about Value opposes only some of the many aspects of the transformation brought about by the Industrial Revolution: such as the erosion of children's and women's rights in the early industrial years, linguistic standardisation, or the destruction of the natural environment.

Second, we must remember that conservative reasons are *pro tanto* reasons. This means that proponents of Conservatism about Value can judge radical progress as bad in one respect but good overall when the new ways of life are significantly better than the old ones. Indeed, in virtue of alleviating poverty alone, the Industrial

Revolution appears to have brought about improvements to the human condition so profound as to trump the destruction of some valuable modes of existence.¹³ So, the conservative critique of the Industrial Revolution does not amount to its rejection.

Finally, I think that on deeper reflection, we rarely support radical progress categorically, even in its grandest instances. Instead, our outlook on events such as the Industrial Revolution is inherently divided. We regret the destruction of some ways of life and welcome the appearance of others, even when the new ways of life are significantly better. A major virtue of Conservatism about Value is that it seems to be uniquely equipped to explain this kind of evaluative ambivalence due to its emphasis on preserving valuable objects even in the face of replacement.

The third challenge concerns the tragedy of decline. In particular historical cases, it might seem that the civilisational collapse was attributable entirely to factors beyond human control. For example, it could be that the Easter Island civilisation was always doomed to collapse because of the limited resources available on the island and its extreme isolation (the two nearest lands are Chile, 2,300 miles to the east, and the Pitcairn Islands, 1,300 miles to the west). Or take the bubonic plague pandemic of the 13th century, which killed as much as 50% of the European population and was responsible for a widespread social and economic upheaval.

¹³ In 1820, an estimated 83.9% of the world population lived in extreme poverty (Bourguignon and Morrison 2002). Nowadays, it's less than 10% (World Bank 2020).

These instances of *Decline* seem tragic, and they seem to be worse than the corresponding instances of *Progress*. But, the objection goes, the Conservative Account cannot explain that because no failure of conservation was involved, just bad luck. We can call this the *Bad Luck Challenge*.

My issue with this challenge is that it rests on an overly simplistic understanding of civilisational collapse. The societies in question were not entirely faultless with respect to their decline. For example, while Easter Islanders had no control over how much timber was available on their land and the pace of reforestation, they were undoubtedly responsible for the demand side of the equation. In particular, some historians believe that what pushed their civilisation beyond the levels of sustainability was the construction of the famous stone statues (*moai*), which was driven by the competition for status between local tribes. Close to 1,000 statues weighing up to 75 tons each have been discovered on the island, and erecting each required great amounts of timber for sledges and levers, among other natural resources. Likewise, although Europeans were presumably unable to completely prevent the bubonic plague pandemic, its ramifications would not have been so disastrous if not for the overpopulation and widespread famines in the preceding decades that increased people's vulnerability to pathogens. In general, I believe that closer inspection reveals that historical cases of civilisational collapse were at least as much a product of negligence, internal struggle, lack of foresight,

and other failures of conservation, as they were of external shocks.¹⁴ If that's right, then the Conservative Account does explain why these scenarios strike us as tragic.

5. Conclusions

In closing, let me take stock of the discussion so far, consider some of its implications for matters related to the future of humanity, and then return briefly to the puzzle about the shape of a life.

We started with the following question: is it better if the history of humanity features a pattern of improvement rather than deterioration? My aim in this essay has been to persuade you that the answer is not a simple 'yes' or 'no', but rather 'typically'. According to the Conservative Account, *Progress* is better than *Decline* when and because these courses of history differ in terms of whether certain valuable elements of our culture have been conserved across generations. In cases of *Decline*, the extent of the failure to conserve these things is typically greater than in cases of *Progress*, and such failures are intrinsically bad. Therefore, *Progress* is typically better than *Decline*.

But there are exceptions. In some rare cases, the extent of conservation in *Progress* and *Decline* is the same, so these histories are equally good. We saw this when we discussed the Independence Objection, the Simple Pleasures Objection, and the

¹⁴ For an extensive discussion of this issue, see Jared Diamond's tellingly titled book: *Collapse: How Societies Choose to Fail or Succeed* (2011).

Fitting Pattern Objection. In other rare cases, the extent of conservation in *Progress* is lower than in *Decline*, so the former course of history is worse. For example, we could imagine that *Progress* is a story of extreme negligence and serendipitous hedonic improvement which outweighs the badness of that negligence. In contrast, *Decline* is a scenario in which a civilisation is exceptionally diligent about conservation but unfortunately experiences a major hedonic deterioration over time.

This account of the shape of history carries a number of important lessons for matters of public policy and individual conduct concerning the future of humanity. Let me mention two.

The first lesson concerns the importance of sustainable development. To bring this out, consider the following case.

Compound X. Suppose that we discover an extraordinary resource, Compound X. It can be used to greatly improve the standards of life worldwide and advance our civilisation in many respects. But it will eventually start to run out, and we will not be able to replace it. And when that happens, the condition of humanity will deteriorate.

How should we exploit Compound X and distribute it among different generations? In particular, should we use a lot of it now, reach the heights of our civilisation quickly, and then witness it gradually stagnate as we run out of this

resource? Or should we save most of it for future generations, gradually phase it in, and aim for a steady improvement until the end of our time?

If the shape of history did not matter, then either option would be appropriate, other things being equal. The Conservative Account, however, suggests that we should do the latter, and I believe that it is correct in this regard.

This case strikes me as a helpful metaphor for humanity's current predicament. While no single resource might be quite as powerful and decisive with respect to humanity's fate as Compound X, the totality of the resources that we are currently exploiting—fossil fuels, freshwater stores, fertile soils, biodiverse ecosystems, and even things like clean air, relatively low sea levels, and largely moderate climate—might prove to be. In these circumstances, sustainable development is of paramount importance.

There are two aspects of sustainable development that I want to emphasise here. For one thing, when deciding how to allocate our resources among different generations, we should take care to leave enough for our distant descendants. Crucially, 'enough' here does not mean just 'enough to survive' or 'enough to have lives worth living'. Instead, it means 'enough to sustain and cultivate the valuable elements of our culture'. That is a significantly higher standard.

For another, when choosing how to exploit these resources, we should be mindful not to develop our civilisation and culture too quickly. This is not, as the Temporal Pattern Account has it, because it's better if the best things happen later

in our history. Instead, it's because there is something distinctively bad about a situation in which a society—not unlike the Roman Empire or the Easter Island Civilization—creates an extraordinary amount of cultural artefacts and develops other aspects of their civilisation to a point such that the next generations will foreseeably struggle to maintain and cultivate them.

It is worth noting that, in the case of developing at an unsustainable pace, moral failure seems to lie primarily in the actions of the earlier generations rather than the later ones. In other words, what's bad is not that the later generations have failed to preserve the cultural artefacts created by the earlier ones, but rather that the earlier generations brought about circumstances in which appropriate conservation was difficult or impossible. I believe this is because our conduct in these matters is constrained by what some philosophers call *standard-regarding reasons*. For example, in his discussion of the procreation asymmetry in population ethics, Frick (2020) argues that if one cannot give a person a life worth living, then they have a standard-regarding reason not to create that person. I find it plausible that decisions regarding creating non-instrumentally valuable cultural artefacts are subject to similar moral standards. These moral standards are satisfied when such things are adequately conserved, and they are not satisfied when these things are instead neglected or destroyed. Thus, like bringing a child into this world while knowing that you will not be able to provide for it, creating certain cultural artefacts

with the awareness that they will not be adequately conserved in the future is something that we have a moral reason not to do.

This brings us to the second lesson of my account for matters of public policy and individual conduct. In my discussion up to this point, I have considered different trajectories that the condition of humanity might take during its tenure. But, almost inevitably, this tenure will someday come to an end. It is, therefore, worth asking: what should we do if the prospect of our extinction becomes imminent? Should we go about our lives pretending nothing has changed, or perhaps dramatically alter our ways of life?

To illustrate this issue, consider the famous scenario from P. D. James's novel, *The Children of Men*. Suppose that we discover that all human beings have become infertile. Although no one will die prematurely, there will be no next generation, and our species will go extinct within a century. If the events of the novel are any guide to human psychology, those unfortunate enough to belong to this last generation would, in addition to experiencing grief and distress, find themselves unwilling to open new lines of scientific inquiry, start political movements, build new marvels of architecture, or set foundations for new traditions.

It might be initially tempting to think of this quiescence as blameless—given the extraordinary circumstances—but nonetheless imprudent. After all, the kinds of endeavours described just above greatly enrich our lives, and the people depicted in James's novel would seem to deprive their already impoverished lives of such

opportunities needlessly. However, if I'm right about the importance of conservation and the existence of standard-regarding reasons, then this behaviour is precisely what morality calls for. While we should continue to preserve the valuable things that already exist, we should not create cultural artefacts and initiate collective endeavours if there will be no one to carry the torch further into the future.

In closing, let me briefly comment on two further issues. The first is the relationship between the Conservative Account and *longtermism*, the increasingly popular view that a key moral feature of our actions today is their impact on the far future.¹⁵ The main similarity between these two views is that they endow the long-term consequences of our actions with special importance. But there are significant differences. In the first place, many longtermists believe that the long-term consequences of our actions are *overwhelmingly* important—in the sense that, for practical purposes, we can often ignore the short-term consequences of our actions.¹⁶ The Conservative Account is less radical. It merely says that the long-term consequences are, *in one respect*, more important because they determine whether the history of humanity is that of negligence or conservation. Another difference concerns the premises from which the conclusion about the special importance of the long-term consequences is derived. Though there are variations, the standard

¹⁵ For a key statement and defence of longtermism, see Greaves and MacAskill (2021).

¹⁶ This view is known as 'strong longtermism'.

case for longtermism relies on several controversial assumptions: for example, that the expected number of people is as high as 10^{24} , that creating happy lives is just as important as improving the lives of existing people, or that we should be trying to maximise expected total welfare (Greaves and MacAskill 2021). By contrast, the Conservative Account does not utilise these or other similar assumptions and, in virtue of that, should be more appealing to consequentialists and non-consequentialists alike.

The second issue that I want to address is the shape of a life. One may wonder whether we can also appeal to the Conservative Account to explain why an improving life seems better than a deteriorating life. I believe that we can, but that explanation would require three novel claims. The first concerns the scope of conservative reasons. While discussions of Conservatism about Value have focused on things which are valuable simpliciter (such as cultural artefacts), I think that we also have a distinctive moral reason to conserve certain prudential goods (such as loving relationships and important personal projects). The second claim is about the determinants of the goodness of a life. To explain the significance of a life's shape, we would have to show that failing to conserve certain non-instrumentally valuable things does not just make a state of affairs worse, but that it also makes one's life worse. The third claim concerns our intuitions. I suspect that an improving life is not always better than a deteriorating life, but rather better just in case these patterns of value have to do with the conservation (or neglect) of loving relationships and

important personal projects. *If* these three claims are true, we can explain the significance of a life's shape as follows. An improving life is better than a deteriorating life just in case these lives differ in terms of whether certain prudential goods, such as loving relationships and important personal projects, have been conserved over time. The extent of the failure to conserve these things is typically greater in a deteriorating life than in an improving life, and such failures are bad for a person. Thus, an improving life is typically better than a deteriorating life.

Are these three claims true, though? I find each of them plausible, but the task of defending them and exploring their broader implications for how to direct our lives is best left for another occasion.

Chapter 3: Conservatism about Prudential Value

1. Introduction

During his early years in Paris, Pablo Picasso suffered from discouragement, hunger, and the piercing cold of continental winters. While many paintings survive from this ‘Blue Period’, the artist is believed to have burned several others, desperate to keep his hands warm as he worked on new pieces in his unheated apartment on Boulevard Voltaire. Suppose that these rumours are accurate and that one night Picasso burned his masterful *Absinthe Drinker* (1901) to create another painting that, let’s assume, was even more valuable.

Situations like this have a curious feature. The loss of the *Absinthe Drinker* strikes us as worthy of regret, even though it is replaced by another painting that is no less valuable. Moreover, it seems that Picasso had a distinctive reason not to destroy the *Absinthe Drinker* and that this reason was stronger than the reason to create another beautiful painting.

These kinds of considerations motivate a more general thesis about value. According to *Conservatism about Value*, we have a distinctive moral reason to conserve certain non-instrumentally valuable things, such as architectural wonders, important traditions, and beautiful paintings. At the minimum, this means that we should not destroy them even when they could be replaced by similar things of equal or even

greater value. By contrast, we have no moral reason to preserve merely instrumentally valuable things, such as a generic \$10 bill, when another \$10 bill would replace it.

Proponents of this view include Samuel Scheffler (2007; 2013; 2018), G. A. Cohen (2012), Jacob Nebel (2015; 2022), and Johann Frick (2017). These authors disagree about the precise range of things we have a reason to preserve and the ground of this concern. Still, one aspect that their discussions have in common is the almost exclusive focus on things which are valuable simpliciter, at the expense of those which are valuable for particular people. My aim in this paper is to fill this gap by offering a theory of *Conservatism about Prudential Value*.

In the first part, I provide an initial motivation and a theoretical foundation for this view. In Section 2, I present two examples of prudential goods—loving relationships and important personal projects which culminate in achievements—that we seem to have a moral reason to preserve even when they could be replaced by something at least as valuable. In Section 3, I examine what explains the presence of conservative reasons, how strong these reasons are, which things we have reasons to conserve, what exactly we have reasons to do, and to whom these reasons apply.

In the second part, I offer further support for Conservatism about Prudential Value by showing that it elucidates several puzzling questions and sentiments concerning how we should direct our lives. These include: how to balance exploration and commitment at different stages of life (§4), whether the

reluctance to start new projects and relationships in older age is justified (§5), when it is permissible to move on from a loss (§6), and why an improving life seems better than a deteriorating one (§8).

My discussion of these issues has two striking upshots. On the practical side, it implies that we cannot start a new life whenever we please. Our existing loving relationships and important personal projects warrant a certain kind of loyalty, in virtue of which we have reasons to conserve them. On the theoretical side, it gives us an insight into the normative status of relationships and projects. We should be concerned with the survival and flourishing of these things above and beyond the contribution that they make to our well-being, and some of our reasons to preserve them are moral and agent-neutral, just like our reasons to preserve the Grand Canyon or the *Mona Lisa*.

A few clarificatory remarks before we begin in earnest. First, the presented arguments assume an *objective list theory of well-being* that recognises multiple ingredients of well-being. These include not just pleasant experiences but also loving relationships and personal projects culminating in achievements, among other things.

Second, Conservatism about Prudential Value is meant to be understood as a component of Conservatism about Value. The latter view says that we have a distinctive moral reason to preserve certain non-instrumentally valuable things. The

former points out that these things include loving relationships and important personal projects.

Third, a word on methodology. This paper does not offer anything like a formal proof of Conservatism about Prudential Value. In value theory, and at a certain level of generality, a careful discussion of cases and an application of the method of reflective equilibrium are sometimes the best resources we have. In this spirit, I start with two examples of prudential goods we have a reason to preserve even when a replacement is available. The remainder of my case for this view rests in showing that it provides a compelling answer to a number of ethical puzzles.

2. Two examples

To motivate Conservatism about Value, its proponents often reach for cases in which a non-instrumentally valuable thing is replaced by another thing of the same kind which is equally or more valuable, and yet the entire transaction leaves considerable moral residue. Picasso's tragic decision to set one of his masterful paintings aflame to finish another piece is a case in point. Other standard examples include historical artefacts, cultures, institutions, and great works of architecture. What has been neglected, however, is that there are analogous scenarios involving things which are prudentially valuable rather than valuable simpliciter.

To begin with, consider the case of Jim Brown, a former fullback for the Cleveland Browns. In 1966, Brown shocked the football world by announcing his

retirement from the sport at the young age of 29. His decision was not motivated by an injury or a decline in performance. In fact, he had just won the NFL Most Valuable Player award for the third time in nine years and was on track to receive further accolades. Instead, Brown left football to pursue a different career, acting, where he also enjoyed great success. He became arguably the first Black action star, appearing in over 50 movies and TV shows, including many blockbusters.

If you are like most football fans, Brown's story evokes in you the attitude of *evaluative ambivalence*. On the one hand, there is plenty to cheer for: his post-football life was oriented towards a worthy personal goal and rich in achievements. On the other hand, Brown's decision seems, in one respect, regrettable. Crucially, this is not just because he missed out on athletic accomplishments that lay ahead of him but also because of what he had left behind. He quit a career that was a non-instrumentally valuable personal project that he gave up much for and that still had plenty to give back to him.

We encounter a similar pattern of intuitions in the context of loving relationships. Consider the following fictional case. Ann is in a loving relationship with Beth. One day, during a tough run workout at the golden-surfaced track of the historic Edwards Stadium in Berkeley, she meets Sally, who happens to be very similar to Beth in all respects except one. Like Ann, Sally is a dedicated runner, so the pair could enjoy some of their training together and support each other during

relationships, which would bring them even closer. After some deliberation, Ann ends her relationship with Beth and enters one with Sally.

Here, too, we witness evaluative ambivalence. In one respect, Ann's decision seems wise. Any reasonable theory of well-being will recognise that relationships are among the key ingredients of the good life and that they can be better or worse. From the case description, it is clear that the new bond is better for Ann, though perhaps only slightly so. But there is also something troubling about Ann's decision. That old relationship was valuable in itself and would have continued to be so, for all we know. Thus, Ann's choice leaves us with considerable normative residue.

I think that this normative residue suggests that we have a distinctive reason to conserve loving relationships and important personal projects even when a replacement is readily available. Crucially, this reason is not conceptually tied to anyone's interests. After all, Jim Brown had a reason to preserve his football career, and Ann had a reason to preserve her previous relationship even though the new engagements would be equally good or even better for them. Thus, the reason in question seems to be of moral—rather than prudential—variety. This is precisely what Conservatism about Prudential Value seeks to capture.

Some readers might object to the preceding argument. They might concede that, all-things-considered, Jim Brown and Ann should have preserved their original engagements but deny that we need to appeal to Conservatism about Prudential Value to explain this intuition. As numerous authors have emphasised, what ties you

to a particular person are not just their generic qualities, such as wit, beauty, or intelligence, but also your *shared history*: for example, that you shared many good experiences, endured suffering together, or jointly raised a child.¹ One might therefore suggest that starting a new relationship with Sally would be overall worse for Ann than continuing the old one with Beth because the former lacks that shared history. Likewise, one might argue that what ties you to a particular project is the history of your prior involvement, and thus it was overall worse for Jim Brown to start an acting career than to continue his athletic career.

These remarks are important because they remind us about the features that make our attachments valuable. But I think that the objection does not carry the day. While it is true that a new relationship or project would *initially* lack the relevant historical qualities and thus be less valuable, there is no principled obstacle to these new engagements *eventually* acquiring such attributes, and thus becoming equally or even more valuable. So let's just explicitly assume that the new engagements would be more valuable. When we do so, the intuition that Ann and Jim Brown each had a distinctive reason to preserve their original engagements perseveres. And, since we have stipulated that it was in their personal interest to do otherwise, this reason must be a moral reason. This lends support to Conservatism about Prudential Value.

¹ See Kolodny (2003) and Hurka (2016) for this kind of view. Strictly speaking, both of these authors are concerned with what gives us reasons for love rather than with what makes a relationship valuable, but their remarks seem to extend to the latter setting as well.

3. Theoretical foundations

Having presented the initial motivation for Conservatism about Prudential Value, I now want to explain the ground, character, strength, scope, and content of our conservative reasons.

First, can we ground our moral reasons to conserve prudential goods in more fundamental normative truths? Here is one attractive line of thought. Things which are non-instrumentally valuable command a certain kind of response from members of the moral community. This may involve holding specific attitudes or taking certain actions: respecting, protecting, conserving, savouring, etc. These considerations can give rise to moral reasons. For example, as Cohen (2012) argues, we seem to have a moral reason to preserve the impersonally valuable traditions of Oxford's All Souls College and Michelangelo's *Slaves* when and because this is a fitting response to their value.

The examples featuring Jim Brown and Ann suggest that loving relationships and important personal projects also have this profile: we have a moral reason to conserve them when and because this is a fitting response to their value.² Crucially, one can have a conservative reason to preserve a loving relationship or an

² In a recent paper, Nebel (2022) distinguishes between three ways of grounding conservative reasons: (i) in fitting response or, as he calls it, respect, (ii) in valuing attitudes, as proposed by Scheffler (2007; 2013; 2018), and (iii) in the concern for the good of the valuable things, as suggested by Nebel himself. Each of these accounts could in principle support Conservatism about Prudential Value, but it is helpful to adopt a specific one in this paper.

important personal project even if it's not in their interest to do so. In this respect, a conservative reason to preserve a loving relationship is like a conservative reason to preserve something valuable simpliciter, like the Grand Canyon.

Second, conservative reasons to preserve prudential goods have an *agent-neutral* character. In the case of the Grand Canyon, everyone has a reason to preserve it, not just you and me. Likewise, it is not just me who has a conservative reason not to destroy my loving relationship but also you, a third party, if it is within your power to do so.

The picture that emerges from these considerations is thus reminiscent of the Moorean view about why we should promote people's well-being. Donald Regan (2004, p. 204) characterises it as follows:

On this approach, any agent's fundamental obligation is to bring about the existence of as much good as possible. The 'good' asserted here is an agent-neutral good, not tied conceptually to any person's interests. But in fact the most important goods are certain kinds of events and experiences that occur within individual lives. Therefore, an important part of promoting the good turns out to be promoting certain events and experiences in individual lives.

Of course, Conservatism about Prudential Value is not concerned with bringing about the existence of loving relationships and important personal projects, but

rather with preserving these things. Still, it likewise postulates no profound normative distinction between what is good for particular people and what is good simpliciter. We have a conservative reason to preserve certain non-instrumentally valuable things, some of which happen to occur in people's lives. Therefore, an important part of preserving non-instrumentally valuable things turns out to be preserving prudential goods.

To be sure, it is much more common for a person to have a conservative reason to engage in a particular way with their own relationships and projects than with other people's. But that's just because people are generally in a privileged position to influence the trajectory of their own engagements. Only you can cultivate your relationships and projects, and you have more opportunities to destroy or damage them than I do. Moreover, I might have a stronger overall reason not to ruin my relationship than you do because of the special duties that I owe to my loved one in virtue of being in a loving relationship with them.

Third, conservative reasons are merely *pro tanto* reasons. If a new loving relationship or important personal project would be vastly more valuable (and not just slightly so), then taking it up at the expense of an existing one would plausibly be what you all-things-considered ought to do. After all, the prospective prudential value of the potential replacement also confers reasons on us, and these reasons might prove to be stronger.

Fourth, let's take a moment to consider which prudential goods we have a conservative reason to preserve. In the literature on well-being, it is common to distinguish between hedonic and non-hedonic goods. Hedonic goods include simple sensory pleasures like the feeling of a warm shower on your skin and more sophisticated experiences like enjoying a beautiful painting. Non-hedonic goods include loving relationships and personal projects culminating in achievements, among other things. Accordingly, one might suggest that we have no conservative reason to preserve hedonic goods and that we have a moral reason to preserve non-hedonic goods. After all, there appears to be no reason to be concerned whether the pleasant experience of a warm shower on one's skin continues or whether it is supplanted by another pleasant experience, as long the replacement is at least as good. By contrast, as we have seen, there does seem to be a reason to be concerned with the preservation of loving relationships and important personal projects.

However, this partition does not do the relevant explanatory work, and its extension is only roughly correct. Take the issue of explanation first. To determine whether a particular object warrants preservation, we must investigate its nature. I think that what's distinctive about hedonic goods, and what confers their special value on them, is that they are essentially *transient*. They come and go. In virtue of this feature, resisting replacement is generally not warranted. By contrast, many non-hedonic goods are *enduring* or *historical*, for lack of a better term. That is, their special value depends on the fact that they stand in specific relations to our past and persist

through time. In virtue of this, non-hedonic goods typically warrant preservation in the face of replacement. Thus, I believe that it is the distinction between transient and enduring objects that has explanatory significance, not the distinction between hedonic and non-hedonic goods.

What about the extension of the partition between hedonic and non-hedonic goods? I think that when we reflect on the full range of possible relationships and personal projects, it becomes apparent that we do not always have a reason to preserve them. For one thing, some valuable engagements have a natural conclusion point. For example, it's futile to continue pursuing the goal of being the first person to run a four-minute mile if that has already been done, be it by you or someone else. Likewise, it is perfectly fine to let go of that summer romance when leaves start to change their colours. This contrasts with most careers and committed relationships, which do not typically feature an apparent finish line. For another, only our important personal projects warrant conservation. Minor and tentative engagements, such as half-learned languages and half-begun papers, do not call for continuation in the same way that careers and other life-defining projects do.

Fifth, let's now consider the content of conservative reasons. There are two ways in which Conservatism about Prudential Value is more nuanced in this respect than it may initially seem.

The first issue is that we rarely have a conservative reason to preserve a valuable object 'as is'. To see that, note that any loving relationship has many

features, and only some of them are *value-conferring*. Plausibly, these are the only features we have a reason to preserve. The petty arguments you have with your spouse, like the one about what gift to get your mutual friend for their birthday or the envy you sometimes feel about each other's accomplishments, seem to take away from that relationship's value and preserving them is not something required of you.

Moreover, it can sometimes be appropriate to allow for or make changes to the value-conferring features of a valuable object. For example, you might settle on a more modest goal time for your next race. Allowing such developments to occur is not equivalent to abandoning an existing engagement. After all, like cultures and languages, relationships and projects can naturally evolve over time, so to stop this process would threaten their survival.³ Other changes, especially to the value-conferring features, can destroy valuable engagements altogether or turn them into something different. For example, when love between two people is replaced by hate, they are no longer in a valuable relationship.⁴

The second issue is that what we have a conservative reason to do sometimes goes beyond merely ensuring that certain ingredients of our well-being *continue to exist*. This is so in two related ways.

³ For an insightful discussion of cultural change and its desirability, see Scheffler (2007).

⁴ Of course, there will also be cases in which it is not obvious whether a valuable engagement has been destroyed, or merely significantly altered. For instance, suppose that Andreas and Heidi are initially romantic partners, but later choose to relate to each other merely as co-parents and friends. It's contentious whether their original bond has been severed altogether. To determine that in any given case, we need to examine the relevant persistence conditions.

For one thing, when someone lets an extraordinarily valuable relationship deteriorate so much that it survives at a level that is just barely good, their behaviour clearly falls short of being fitting. This suggests that preserving valuable things *in a flourishing state* is important.⁵ And whether that is achieved depends, at least partly, on how valuable the surviving thing is relative to how valuable it was.

Moreover, it is essential to *cultivate* loving relationships and personal projects. Someone who last ran years ago can hardly call themselves a runner. And someone who hasn't come up with any jokes or performed in a long time can't call themselves a comedian. Whatever engagement you take on, you have to keep earning the title.

This concludes my characterisation of the key features of Conservatism about Value. To summarise, on this view, we have a *pro tanto* moral reason to conserve certain prudential goods when and because this is a fitting response to their value. These goods include a range of enduring goods, such as loving relationships and important personal projects. We have a reason to preserve these things even when a superior replacement is available, but we rarely have a reason to preserve them 'as is' or merely to ensure that they continue to exist.

There is, of course, more to say about these features of Conservatism about Prudential Value. Relationships and projects come in many types, and we may want to know precisely what conservative efforts are called for in each case. We will make

⁵ Frick (2017) makes a similar observation about our reasons to ensure humanity's survival.

some progress on these matters in the remainder of the paper, as we investigate a set of puzzles concerning directing one's life in youth, adulthood, and older age; moving on from a loss; and the significance of a life's shape.

4. Exploration and commitment

In the remainder of the paper, I want to offer further support for Conservatism about Prudential Value by showing that it elucidates several old and new puzzles about how we should direct our lives.

The first issue that I will examine is a novel puzzle about the balance between exploration and commitment at different stages of life.

Start with *youth*. Conservatism about Prudential Value might appear to generate a dilemma at this stage of life. On the one hand, if you take on a valuable engagement at a younger age, this choice seems to carry special significance into the future. A budding cellist, for example, appears to be under considerable normative pressure to move forward with her musical education, even if she wishes to turn to boxing instead. Similarly, one appears to have a weighty moral reason to cultivate that first relationship for as long as possible, even when further, more valuable romantic opportunities arise. This may strike us as objectionably constraining.

On the other hand, although we could escape the force of conservative reasons by abstaining from relationships and important personal projects, that

would greatly impoverish our lives. After all, these goods are widely recognised as some of the most significant contributors to our well-being.

Thus, younger people appear to face the unfortunate choice between putting on conservative shackles and missing out on some of the best things in life. We can refer to this predicament as *the Dilemma of Youth*.

This problem admits several possible responses. Let's start with two that target the first horn of the dilemma. One option is to deny outright that the youthful pursuits in question count as genuine, non-instrumentally valuable projects and relationships. If they don't, they do not fall under the scope of Conservatism about Prudential Value, and engaging in them does not constrain our future choices at all. Call this *the No Value Solution*.

However, this solution comes at a high cost. Although it does not compel us to abstain from valuable engagements in youth, it does imply that this part of our life is significantly impoverished anyway, which is difficult to square with our intuitions.

The second response to the Dilemma of Youth is subtler. It concedes that our adolescent engagements are non-instrumentally valuable and instead postulates that this part of life is covered by a grace period. That is, conservative reasons do not apply in youth, and we can take on relationships and projects without the fear of having to keep them up once new opportunities arise. Call this *the Grace Period Solution*.

The Grace Period Solution may seem attractive, but on reflection, I think that it proves too much. While not all youthful engagements warrant conservation, some certainly do. It would be regrettable, for example, for the teenage Saul Kripke to abandon a proof in modal logic a couple of steps away from completion, and not only because he would thereby miss out on a major achievement. More mundanely, perhaps, many of us have been fortunate enough to be at a young age in a loving relationship characterised by a mutual valuing of that bond, emotional vulnerability, and intuitive understanding of the reasons this connection generates.⁶ These kinds of youthful engagements seem to warrant conservation.

In light of these unsuccessful attempts to tackle the Dilemma of Youth, let's ask not if the premises of the argument comprising the dilemma's first horn are true but whether the advertised conclusion follows from them. That is, is it true that the implications that Conservatism about Prudential Value has in this context are objectionably demanding?

To address this question, I want to begin by noting that what's distinctive about youth is that we are still learning a lot. For one thing, we are yet to fully appreciate what's out there: what possible engagements the world can offer us, how valuable they are, how much effort and time they require, and so on. But there is also an important self-knowledge component. We must cast our net widely and

⁶ For this way of thinking about the nature of love, see Kolodny (2003).

explore different possibilities to figure out what we are good at, what brings us joy, and how much we can fit in a day, a year, or a lifetime. As we collect more information, we put ourselves in a better position to decide which engagements to take up in adulthood.

In axiological terms, this means that cycling through many projects and relationships at this stage of life has extraordinary instrumental value for us. Once we take this value into account and acknowledge that the prudential reasons it confers on us will often outweigh the moral reasons to preserve our initial relationships and projects, the implications that Conservatism about Prudential Value has for this stage of life no longer appear to be objectionably demanding. We can call this *the Learning Phase Solution*.

Turn now to *adulthood* and *older age*. The preceding insights also apply to these stages of life. By the time we transition into adulthood, the benefits associated with learning about ourselves and the external world start to diminish rapidly—at least if things go well in one’s youth and one successfully engages in multiple relationships and projects, which admittedly is not always the case. In contrast, the moral reasons to preserve existing engagements retain their weight. As a result, the conservative grip tightens around us at this stage of life. This accounts for our pre-theoretic intuitions about this matter: we tend to accept that our adult engagements place significant and binding demands on our future conduct, unlike those pursued in our youth.

Thus characterised, our decisions concerning which relationships and projects to take on, and whether to preserve them or move on, are reminiscent of *the optimal stopping problem* familiar from mathematics. Brian Christian and Tom Griffiths (2016) illustrate this problem as follows:

You interview the applicants in random order, one at a time. You can decide to offer the job to an applicant at any point and they are guaranteed to accept, terminating the search. But if you pass over an applicant, deciding not to hire them, they are gone forever.

The question, naturally, is when to stop and commit to hiring a candidate. As it turns out, there is a unique solution to this problem. You set a predetermined fraction of the candidate pool—37%, to be precise—for the ‘look’ phase, in which you learn about your options and categorically do not choose anyone. After that point, you enter the ‘leap’ phase, during which you commit to the first option that is superior to all you have already encountered. This strategy maximises the probability of selecting the best candidate. Call this *the 37% Rule*.⁷

Of course, the optimal stopping problem makes several structural assumptions which may or may not be replicated in real life (for example: if you pass

⁷ See Christian and Griffiths (2016), Ch. 1, for an accessible explanation of this result.

over a candidate, they are gone forever). But the general lesson remains useful for our purposes. As we go about our lives, we need to spend a considerable amount of time exploring different loving relationships and personal projects, or else we run a high risk of never discovering the engagements which would be most valuable for us. But it is not permissible for us to keep up this approach for too long because the prudential reasons associated with the instrumental benefits of exploration quickly diminish, and the moral reasons to preserve existing engagements retain their weight. So, there will come the point when the balance of reasons favours conserving an existing engagement over replacing it.

Indeed, if something like the 37% Rule turned out to be the optimal approach to navigating valuable engagements in our lives, that would dovetail with our intuitions. For it would mean that a typical person expecting to live for 80 years should, ideally, spend roughly the first 30 years of their life exploring their options, after which point it's best to commit.

5. New engagements at the end of life

The second puzzling issue that I want to investigate through the lens of Conservatism about Prudential Value concerns older age. In particular, the elderly are often reluctant to engage in new loving relationships and important personal

projects, and in doing so they seem to needlessly deprive themselves of things that greatly enrich our lives.⁸

While Conservatism about Prudential Value is mainly concerned with our existing attachments, I believe that it can also help vindicate this forward-looking attitude. To see that, note that older age is often marked by cognitive and physical decline, which may make it impossible for a person to maintain a valuable relationship or important personal project in a flourishing state. And, intuitively, we have a reason not to take on new commitments when we are aware that this is likely to happen.

The reason in question seems to be an instance of what Johann Frick (2020) calls *standard-regarding reasons*, which, he argues, govern our reproductive choices:

I claim that any outcome in which I create a new person S is subject to a moral standard (that of S's well-being), which is grounded in the existence of S, a being with moral status. This moral standard is satisfied if S has a life that is worth living, and failed if she has a life that is not worth living. ... If I am unable to give S a life that is worth living, I have a standard-regarding reason not to create S. (p. 73)

⁸ To be sure, this attitude is also common among those who are not old in the conventional sense, but instead find themselves at the end of life because of a terminal illness. For the sake of simplicity, however, I focus here on the seniors.

Loving relationships and important personal projects also seem to be subject to moral standards. These moral standards are satisfied when our valuable attachments are maintained in a flourishing state and not satisfied when they are neglected or destroyed. Thus, like bringing a child into this world while knowing that you will not be able to provide for it, taking on a new valuable engagement with the awareness that you will not be able to cultivate it is something you have a reason against.

A notable corollary of this is that we should take special care to secure for the later stages of life certain background conditions (health, wealth, etc.) essential to maintaining loving relationships and personal projects. If we fail to make such provisions, we may not only neglect the things which already contribute to our well-being, but also find ourselves unable to take on new valuable engagements.

Having said that, it is important to understand that Conservatism about Prudential Value does not imply that we should be completely passive and austere in older age. First, as noted earlier, conservative reasons are *pro tanto* reasons. Even at this stage of life, it seems possible to enter a loving relationship or start a project which would make a temporary contribution to our well-being so significant that it ought to be pursued even in the face of impending decline. Similarly, we may sometimes have overall reason to hold on to our existing attachments, even if we cannot maintain them for much longer.

Second, as we have also discovered, not all relationships and projects must be preserved indefinitely or for a very long time. Many attachments have a natural conclusion point: a summer romance, perhaps, or a project that will take a few months rather than a few years to complete. Conservative standards are easier to satisfy in such cases. Thus, strikingly, seniors have no excuse for not engaging in fleeting relationships and projects.

6. Moving on from a loss

The third puzzling issue that I want to address concerns loss and when it is appropriate to move on.

Suppose that, after many years of marriage, your spouse dies unexpectedly. While this relationship has come to an end, many people share the sentiment that you should abstain from entering a new loving relationship, at least for some time. Arguably, our important personal projects are similar: you have a moral reason to refrain from taking on a new engagement immediately after the one you are currently pursuing shatters, as when a person aiming to become the first to run a four-minute mile is beaten to this accomplishment.

These kinds of attitudes also extend to certain things which are valuable simpliciter, such as architectural marvels. Consider the example of the 18th-century Saxon Palace in downtown Warsaw, a former royal residence, which was destroyed by German troops in retaliation for the Warsaw Uprising of 1944. Close to eight

decades later, this space remains unutilised, and many citizens argue that it should never be, even if it could be replaced with a new structure which would be architecturally as impressive as the original and serve as an equally good reminder of the horrors of war.

What's common to these three cases is that something that was once non-instrumentally valuable but no longer exists continues to call for a certain kind of response from members of the moral community. In particular, it calls for preservation—though not so much of the valuable thing itself, but rather of the space it used to occupy in our lives. This suggests that the scope of Conservatism about Prudential Value is even broader than we might have assumed: we seem to have a moral reason to conserve not only certain *existing* valuable engagements but also some *past* valuable engagements that have ended prematurely.

That said, it is important to recognise that not all attachments have this profile. Suppose that you start a new friendship after the circumstances of life put an end to a previous one. Or imagine that a woodworking accident puts an end to your favourite pastime, climbing, and you fill the gap by taking up running. In both cases, a valuable engagement comes to an end through no fault of your own, and yet moving on does not strike us as disloyal or otherwise morally objectionable.

I think that we can explain this difference in the normative status of our valuable attachments by pointing to their disparate characters. Typically, romantic relationships and life-defining projects are *exclusive* in nature. This means that a new

romantic relationship or a new life-defining project is bound to, at least partly, take the place of the old engagement. By contrast, friendships and many lesser personal projects are *inclusive*, in the sense that they can be reconciled with each other.

This distinction between exclusive and inclusive engagements is especially natural in the synchronic context. For example, entering a new romantic relationship while the old one is still going and pursuing both at the same time often seems objectionable, whereas having multiple friends does not. Likewise, orienting your life towards a Hollywood career while still playing football at the highest level—as Jim Brown did—seems troubling in a way that combining passions for recreational climbing and running does not.

But I think that the same distinction also plays an important role in the diachronic context. In particular, when an attachment is exclusive and comes to an end prematurely, it calls for the preservation of the space that it used to occupy in our lives, and we have a conservative reason not to move on. The cases with which this section started are of this kind. On the other hand, when an inclusive engagement is terminated, we are free to press on with our lives and seek out new relationships and projects.

I believe that these considerations help us resolve a long-standing puzzle concerning loss. Return to the case of your spouse's unexpected death. It seems inappropriate to move on, say, just one day after the death of your loved one, but perhaps it is not inappropriate to do that a year or two later. Moreover, it does not

seem permissible to move on to just any relationship that would be good for you; the new relationship must be *good enough*. How can we account for these intuitions?

One potentially helpful idea is that conservative reasons diminish in strength over time and possibly even expire eventually. This kind of suggestion has been made in a closely related context: about reasons for grief. In particular, Krister Bykvist (2009, p. 16) argues that the fittingness of certain sentiments “depends on the ‘distance’ between oneself and the object”, where that includes “personal, modal, and temporal distance”.

However, as Berislav Marusic (2018, p. 12) points out, Bykvist’s view has implausible implications in cases like the following:

Suppose that you are in a car accident with your mother. You survive, but your mother does not. However, you spend years in a coma. When you wake up, you are informed that your mother has died. You are not relieved, and have no reason to be relieved, to hear the accident happened ten years ago. The mere passage of time makes no difference to your reasons.

The lesson of this example carries over to the present context. The strength of conservative reasons likewise does not seem to diminish with time.

A better explanation for why the appropriateness of moving on varies with time appeals to a curious feature of our psychology. C. S. Lewis (2001) describes it aptly in his reflections on the death of his wife:

I think I am beginning to understand why grief feels like suspense. It comes from the frustration of so many impulses that had become habitual. Thought after thought, feeling after feeling, action after action, had H. for their object. Now their target is gone. I keep on through habit fitting an arrow to the string, then I remember and have to lay the bow down. (p. 47)

As Lewis observes here, even after a relationship ends, it continues to occupy our attention and emotions. In fact, it often makes an even greater claim on our mental life than it used to when the other person was still around. Lewis notes that, too, in the same piece:

Part of every misery is, so to speak, the misery's shadow or reflection: the fact that you don't merely suffer but have to keep on thinking about the fact that you suffer. I not only live each endless day in grief, but live each day thinking about living each day in grief. (pp. 9-10)

These features of grief profoundly affect our ability to form and sustain new attachments. For most people, starting a new valuable relationship is impossible when they continue to act as if the other person is still present in their life and continue to think about their loss. In these circumstances, one cannot exhibit the kind of emotional vulnerability that is required for genuine love. We do, eventually, extricate ourselves from this state of mind. But this process takes time. For some people, it is a couple of months; for others, it might be a decade. Only then are we finally able to take on a new valuable engagement.

These considerations alone suffice to explain the first aspect of the puzzle: why moving on from a loss seems more appropriate with the passage of time. But we need to appeal to Conservatism about Prudential Value to explain the second aspect of the puzzle: that the new relationship must be ‘good enough’, not just ‘good’. Since conservative reasons do not expire with time, only sufficiently strong prudential reasons conferred by the prospective value of the new relationship will be able to outweigh them. That’s why not just any valuable relationship will do; it must be sufficiently valuable.

7. Prudence and morality: an interlude

In the preceding three sections, I have argued that Conservatism about Value can illuminate many otherwise puzzling intuitions and sentiments about starting,

preserving, and ending our loving relationships and important personal projects at various stages of life.

This discussion has a curious feature. Antecedently, these puzzles may have appeared to belong squarely in the domain of prudence. After all, they are all concerned with how to pursue, prioritise, and organise *prudential* goods over the course of a life. Accordingly, we might have expected to resolve them entirely within the prudential domain, perhaps by identifying some previously overlooked prudential reasons or by weighing prudential considerations in a novel way.

We have discovered something quite different, however. Our conduct towards our own relationships and projects is constrained, and must be informed, not just by prudential but also by moral reasons. This is so even when other people's interests are not at stake. Indeed, I believe that the issues discussed in the preceding sections are puzzling largely because they involve a previously unacknowledged conflict between our prudential and moral reasons. For example, while we continue to have some prudential reason to explore new relationships and projects even late in life, at that point, we typically have a stronger moral reason to commit to a select few.

Having said that, there is a way to bring all of these issues back into the domain of prudence while retaining the core of the conservative insight that we have a moral reason to preserve our loving relationships and important personal projects.

What we need is something to bridge the gap between what we have moral reasons and prudential reasons to do. One attractive way to do that is to adopt a less restrictive view about what makes our lives go well. As Derek Parfit (2016) points out, certain acts seem in themselves bad for the person who performs them, in the sense that they are in one way bad for them even when they have no further bad effects on their well-being. For example, your life seems to go worse when you deceive other people, coerce them, act unjustly, or break promises. As Parfit puts it:

Reviewing what we have made of our lives, in which our main projects have failed, we might be consoled by the thought that we have at least done some good, and that we haven't often acted in ways that were seriously wrong. These would be partial achievements, to set against our failures to achieve our other aims. On this view, if we are morally bad and act wrongly, these facts would not merely cause our lives to go worse, but would be in themselves ways in which our lives would go worse. (pp. 401-2)

Now, while Parfit's discussion focuses on the moral significance of treating *people* in specific ways, I believe that this is just one instance of a broader issue: the moral significance of responding appropriately to what is non-instrumentally valuable. Just as we fail to respect people when we deceive them, we fail to respond

fittingly to the value of relationships and personal projects—and it makes our lives go worse—when we neglect them. We can call this *the Bad Acts Thesis*.

Armed with the Bad Acts Thesis, we can recast the issues discussed earlier as involving conflicts between prudential reasons sourced in different aspects of well-being. To give just one example, in the case of the loss of a spouse, while we have some prudential reason find a new partner soon after, we might have a stronger *prudential* reason to abstain from entering a new relationship grounded in our interest to respond appropriately to the special value of the previous relationship.

The Bad Acts Thesis is bound to strike some people as controversial. For what it's worth, the reader is free to decide for themselves if they are willing to accept it and, thus, which way of hashing out the puzzles discussed in the previous sections they prefer. My principal aim has been to show that Conservatism about Prudential Value can elucidate these puzzles, and that remains true regardless of whether we take this view to give us just moral reasons or also, indirectly, prudential reasons.

Having said that, there is a significant advantage to accepting the Bad Acts Thesis. If we do that, we can appeal to Conservatism about Prudential Value to resolve a long-standing puzzle concerning the evaluative significance of a life's shape. The next section explains this.

8. The significance of a life's shape

After examining youth, adulthood, and older age in isolation, let's now turn to a puzzle about life considered as a whole. To get a grip on this puzzle, consider two ways in which your life might go.

Improvement: Your life begins in the depths and has an upward trend: misery in the early years, mixed fortune in the prime of your life, followed by flourishing in older age.

Deterioration: Your life begins at the heights and has a downward trend: an early period of flourishing, mixed fortune in the prime of your life, followed by misery in older age.

In describing these cases, I have used vague expressions like 'hardship' and 'flourishing' because the puzzle I am about to describe arises for a wide range of theories of well-being. As noted in the introduction, however, I am assuming an objective list theory of well-being for the sake of concreteness. We should therefore think of an upward trend in *Improvement* as corresponding to an improvement in one or more of the objective goods, and conversely for *Deterioration*.

How do these lives compare? When confronted with this choice, many people have the intuition that the improving life is better. So, for example, if you

had to choose one of these lives for your child or a loved one, you should choose the former.⁹ But this judgment is puzzling. Notice that these two lives are permutations of each other—in the sense that for every period of flourishing in one life, there is exactly one such period in the other, and so on. This, on one kind of view, strongly suggests that *Improvement* and *Deterioration* are equally good.

My aim in this section is to resolve this puzzle by developing a theoretical rationale for the intuitive judgment that *Improvement* is indeed better. This rationale is rooted in Conservatism about Prudential Value and the Bad Acts Thesis.

To get a grip on this solution, consider the following more specific cases that feature contrasting patterns of temporal well-being.

Lost Love: In his youth, Alex has a number of highly fulfilling relationships.

Over the years, however, he lets these attachments crumble, and they are replaced by other, more superficial relationships.

Found Love: Alex's youthful relationships are relatively superficial, but over the years, they deepen, and he continues to make new, highly meaningful connections.

⁹ This intuition is recorded in the literature by Slote (1983), Bigelow et al. (1990), Velleman (1991), Kamm (1998; 2003), Portmore (2007), Temkin (2012), Kauppinen (2012), Glasgow (2013), Dorsey (2015), and Hirose (2015).

Let's assume that these two lives are otherwise very similar: for example, Alex does not realise that his relationships are getting worse in the first case, or, if he does, this realization does not cause him any distress. Even with these assumptions, I believe, Found Love is better than Lost Love.

A natural explanation of this judgment draws on Conservatism about Prudential Value. In Lost Love, Alex fails to preserve existing valuable relationships (and they are replaced with others which are also less good), which—according to the Bad Acts Thesis—is in itself bad for him. This is not so in Found Love. Thus, even if Alex's life ultimately contains the same relationships in either case, Lost Love is worse for him than Found Love.

An analogous gloss can be given for the comparison between two lives which differ in terms of when one's greatest achievements occur.

Early Breakthrough: Ada takes up a career in academic philosophy. She writes her most original and insightful article shortly after finishing her graduate studies. Her subsequent publications continue to make valuable contributions, but their quality steadily diminishes over time.

Late Breakthrough: Ada begins her career in academic philosophy with a relatively pedestrian article, but the originality and insight of her work steadily improve over time.

Here, as before, there is something tragic about a situation in which great accomplishments at the start of one's life are not replicated further down the road. I believe that Conservatism about Prudential Value is uniquely positioned to explain this, in the following way. When Ada sets her sights on a career in academic philosophy, she takes on a valuable personal project. Her plan is presumably not to write just one outstanding article, but rather to make significant scholarly contributions throughout her career. As noted earlier, our important personal projects must be preserved in a flourishing state. Ada apparently fails to do that in Early Breakthrough since the quality of her papers deteriorates over time, which is in itself bad for her. By contrast, this is not the case in Late Breakthrough. So, Ada's life is worse in Early Breakthrough than in Late Breakthrough, even if these scenarios ultimately feature the same achievements.

These considerations lead me to propose the following account. *Improvement* is better than *Deterioration* just in case and because these lives differ in terms of whether certain prudential goods, such as loving relationships and important personal projects, have been conserved over time. In cases of *Deterioration*, the extent of the failure to conserve these things is typically greater than in cases of *Improvement*,

and such failures are intrinsically bad. So, *Improvement* is typically better than *Decline*. Call this *the Conservative Account*.

Two sets of considerations will help us appreciate the nuance and strength of this solution. First, I would like to stress that although the above examples are stylized, they are not cherry-picked. The same feature—the failure to conserve certain prudential goods—is at least implicit in all key examples discussed in the literature on this topic. For example, Michael Slote (1983), to whom the contemporary debate over the significance of a life’s shape can be traced, centres his discussion on the following case that involves contrasting achievement patterns.

Political Careers: “A given man may achieve political power and, once in power, do things of great value, after having been in the political wilderness throughout his earlier career. He may later die while still ‘in harness’ and fully possessed of his power, at a decent old age. By contrast, another man may have meteoric success in youth, attaining the same office as the first man and also achieving much good; but then lose power, while still young, never to regain it. Without hearing anything more, I think our natural, immediate reaction to these examples would be that the first man was the

more fortunate, and this seems to suggest a time preference for goods that come late in life.” (pp. 23-4)¹⁰

I will return to Slote’s preferred diagnosis in a moment. For now, let me just note that the Conservative Account offers an alternative explanation of the advertised intuition: the description of the second politician’s life suggests that he failed to adequately cultivate his political career, which was in itself bad for him.

Second, I believe the Conservative Account compares favourably with three popular views offered in the literature: *the Temporal Location Account*, *the Pattern Account*, and *the Narrative Account*. Let’s consider these in turn.

According to Slote’s (1983) Temporal Location Account, an improving life is better than a deteriorating life because what happens later in life has greater weight in determining the life’s overall value. In other words, any given prudential good will contribute more to the goodness of a life if it occurs in one’s adulthood or older age rather than youth. That is why an improving life, in which the best things happen late in life, is better than a deteriorating life.

The Pattern Account offers a different take. On this view, the fact that your life features an improvement is itself a source of well-being (and that your life

¹⁰ Velleman (1991) also discusses this case in his influential article on the topic. So does Dorsey (2015), who supplements it with a similar example involving O. J. Simpson (who started as a successful football player and ended as a convict) and J. O. Nospmis (a fictional character whose life proceeded in the reverse order).

features a deterioration is a source of ill-being), quite independently of any attitudes you might have towards these trends and their impact on your well-being. This view has been endorsed by the likes of Frances Kamm (2003), Larry Temkin (2012), Joshua Glasgow (2013), and Iwao Hirose (2015).

An undeniable attraction of these two views is their simplicity. But it is also the source of their main flaw: they wrongly predict that an improving life is *always* better. Let's take a moment to explain why this is a vice, and not a virtue, of these accounts. To begin with, consider the following case due to Dale Dorsey (2015, p. 319):

Hedonic Experiments: “[C]ompare two different experiments. The first person is hooked up to a pleasure-stimulating computer with software designed to start his life at a neutral level: no pleasure, no pain. Gradually, say, twice a year, the pleasure is increased in a linear fashion. The second is precisely the opposite: the person’s machine starts out with quite a lot of pleasure. Gradually, also twice per year, the software decreases the pleasure in a linear fashion, such that both people, over the course of their lives, will generate the same amount of sensory pleasure.”

Many people find it implausible that an improving life is better than a deteriorating life in this scenario, and I share that intuition. Moreover, the shape of

a life does not seem to matter when it aligns with the character of one's engagements. For example, the life of a philosopher who does their best work later in life does not appear to be automatically better than that of a typical athlete who peaks in their late twenties. The Temporal Location Account and the Pattern Account cannot accommodate these intuitions.

By contrast, the Conservative Account gets these cases right because it recognises that what we have a reason to do depends on the nature of the valuable object. As we have seen, hedonic goods are merely to be savoured while they last, so it does not matter whether we experience greater pleasures earlier or later in life. Moreover, an athletic career has an early natural conclusion point marked by an impending decline in physical ability, which makes participation at the elite level impossible, so there is no conservative reason to cultivate it beyond that point.

In addition to demonstrating that the Temporal Location Account and the Pattern Account are inadequate, these examples carry a more general lesson. The original intuition that might have had, that *Improvement* is better than *Deterioration* regardless of what happens in these two lives, turns out to be incorrect. The refined intuition is that an improving life is typically better than a deteriorating life, and this is what we must ultimately explain.

Turn now to the third prominent account offered in the literature: the Narrative Account. It holds that *Improvement* is better than *Deterioration* when and because these lives differ in terms of certain narrative relations between discrete

events, such as when one learns from a past misfortune or when an eventual success renders a particular effort a good investment rather than a waste. David Velleman (1991) defends this account, and its variants have been endorsed by Jeff McMahan (2002), Douglas Portmore (2007), Antti Kauppinen (2012), and Dale Dorsey (2015).

The Narrative Account is closer to the truth in that it recognises, first, that an improving life is only typically better than a deteriorating life and, second, that to determine whether it is, we have to look at relations between discrete events in one's life. Still, it appears to be wanting in at least three respects.

First, the Narrative Account cannot accommodate our intuitions in many cases, such as the contrast between Found and Lost Love or between Late and Early Breakthrough. These scenarios do not differ—or at least need not differ—in terms of whether the agent learned from their past mistake or their efforts turned into a success. To be sure, learning from past mistakes and turning struggles into a success do have something in common: for example, they are aggregates of discrete events in a person's life or a course of history. But not all such aggregates of events have evaluative significance. For example, the badness of the pain I experienced when I stubbed my toe yesterday is not cancelled by the fact that I had a delicious kale salad for lunch today. So there must be some further, more specific feature that makes learning from past mistakes and turning struggles into success evaluatively significant. But it is at least unclear from Velleman's discussion what this feature is. This is in stark contrast to the Conservative Account, which identifies a single

evaluatively significant feature in all relevant comparisons: whether certain prudential goods have been conserved over time.

Finally, the scenarios used by proponents of the Narrative Account to motivate their view often lend themselves to more natural elucidations via the Conservative Account. For example, Velleman (1991, p. 55) extensively discusses the following case.

In one life your first ten years of marriage are troubled and end in divorce, but you immediately remarry happily; in another life the troubled years of the first marriage lead to eventual happiness as the relationship matures.

Velleman believes that the second life would be better and claims that this points to the importance of learning from past mistakes. However, when I reflect on this case, the feature that stands out to me is that the first loving (if troubled) relationship has been conserved, whereas the second has been simply abandoned and replaced.

Overall, these considerations lead me to conclude that the Conservative Account provides a better explanation of the significance of a life's shape than the Temporal Location Account, the Pattern Account, and the Narrative Account.

9. Concluding remarks

If Conservatism about Prudential Value is true, we cannot start a new life whenever we please. Our existing loving relationships and important personal projects warrant a certain kind of loyalty, in virtue of which we have a moral reason to conserve them.

As we have seen, this view best explains the evaluative ambivalence we witness in cases of replacing loving relationships and important personal projects, and it vindicates a range of our otherwise puzzling intuitions and sentiments. These include: that it is permissible to cycle through many different projects and relationships in youth, but not at later stages of life; that the reluctance to start new engagements at the end of life is often justified; that we should not move on from a loss of a loved one too quickly and to a relationship that is just barely good for us; and that an improving life is typically better than a deteriorating life even if these lives contain many of the same things.

Chapter 4: Extension and Replacement

1. Introduction

As Gustaf Arrhenius (2008, p. 211) observes,

It seems to be a widespread opinion that increasing the length of existing happy lives is better than creating new happy lives although the total welfare is the same in both cases, and that it may be better even when the total welfare is lower in the outcome with extended life.

Despite the popularity of this view, one would be hard-pressed to find a fully compelling justification for it in the literature. Arrhenius himself considers three candidates, but eventually rejects all of them. In this essay, I want to propose a novel account of why and when extension is better than replacement which appeals to a broadly conservative view about value.

Before I review a handful of rival justifications and introduce the details of my account, I would like to clarify three aspects of my discussion. First, I will understand the claim that extension is *better* than replacement in the wider, reason-implicating sense rather than in the narrower, axiological sense. That is, I take the relevant sentiment to be that we have a weighty moral reason to increase the length

of existing happy lives rather than create new happy lives, although the total welfare is the same in both cases. This may be because the former state of affairs has overall more value than the latter, but it need not be. Indeed, as I will argue, we often have a moral reason to choose extension over replacement, even if the latter would result in a more valuable state of affairs.

Second, while Arrhenius focuses on persons, I want to include other kinds of entities in my discussion. It seems to be an equally common view that extending *humanity's* tenure would be better than allowing another intelligent and sentient species to replace us. Likewise, we should reflect on whether extending the life of a happy *non-human animal* (hereafter just 'animal') is better than replacing it with another happy animal. (Opinions on this latter issue are admittedly more divided.) To my knowledge, these three questions have not been discussed together. This strikes me as a missed opportunity. As I propose here, there is a common ground for an important class of our moral reasons to favour extension over replacement with respect to animals, persons, and humanity. I will first derive this ground in the context of humanity as a whole, where the case for extension is arguably the most intuitive, and then expand my account to cover persons and animals.

Third, my aim in this paper is to identify *one important source* of our moral reasons to choose extension over replacement. It is not, however, to identify a *unique source* of such reasons. I choose to restrict my aim in this way mainly because concrete cases depicting the choice between extension and replacement typically

have multiple features that could potentially ground a moral reason for extension. We could, of course, try to rule out many of these features by constructing increasingly intricate vignettes, but this approach runs the risk of foregoing tractability. Instead, I propose that we take the presented cases to illustrate the kinds of situations to which the proposed account applies, assess how common or rare these situations are, and in that way, gauge this account's explanatory power.

With these clarifications in place, my central question takes the following form:

Extension vs Replacement: What, if anything, grounds our moral reasons to choose extension over replacement with respect to non-human animals, persons, and humanity as a whole, when the total welfare is the same in both cases?

The answer I propose appeals to a broadly conservative view about value. This view emphasises the moral importance of preserving certain non-instrumentally valuable elements of humanity's culture (such as beautiful artworks, cherished traditions, just institutions, bodies of scientific knowledge, and relations of equality) and of individuals' lives (such as loving relationships and important personal projects). In essence, my account holds that we have a moral reason to choose the extension of the life of an animal, person, or humanity over its

replacement when and because this is necessary and sufficient for the preservation of those non-instrumentally valuable things.

Let me say three things about this view, in an effort to anticipate some natural questions. First, the proposed account implies that we *typically* have a moral reason to choose extension over replacement with respect to persons and humanity, but it does not imply that we *always* do. This, I believe, aligns with our intuitions. This account also goes some way towards explaining why people's opinions about the same choice regarding animals are divided. As I will explain in Section 4, it turns out that whether we have a conservative reason to extend animal lives depends on a largely empirical premise (namely, whether animal lives feature loving relationships and important personal projects which warrant preservation), and folk intuitions about this premise are divided.

Second, my answer has a *reductive* character. In my view, our conservative reasons to favour extension over replacement are grounded in the importance of preserving the valuable contents of an individual's life or the valuable elements of humanity's culture. They are not grounded in the importance of preserving life or the human species themselves, as a *non-reductive* account would have it. As I will explain in Section 5, while perhaps strange at first blush, the reductive account has two main advantages. For one thing, it's more parsimonious: to explain why we typically have a moral reason to choose extension over replacement, if we already believe that certain prudential goods and cultural artefacts are worth preserving,

there is simply no need to postulate that, in addition, lives or the human species themselves are worth preserving. For another, this account better captures our considered judgments. I submit that there is no conservative reason to favour extension over replacement when all of the valuable elements of an individual's life or humanity's culture would be eradicated. In a slogan, what matters is the *continuity* of life rather than its mere *continuation*.

Finally, I would like to review a handful of views in population ethics which might seem to provide straightforward explanations of when and why extension is better than replacement. Since many of the criticisms that I want to mention have been discussed extensively in the literature, I will be very brief.¹

Consider, first, two simple impersonal views. According to the average view, one outcome is better than another just in case it features higher average welfare. According to the critical level view, one outcome is better than another just in case it features higher total contributive welfare, where each individual's contributive value equals their welfare minus some positive constant representing the critical level. Both of these views imply that it is better to extend the life of a person or an animal rather than replace them with another individual because they prefer welfare to be spread among as few individuals as possible.

¹ For overviews of these views and criticisms, see Arrhenius (2008) and Greaves (2017).

Alternatively, we might want to appeal to a person-affecting view. There are several ways in which such views can be spelled out—in terms of conditional goodness, harm-minimisation, or an asymmetry between comparative and non-comparative benefits and harms—but the animating idea appears to be that it's better to extend the life of an individual rather than allow for their replacement because the former option is better for someone and worse for no one.²

I think that none of these views provides a fully compelling justification of the intuition that extension is better than replacement. First, they all have well-known troubling implications for other issues in population ethics. For example, the average view and the critical level view entail the sadistic conclusion that it can be better to add to a population individuals with negative welfare rather than individuals with positive welfare. Moreover, person-affecting views struggle to explain why we have a moral reason not to create individuals with bad lives.³ Second, and this is more important in the present context, these views cannot provide a *unified* answer to my question. This is because they fail to explain why it is better to extend the tenure of humanity rather than allow its replacement with another intelligent and sentient species in cases when the continued existence of any particular individual is not at stake. Third, even if one of these views proves to

² See, for example, McMahan (2013) and Frick (2020).

³ Though see Frick (2020) and Bader (2022) for recent attempts to accommodate this intuition within the person-affecting framework.

correctly identify *a* moral reason to choose extension over replacement, this might not be the only or the most important reason that we have in this context. Indeed, the ensuing discussion seeks to show that an important class of our moral reasons to favour extension over replacement has been largely overlooked.

2. Humanity

What, if anything, grounds our moral reason to choose extension over replacement with respect to *humanity* as a whole when the total welfare is the same in both cases?

To get a better grip on this question, consider the following example. Suppose that you have two options. If you do X, humanity will continue to exist for another 500,000 years. If you instead do Y, humanity will go extinct, but another intelligent and sentient species will take its place and remain in existence for 500,000 years. Assume that these two futures would not differ in terms of total welfare, that doing Y does not amount to killing anyone, and that no one would die prematurely. Perhaps humans have recently become infertile, and you can either remedy this or create conditions under which another intelligent and sentient species would evolve on or settle our planet once all traces of our civilisation have disappeared. What, if anything, grounds your moral reason to do X rather than Y in this situation?

This question has received little philosophical attention in the literature.⁴ But there is another, related question which has been discussed more widely:

⁴ For one exception, see Williams (2006).

Extension and Extinction: What, if anything, grounds our moral reason to extend the length of humanity's tenure rather than to let it go extinct unreplaced?⁵

For example, in his recent book *Why Worry About Future Generations?* (2018), Samuel Scheffler argues that among the reasons to prevent the extinction of humanity are what we can call 'conservative reasons'. He writes:

All of the many things we value that consist in or depend on forms of human activity will be lost when human beings become extinct. No more beautiful singing or graceful dancing or intimate friendship or warm family celebrations or hilarious jokes or gestures of kindness or displays of solidarity. Other things that we value—physical artifacts, for example—may survive for a while, but with no one to appreciate their value, for in addition to the disappearance of *valuable things*, the extinction of the human race will mean the disappearance of *valuing* from the Earth. (pp. 69-70)

⁵ Henceforth, 'extinction' will refer to 'extinction without replacement'.

In this passage, Scheffler sketches an ingenious account of one important class of moral reasons to favour the extension of humanity's tenure over its extinction. It can be reconstructed as follows. We have a moral reason to favour extension over extinction because humanity's survival is necessary for the continued existence of valuing and certain valuable elements of humanity's culture, and we have an antecedent moral reason to preserve these things.

It is worth noting at the outset that while Scheffler's remarks focus on the *necessity* of the survival of humanity for the preservation of valuable things and valuing, an additional element is needed. Specifically, in any given set of circumstances, the survival of humanity must also be *sufficient* for preserving valuing and some valuable elements of humanity's culture are preserved. If the disappearance of valuing and the valuable elements of our culture was inevitable, there would be nothing to ground a conservative reason to prevent the extinction of humanity.

With this small amendment in place, let's consider: can the same kinds of considerations be used to support our moral reason to choose the extension of humanity's tenure over the replacement of humanity with another intelligent and sentient species?

There are two main grounds for hesitation. To begin with, consider valuing. On Scheffler's own view developed elsewhere, valuing X involves four elements: (i) a belief that X is valuable; (ii) a susceptibility to experience a range of context-

dependent emotions regarding X; (iii) a disposition to experience these emotions as being appropriate; and (iv) a disposition to treat certain X-related considerations as reasons for action in relevant deliberative contexts (2010, p. 29). Based on this description, there seems to be nothing distinctively human about valuing that could not, or would not, be replicated in another intelligent and sentient species. Consequently, it looks like our concern for the continued existence of valuing cannot support a concern for the continued existence of humanity in particular, as opposed to the continued existence of intelligent and sentient life in general.

What about the valuable elements of humanity's culture, such as beautiful singing, intimate friendship, and displays of solidarity? One might likewise suggest that these things would not disappear altogether if humanity was replaced by another intelligent and sentient species. There is no principled reason to think that members of that species would not be capable of creating and performing music, establishing intimate connections with one another, and displaying virtues such as solidarity. Thus, it may look like our concern for the continued existence of the valuable elements of humanity's culture cannot support a moral reason to extend humanity's tenure when replacement is the alternative.

But this second strand of scepticism rests on a misunderstanding of the conservative view about value that underlies Scheffler's account of our moral reasons to prevent extinction. On this view—also espoused by the likes of G. A. Cohen (2012), Johann Frick (2017), and Jacob Nebel (2015)—we have a moral

reason to preserve certain valuable things even when they could be replaced by similar things of equal or greater value. Crucially, what we have a reason to preserve is not the existence of valuable things *in general*, or even certain *types* of valuable things, but rather *tokens* of valuable things. For example, we have a reason to preserve the iconic Golden Gate Bridge even if it could be replaced by another, even more impressive structure. This is important for thinking about the choice between extension and replacement. If another intelligent and sentient species were to replace us once all traces of our civilisation have disappeared, as we are asked to imagine, tokens of many valuable things present nowadays—*human* forms of singing, friendship, and solidarity—would not be preserved. Therefore, our concern for the continued existence of such things can, in fact, support a moral reason for extending humanity's tenure when the alternative is replacement with another intelligent and sentient species.

With these considerations in place, the conservative answer to my central question, as it concerns humanity as a whole, takes the following shape. We have a moral reason to extend the length of humanity's existence rather than allow its replacement with another intelligent and sentient species when and because this is necessary and sufficient for preserving certain valuable elements of our culture. These include distinctively human forms of singing, friendship, or solidarity, but also tokens of other non-instrumentally valuable things, such as beautiful artworks,

sublime landscapes, unique languages, or relations of equality between members of the moral community.

Let's take a moment to emphasise two features of this view. First, conservative reasons are *pro tanto* reasons and must be weighed against other relevant moral considerations. For example, if the replacement of humanity with another intelligent and sentient species would eventuate in a state of affairs that would be vastly more valuable, then this might be what we should all-things-considered choose.

Second, conservative reasons have *varying strengths*. This strength is plausibly a function of both the quantity and the value of the things that are such that humanity's survival is necessary and sufficient for their preservation. One corollary of this is that our conservative reason to prevent extinction will likely be stronger than our conservative reason to prevent replacement. This is because, as I have suggested above, at least one of the things that are worth preserving—valuing itself—would disappear in the case of extinction but would not (or at least need not) disappear in the case of replacement.

Return to the main thread of the argument. The preceding discussion explained *why* we sometimes have a moral reason to choose extension over replacement. Now, let's examine *when* that's the case. To that end, it is helpful to single out three conditions which must be satisfied for such reasons to arise: *presence* (of non-instrumentally valuable things which warrant further preservation), *necessity*,

and *sufficiency*. I think that these conditions are satisfied in most practically-relevant cases, but there are some noteworthy exceptions. Let me mention one example for each condition.

Take ‘presence’ first. As things stand, there is an abundance of non-instrumentally valuable things that warrant further preservation: thousands of beautiful artworks, sublime landscapes, unique languages, important traditions, collective endeavours, and other cultural artefacts. While some of them inevitably perish every year, the total number of these things keeps increasing, as a seemingly necessary by-product of human activity. But one day, all of this might change. A global nuclear conflict or a natural disaster of an unprecedented scale could reduce our planet and the fruits of our civilisation to rubble and dust.⁶ If humanity were to survive a catastrophe of this kind and find itself completely deprived of valuable things that warrant further preservation, there would be no conservative reason to choose extension over replacement (or extinction, for that matter). That is, even if we were sufficiently prepared (or lucky!) and had the opportunity to build a new world on top of the ruins of the old one, in the absence of genuine continuity of the human civilisation, extension would not be preferable to replacement from the conservative point of view.

⁶ For recent illuminating discussions of how civilisational collapse could come about and what we could do to prevent it, see Ord (2020) and MacAskill (2022).

Turn to ‘necessity’ now. For all we know, humanity is alone in our little corner of the universe. No one will rescue us if we get into trouble, and no one will take care of our business once we’re gone. But suppose, for the moment, that were not true. Perhaps an intelligent and sentient alien species has been observing us from afar and developed an appreciation for our ways of life. As long as things are going sufficiently well for us, these aliens will not interfere. But in the event of human extinction, their emissaries will rush to Earth and seamlessly take custody of as many valuable things as possible. They will safeguard our art, learn and practice our languages and traditions, and work hard to advance our collective endeavours. If that were to happen and nothing of value would be lost in this transition, there would be no conservative reason for extension over replacement either.⁷

Finally, consider ‘sufficiency’. We can often expect humanity’s survival to be sufficient for the preservation of many valuable elements of our culture. Some valuable elements of our culture are straightforward to preserve. Languages and traditions, for example, just need to be practised by people, and no special incentives are usually necessary for them to do that. In other cases, we already have elaborate systems in place: museums and galleries that house our artworks and whole

⁷ Even if a few valuable elements would be lost in the transition (as they would be, perhaps, if upon all philosophers’ sudden disappearance, philosophically-minded economists, physicists, and sociologists took over our positions and research agendas), the conservative reason for extension over replacement would be quite weak.

professions dedicated to maintaining them in good condition and on display for people to appreciate.

But there are outliers. Imagine a scenario in which humanity faces a global catastrophe, and the only way for our species to survive is to forego everything of value. Perhaps a small number of people could find shelter in an underground bunker, where they would live in primitive conditions while everything on our planet's surface would perish. If truly none of the valuable elements of our culture would remain in this world, then there would be no conservative reason to extend humanity's tenure rather than allow for our replacement. This is so even if the surviving population could at a later time, when the conditions on the surface have sufficiently improved, create completely new cultural artefacts.⁸

This concludes the first part of our discussion. We can see that while we typically have a conservative moral reason to choose extension over replacement, there are exceptions. This strikes me as intuitively correct. I believe that what we ultimately care about—at least from the conservative point of view—is not the preservation of the human species itself, but rather the preservation of the valuable elements of our civilisation. These two things usually go together, but they can also come apart. Our species can survive even if our civilisation crumbles, and (as the

⁸ I will return to this issue in Section 5.

scenario involving alien custodians is meant to illustrate) the fruits of our civilisation might survive even if our species goes extinct.

It is also worth noting that whether there is a conservative reason to choose extension over replacement is largely independent of welfare considerations. In particular, we could have a conservative reason to choose extension even if replacement would result in a state of affairs that would be more valuable. In fact, we could have a conservative reason for extending humanity's tenure even if it would result in a state of affairs with negative total welfare (for example, if future people suffered from some painful condition, but nonetheless preserved many valuable elements of our culture). That said, it is important to keep in mind that conservative reasons are merely *pro tanto* reasons. Even if they support extension in such cases, they do not guarantee that this is what we should all-things-considered choose.

3. Persons

Turn now to the second part of my central question: what, if anything, grounds our moral reason to choose extension over replacement with respect to *persons* when the total welfare is the same in both cases?

This choice can be illustrated as follows. You have two options. If you do X, Ada will continue to exist for another 50 years. If you instead do Y, Ada will die, but Adam will come into existence and live for 50 years. Assume that Ada's extra

50 years and Adam's 50-years-long life would not differ in terms of welfare, and that doing Y does not amount to killing or failing to rescue Ada. Perhaps you must allocate one unit of some scarce drug, and that drug will eventually either extend Ada's life or else cause Adam to exist. What, if anything, grounds your moral reason to do X rather than Y in this situation?

I believe that, in this context as well, the answer has to do with the importance of preserving valuable things. In particular, we have a moral reason to extend a person's life rather than create another life when and because this is necessary and sufficient to preserve certain valuable things.

What are the relevant valuable things? On the one hand, there are things which are *good simpliciter*. These are the things we discussed in the previous section, such as beautiful artworks, important traditions, and relations of equality between members of the moral community. In some cases, the continued existence of a particular person will be both necessary and sufficient for the preservation of such things. For example, we could imagine that preserving a magnificent but crumbling cathedral depends on the continued oversight of a renovation project by a particular architect with unique expertise. Likewise, the maintenance of peace in some region could require the continued tenure of a particular, extraordinarily skilful politician. We would have moral reasons to ensure the continued existence of these individuals, rooted in the concern for preserving valuable things. That said, such cases will be rare.

More commonly, our moral reasons to extend a person's life will be grounded in the importance of preserving *prudential goods*: things which are valuable for particular people. While discussions of conservatism about value in the literature have focused almost exclusively on things which are good simpliciter, I have argued in other work that we also have a distinctive moral reason to preserve certain prudential goods, such as loving relationships and important personal projects, even when a superior replacement is available.⁹

Two main examples motivate this idea. The first concerns Jim Brown, a former fullback for the Cleveland Browns. In 1966, Brown shocked the football world by announcing his retirement from the sport at the young age of 29. His decision was not motivated by injury or a decline in performance. In fact, he had just won the NFL Most Valuable Player award for the third time. Instead, Brown left football to pursue another career, acting, where he also enjoyed great success. He became the first Black action movie star, appearing in over 50 movies and TV shows, many of which were blockbusters.

The second example is fictional. Suppose that Ann is in a loving relationship with Beth. One day, Ann meets Sally, who happens to be very similar to Beth in all respects except one. Like Ann, Sally is a dedicated runner, so they could enjoy their

⁹ See Chapter 3 of this dissertation.

training together and support each other during races, bringing them even closer. After some deliberation, Ann leaves the relationship with Beth for Sally.

What should we think about these cases? If you are like most people, Brown's decision seems, in one respect, regrettable. Crucially, this is not just because he missed out on the many athletic accomplishments that lay ahead of him, but also because of what he had left behind. He quit a career that was a non-instrumentally valuable personal project that he gave up much for and that still had plenty to give back to him. Likewise, there is something troubling about Ann's decision. That old relationship was non-instrumentally valuable and would have continued to be so, for all we know. In other terms, despite the presence of an equally good or even better replacement, both choices leave us with considerable normative residue. This normative residue, I believe, indicates that we have a distinctive moral reason to conserve important personal projects and loving relationships—just as we have a distinctive moral reason to conserve beautiful paintings, important traditions, or relations of equality.¹⁰

If this is right, then the conservative account can potentially extend to a much wider range of cases than we might have originally assumed. Few people are indispensable for the preservation of things which are valuable simpliciter, such as

¹⁰ It is worth emphasising that this reason is of moral, not prudential, variety. After all, Brown had a reason to preserve his first career, and Ann had a reason to preserve her first relationship even though their new engagements would have been at least as good for them.

architectural marvels or peaceful relations in any given region. But almost everyone has some loving relationships or important personal projects, and everyone who does is indispensable for their preservation. Thus, it's almost always the case that we have a conservative reason to choose extension over replacement with respect to persons.

There are exceptions, however. Each of the three conditions—presence, necessity, and sufficiency—can be frustrated. I will give one example for each.

First, in some instances, the kinds of goods which warrant preservation were once present in a person's life but no longer are. This is sometimes true of elderly people: they could have already lost all their loved ones, either completed or abandoned their important personal projects, and severed their connections with valuable cultural artefacts. In such cases, there is no relevant ground for a conservative reason to choose extension over replacement.

Second, turn to the necessity condition. As suggested earlier, this condition is trivially satisfied with respect to prudential goods. The fruits of human civilisation could perhaps survive without the human species, but one's relationships and personal projects could not survive without them. However, the necessity condition is not trivially satisfied with respect to things which are good simpliciter. Consider again the example of an architect who oversees the efforts to preserve a magnificent but crumbling cathedral. If there is another expert who would seamlessly take over

this role in the event of the first architect's death, then there is no corresponding ground for a conservative reason to choose extension over replacement either.

Finally, consider sufficiency. We have previously discussed a scenario in which humanity can only survive in some diminished state or at the expense of the fruits of our civilisation. There is a familiar analogue in the present context. When people describe their experience with chemotherapy or other aggressive forms of cancer treatment, they often speak about fatigue, brain fog, mood changes, stress, and loss of sexual function. These side effects might seem relatively insignificant when considered in isolation, but they tend to carry further deleterious consequences, especially when they occur in combination. For instance, they might prevent a person from engaging in fulfilling social interactions, valuing their professional and athletic goals, or maintaining serious intellectual engagements. That is, as a result of undergoing these forms of treatment, a person could lose all of their loving relationships and important personal projects. If that were the case, there would be no conservative reason to choose extension over replacement.

To summarise this part of the discussion, the conservative account implies that we *almost always* have a conservative moral reason to extend a person's life rather than to create a new life. The only exceptions are cases in which nothing of value would be preserved alongside the person's life.

Crucially, these conservative reasons are *agent-neutral*. We have them regardless of whether we are personally involved in the relationships or important

projects of the person whose life we are in a position to extend. In this respect, preserving a life which features loving relationships and important personal projects is like preserving the Grand Canyon or the Mona Lisa.

4. Animals

Turn now to the third and last entity on our list, non-human animals. We can consider a similar choice in this context. You have two options. If you do X, a cow named Moo will continue to exist for another 5 years. If you instead do Y, Moo will die, but another cow named MooTwo will come into existence and live for 5 years. Assume that Moo's extra 5 years and MooTwo's 5-years-long life would not differ in terms of total welfare. Suppose also that doing Y does not amount to killing or failing to rescue Moo. Perhaps you must allocate one unit of a scarce drug, and that drug will either eventually extend Moo's life or else cause MooTwo to exist. What, if anything, grounds your moral reason to do X rather than Y in this situation?

In my experience, people's intuitions about this case are more divided than about the previously considered cases. Some people, including me, believe that there is a moral reason to extend the lives of at least some animals—maybe not flies or shrimp, but certainly cows, dogs, and chimpanzees. Others, including many of those who accept the practice of 'humane omnivorism', believe otherwise.¹¹

¹¹ Humane omnivorism is the practice of raising animals in humane conditions, killing them for consumption, and replacing them with individuals, which have lives that are at least as

What does the conservative account say about this issue? As before, we have a moral reason to choose the extension of an animal's life over its replacement with another animal when and because this is necessary and sufficient to preserve certain valuable things. Needless to say, this is a conditional claim, so we need to examine when, if ever, the antecedent holds true.

Start with things which are valuable simpliciter, such as beautiful artworks or relations of equality between people. There are some possible cases in which a particular animal's continued existence is essential to the continued existence of these things. Perhaps if a dog named Fido dies, then its architect-owner will fall into the grips of depression, and the magnificent cathedral whose renovation he has been overseeing will crumble. But this kind of dependence is of limited philosophical interest in the present context and such cases would be extremely rare anyway.

What about prudential goods? Take loving relationships. Empirical research on animal behaviour strongly suggests that many animals develop lasting ties with each other (for example, orangutan mothers stay with their offspring for eight to ten years), experience emotions characteristic of loving relationships (such as grief), and are willing to make sacrifices when others of their own kind are at risk of harm.¹²

good. For overviews of arguments in favour and against this practice, see McMahan (2008) and Delon (2016).

¹² For an informative overview of some of this empirical work, see Gruen (2021).

The issue, however, is whether these bonds count as ‘genuine loving relationships’. Presumably, only these kinds of relationships warrant preservation. By contrast, mere acquaintances and fleeting connections do not seem to have this profile.

According to one popular view, defended by Niko Kolodny (2003, p. 150),

Love is a kind of valuing. Valuing X, in general, involves (i) being vulnerable to experience certain emotions regarding X, and (ii) believing that one has reasons both for this vulnerability to X and for actions regarding X.

Some animals satisfy the first condition, as evidenced by their displays of emotions such as grief. What’s more controversial is whether they satisfy the second condition: that is, whether they see their relationships as reason-giving. Although, as noted earlier, there does not seem to be anything distinctively human about valuing, certain sophisticated cognitive capacities do seem to be required to recognise one’s relationship as a source of reasons. Kolodny himself suggests that small children likely lack the relevant capacities and therefore do not stand in genuine loving relationships with their parents (2003, fn. 22). If that’s right, then animals presumably also lack these capacities and thus fail to have genuine loving relationships.

Turn now to personal projects. Empirical research has shown that many animals engage in temporally extended and goal-oriented activities. To give just one example, certain species of birds (such as scrub jays) cache store food to recover it hours, days, or even weeks later when hungry, which suggests that they can anticipate and plan for future contingencies.¹³

But do these activities count as personal projects of the kind that is worth preserving? According to a popular view, genuine relationships and projects are similar in that they must be valued by an individual. For example, Scheffler (2010, p. 48) writes that

Valuing a personal project, like valuing a personal relationship, involves seeing it as reason-giving. In other words, to value a project of one's own is, among other things, to see it as giving reasons for action in a way that other people's projects do not, and in a way that other comparably valuable activities in which one might engage do not.

Overall, these considerations reveal that whether we have a conservative reason to extend rather than replace animal lives hinges on a largely empirical matter: whether animals have the kinds of cognitive capacities that are required for genuine

¹³ See Correia et al. (2007) for a recent study and Dickinson (2011) for an overview of the literature on goal-directed behaviour and future planning in animals.

loving relationships and personal projects. I will not attempt to settle this question here because it requires expertise in animal studies that I do not have. Instead, I will end this section by emphasising two insights from the preceding discussion.

The first insight is that those who find the practice of humane omnivorism morally troubling, as I do, now have a new line of critique at their disposal. They can argue that we have a moral reason to extend rather than replace animal lives when and because these animals' continued existence is necessary and sufficient for preserving certain valuable things like these animals' loving relationships and important personal projects. 'All' they have to do is show that these valuable things are, in fact, present in animal lives.

Of course, this might not be enough to conclude that humane omnivorism is always impermissible. Conservative reasons are merely *pro tanto* reasons, so they must be weighed against moral considerations that arguably support the practice of humane omnivorism, such as those concerning the pleasure some people derive from meat consumption. But I believe that conservative reasons are strong enough to make it difficult to outweigh them in this way.

Moreover, one attractive feature of this line of argument is that conservative reasons apply to a wide range of ways in which humane omnivorism could be practised. For example, they apply not only to an arrangement in which animals are killed prematurely but also to a situation in which animals are genetically engineered to drop dead when in prime condition for consuming their flesh. From the

conservative point of view, what matters is not whether the animal would have survived otherwise (had we not consumed it), but whether it is in our power to ensure that it stays alive. In particular, if a pill or some other form of treatment could mitigate the effects of genetic engineering and secure additional years of life for the animal, we would have a *pro tanto* conservative reason to provide it.

The second insight is of sociological nature. I noted earlier that people's intuitions are divided with respect to whether we have a moral reason to choose extension over replacement for animals. The conservative account suggests that this normative disagreement might have its root in a disagreement about a largely empirical matter: whether animal lives feature genuine, non-instrumentally valuable loving relationships and important personal projects.

5. Continuation and continuity

As should be clear by now, the account I have proposed has a reductive character. Our conservative reasons to extend the life of a person or an animal are grounded in the importance of preserving the valuable contents of their life, such as their loving relationships and important personal projects. Likewise, our reasons for extending humanity's tenure are grounded in the importance of preserving the valuable elements of humanity's culture, such as important traditions, beautiful artworks, and unique languages. On this view, what we ultimately care about from

the conservative point of view is the *continuity* of an individual's life or humanity's tenure rather than the mere *continuation* of their existence.

Some people who are sympathetic to conservatism about value might find this approach strange. In particular, they might insist that, in addition to preserving the valuable contents of an individual's life and the valuable elements of humanity's culture, conservatives about value should care about preserving the individual's life and humanity (understood as the human species) in themselves. To deny that, the thought goes, would be to unduly instrumentalise the value of the individual's life and humanity. We can call this *the non-reductive view*.

One proponent of the non-reductive view is Johann Frick (2017, p. 359), who asserts the following in his discussion of the reasons to prevent the extinction of humanity:

It is commonplace to claim of a wide range of things that they have final value in this sense: wonders of nature, great works of art, animal and plant species, languages, culture, etc. The suggestion that *humanity* too, with its unique capacities for complex language use and rational thought, its sensitivity to moral reasons, its ability to produce and appreciate art, music, and scientific knowledge, its sense of history, and so on, should be deemed to possess final value, therefore strikes me as extremely plausible. I do not, however, have the space to argue this claim in this article. I will ask you to

grant it to me as a premise, in the interest of seeing whether the final value of humanity may ground a moral reason to ensure humanity's survival. What I shall argue in the following is that there is a link between responding *appropriately* to the final value of humanity and being at least disposed to ensure its survival. (p. 359)

In the same spirit, one might think that individual humans (and perhaps animals) have final value in virtue of their unique rational and emotional capacities, and that we have conservative reasons to preserve them that go above and beyond our reasons to preserve the valuable contents of their lives. On this view, the mere continuation of an individual's or humanity's existence is within the scope of the conservative concern (as long as the relevant capacities are retained).

I think that we should reject the non-reductive view for two main reasons. The first reason has to do with theoretical parsimony. As evidenced by the quote above, proponents of the non-reductive view accept that the valuable elements of our culture (and presumably the valuable contents of our lives) are in themselves worth preserving. But, as my argument shows, that's all we need to explain the intuition that we almost always have a weighty moral reason to choose extension over replacement. There is simply no need to postulate that individual animal lives, individual human lives, and humanity are *also* worth preserving in themselves. Thus,

on the grounds of theoretical parsimony, we should reject the non-reductive view in favour of the reductive view.

The second reason has to do with extensional adequacy. I believe that the non-reductive view over-generates: there are cases in which we intuitively lack a conservative reason to choose extension over replacement, but the non-reductive view implies that we do.

To be sure, the non-reductive view need not imply that we always have a moral reason to extend the life of an animal, person, or humanity. For example, in the previously mentioned paper, Frick (2017) suggests that we would not have a conservative reason to ensure the survival of humanity if it were to irreversibly lose its unique capacities, such as those for producing and engaging with valuable cultural artefacts. Presumably, proponents of the non-reductive view would also say that we would have no conservative reason to extend the life of a person (or an animal) if they were to lose the rational and emotional capacities that allow them to initiate and maintain loving relationships and important personal projects. In this respect, the non-reductive view and the reductive view are in agreement.

However, the non-reductive view implies that we would have a conservative reason to extend humanity's tenure rather than allow its replacement with another intelligent and sentient species even if every single valuable element of humanity's culture would be eradicated: all of our traditions, languages, works of art, technologies, written and oral histories, and ways of life. Likewise, the non-reductive

view implies that we would have a conservative reason to extend the life of a person (or an animal) rather than allow their replacement with another person even if every single valuable element of their life would perish, including all of their loving relationships and important personal projects. All that appears to matter to the non-reductivist in these cases is a brute fact about numerical identity: that the existence of a particular individual or particular species is extended.

In this respect, I believe, the non-reductive view betrays the conservative ethos. Take the example of a beautiful painting. Suppose that we can either erase it and create another beautiful artwork on the same canvas, or else destroy the first canvas entirely and create another artwork on a different canvas. There is no conservative reason to do the former rather than the latter. I submit that the mere continued existence of a particular individual or species is nothing more than the continued existence of a canvas. We have no conservative reason to prefer the continued existence of an individual whose life has two discontinuous halves and features entirely different contents (effectively, an individual with two distinct lives) over the consecutive existence of two individuals with corresponding lives. Likewise, we have no conservative reason to prefer the continued existence of a single (human) species whose history has two discontinuous halves and features entirely different cultures (effectively, a species with two distinct civilisations) over

the consecutive existence of two species with corresponding cultures. This, I think, gives us a decisive reason to reject the non-reductive view.¹⁴

6. Further issues

The first part of this paper identified a common ground for our moral reasons to favour extension over replacement with respect to humanity, persons, and animals. In this section, I want to address two further issues connected to that discussion.

The first issue concerns the methods that could be used to extend a person's (or an animal's) life. In recent years, we have made significant progress in our scientific understanding of nutritional, lifestyle, pharmacological, and genetic interventions that can potentially slow down or even reverse ageing.¹⁵ However, I want to focus on two more speculative routes to a longer life that continue to feature

¹⁴ Suppose, for the moment, that I am mistaken, and we have good reasons to accept the non-reductive view. Does that invalidate my project of developing the reductive view? I don't think so. Strictly speaking, the non-reductive view and the reductive view are compatible with each other. If that's the case, then in typical cases, we would have a conservative reason to extend humanity's tenure in virtue of the fact that the human species is itself worth preserving *and* a conservative reason to do that in virtue of the fact that its survival is necessary and sufficient for the continued existence of certain valuable elements of our culture which are in themselves worth preserving. This would be just another instance of the commonplace ethical phenomenon of *normative overdetermination*. Crucially, overdetermination does not mean redundancy here. If there are both non-reductive and reductive conservative reasons to choose extension over replacement, then the conservative support for extension is in one way stronger. This is important because, as I noted earlier, conservative reasons are *pro tanto* reasons and have to be weighed against other moral considerations, such as the prospective value of the replacement outcome. So, perhaps, we should choose extension over replacement even if the total welfare in the latter scenario is much higher, and not just slightly higher.

¹⁵ For overviews of recent developments in this field, see Kennedy et al. (2014), Cohen (2015), Newman et al. (2016), and Partridge et al. (2020).

prominently in the popular imagination: cryonics and uploading. Cryonics is the process of storing a person's brain or whole body at a very low temperature after their legal death in the hope that scientific progress will one day make revival possible. Uploading, on the other hand, involves scanning a person's brain and uploading this information to a computer, where their brain is then simulated.

Concerns about feasibility aside, philosophical discussions of these potential methods of life extensions have focused on two issues. The first is whether these methods would preserve a person's consciousness.¹⁶ The second is whether they would preserve one's memories and thus their psychological continuity, which many philosophers regard as necessary for personal identity.¹⁷

My conservative account puts a spotlight on a third, novel kind of worry. In a nutshell, even if these life extension methods were consciousness- and identity-preserving, something else could be lost in the process: one's loving relationships and important personal projects.

To see that, note that our loving relationships and important personal projects are embedded in the external world. This is important in a number of respects. Most obviously, our loving relationships are relationships with other people. If a person undergoes cryonics and is revived decades later in a world from

¹⁶ See Chalmers (2010).

¹⁷ For this worry, see Chalmers (2010), Aaronson (2016), and Doyle (2018). For an overview of the literature on personal identity, see Olson (2022).

which their loved ones are absent, their loving relationships will not survive either. Likewise, if a person is uploaded, but their loved ones stay in the physical world, it will be impossible, or at least more difficult, to maintain these loving relationships. Similar considerations apply to our important personal projects. While they need not involve other people, they generally depend for their purpose and value on other features of the external world. For example, the project of being the first person to run the marathon under two hours ceases to make sense if that has already been done by someone else (as will likely be the case if you spend decades in a cryogenic tank) or if you no longer have a physical body. Thus, extending one's life in one of these ways would often involve sacrificing these prudential goods, even if one's consciousness and identity were preserved. In such cases, we would lack a conservative reason to choose the extension of a person's life through cryonics or uploading over their replacement.

The second issue connected to my discussion of the choice between extension and replacement is this. We have previously considered cases, including the ones just above involving cryonics and uploading, in which extending the existence of a person or humanity is not sufficient for the preservation of any valuable things, and so there is no conservative reason to choose extension over replacement. But perhaps an even stronger conclusion is warranted in these scenarios: namely, that we have a conservative reason to choose replacement over extension.

To assess that, let's return to the case in which a person can undergo an aggressive form of cancer treatment that would make them suffer from severe fatigue, brain fog, mood changes, stress, and loss of sexual function. These side-effects, we have stipulated, would prevent a person from engaging in fulfilling social interactions, valuing their goals, and maintaining serious intellectual engagements. Suppose now that this person would not entirely lose their loving relationships and important personal projects, but they would instead survive in some significantly diminished state.

What does the conservative view about value say about this case? Plausibly, conservative reasons go beyond merely ensuring that valuable things continue to exist. For example, when someone lets an extraordinarily beautiful painting deteriorate so much that it survives at a level at which it is just barely attractive, their behaviour clearly falls short of being a fitting response to the painting's value. This suggests that it is important to preserve valuable things in a prosperous condition. Now, in the case of the cancer patient, the concern is precisely that, in virtue of staying alive, that person would fall short of this standard and thereby mistreat their relationships and projects. This suggests that there would actually be a conservative reason to choose replacement over extension in this case.

To be sure, if the patient does not undergo the treatment and dies soon, their loving relationships and important personal projects will come to an end. This might seem like an even worse fate for these valuable things compared to surviving

in a diminished state. However, what matters for the conservative assessment of these scenarios are not just the *end states*, but also the *manner* in which they will come about. It is one thing to let an important text or a religious emblem tatter and find its way to a landfill, and another to bury or burn it in a ceremonial way. Likewise, it seems more fitting to allow our loving relationships and important personal projects to end when they are still thriving than to have them survive in a scarred condition.¹⁸

Analogous considerations apply to humanity as a whole: in some circumstances, we may have a conservative reason to let it be replaced by another intelligent and sentient species rather than to extend its tenure. This would be so if ensuring the survival of humanity was possible only at the cost of having many valuable elements of our culture survive in a significantly diminished state. I won't repeat the entire argument for this conclusion.

Instead, I want to use this opportunity to connect the discussion of this issue with a famous claim made by Derek Parfit. In the closing section of *Reasons and Persons* (1984), Parfit asks us to compare three outcomes: (1) peace; (2) a nuclear war that kills 99% of the world's population; and (3) a nuclear war that kills everyone. He claims that while (2) is worse than (1), (3) is 'very much' worse than (2) because a nuclear war that kills everyone would foreclose human history, which might

¹⁸ Sometimes, one might have a third option: to extend their life but end all of their engagements before they perish. However, it is not possible to terminate in a dignified way all of one's valuable attachments with the intention of, say, experiencing a few extra years of pleasure. So, conservative reasons tell against this alternative.

otherwise involve billions of extraordinarily happy lives and many spectacular achievements.

Parfit might be ultimately right here. But I think that it is important to recognise that, *in one respect*, the scenario in which humanity goes completely extinct could be better than the scenario in which a tiny fraction of the world's population survives. This would be so when and because the survival of humanity in this diminished state would degrade the valuable elements of our culture, such as important traditions, great works of art, just institutions, scientific knowledge, and relations of equality between members of the moral community. As in the case of beautiful paintings, loving relationships, and important personal projects, it seems more fitting to allow the valuable elements of our culture to disappear when they are still thriving than to have them survive in a scarred condition.

7. Concluding remarks

What, if anything, grounds our moral reasons to choose extension over replacement with respect to non-human animals, persons, and humanity as a whole?

According to the conservative account that I have presented, we have a weighty moral reason to choose extension over replacement when and because extension is necessary and sufficient for the preservation of certain valuable things, such as cultural artefacts, loving relationships, or important personal projects.

This conservative account does not imply that we always have a moral reason to choose extension over replacement. However, in the context of persons and humanity as a whole, it implies that we almost always do. In the case of animals, on the other hand, the jury is still out: we must first determine whether their lives feature genuine loving relationships and important personal projects of the kind that warrants further preservation.

Part III: Justice and Time

Chapter 5: Healthspan Extension, Completeness of Life, and Justice

1. Introduction

Recent progress in geroscience holds the promise of significantly slowing down—or even reversing—the process of ageing and the development of age-related diseases, such as cancer, heart disease, and dementia. A number of interventions have been shown to have this effect in species ranging from yeast to non-human primates. These include pharmacological and genetic interventions which aim to prevent the shortening of telomeres, mimic the caloric restriction diet, increase the number and potency of our stem cells, modulate the genes and genetic pathways that regulate cellular ageing, or improve mitochondrial activity, among other things. Some of these research programmes have already entered or are about to enter human trials, and could become a part of medical practice within a decade or two.¹

Crucially, these interventions have the potential to extend not just our *lifespans* but also our *healthspans*. That is, in addition to increasing the number of years we are alive, they could also increase the number of years we remain in the state of physical, cognitive, and reproductive health. In doing that, they would extend the prime of our lives while leaving the other two nominal stages—childhood and older

¹ For overviews of recent developments in this field, see Kennedy et al. (2014), Cohen (2015), Newman et al. (2016), and Partridge et al. (2020).

age—intact. Alternatively, these interventions could increase our healthspans without extending our lifespans, by compressing morbidities and disabilities towards the end of life.²

Suppose that these visions of the future are warranted. In particular, assume that we could develop a technology that would allow people to add, in one of these ways, at least 20 healthy years to their life. (Of course, more might eventually become possible, but this number represents a reasonable starting point for ethical theorising.) Call this *healthspan extension*.

In this paper, I argue that, as a matter of justice, we have a duty to develop and widely distribute such technology. This is because justice requires that each person be provided with sufficient opportunities to have a ‘complete life’, many people currently lack such opportunities, and increasing people’s healthspans would substantially improve the status quo in this respect. Along the way, I examine a range of issues intersecting longevity, social disadvantage, and welfare.

It is worth emphasising that my central claim is not that we would *benefit* people by giving them access to healthspan extension technology (though that might be true as well). Instead, the claim is that providing access to this technology is something that we *owe* to people as a matter of justice. This distinction is important

² It is worth noting that a compression of morbidity and disability naturally occurs among the longest-lived humans (Andersen et al. 2012). Thus, we can perhaps expect anti-ageing technology to increase our healthspans both by extending our lifespans and by compressing morbidity and disability.

because duties of justice are generally understood to be more stringent than duties of beneficence. For example, we cannot escape these duties simply by citing the high cost of satisfying them and failing to satisfy these duties often amounts to wronging someone.

2. Justice and Completeness of a Life

The starting point of my argument is a broadly egalitarian conception of justice which holds that, at the minimum, justice requires that each person be provided with a fair share of opportunities.³ This formulation of the ideal of justice is, of course, incomplete. We need to specify what counts as a fair share and what are the opportunities in question.

Consider the first issue. There are three main positions here. On a (strictly) *egalitarian* view, fairness requires distributive equality.⁴ On a *prioritarian* view, fairness requires that we give priority to those who are worse off.⁵ On a *sufficientarian* view, fairness requires that everyone has enough, i.e., that they clear a certain threshold.⁶

Much ink has been spilled over the relative merits of these three distributive principles, and I do not attempt to adjudicate between them here. The main claim

³ For prominent views along these lines, see Arneson (1989) and Cohen (1989).

⁴ See Temkin (1993; 2003) for a defence.

⁵ For influential statements of this view, see Sen (1973) and Parfit (1995).

⁶ See Frankfurt (1987), Crisp (2003), Casal (2007), and Timmer (fc) for key discussions.

of this paper—that justice requires that we develop and widely distribute healthspan extension—follows from any plausible distributive principle.

However, to make our discussion tractable, it is essential to adopt a particular view about what counts as a fair share. In particular, let's make the plausible assumption that *it is a matter of special significance that each person have sufficient opportunities*. This position is much weaker than sufficientarianism: it leaves open, for example, that equality matters above the sufficiency threshold, or that priority should be given to those below the threshold who are worse off. But, as we will see, it is enough to generate the conclusion that justice requires that we develop and widely distribute healthspan extension technology.

Turn now to the second question. Common examples of opportunities within the purview of distributive justice are things like political opportunities (e.g., to run for office), educational opportunities (e.g., to attend a university), and economic opportunities (e.g., to find meaningful employment), or opportunities for welfare in general. But equally important, or perhaps even more important, are people's opportunities to have a *complete life*—which we can define as a life that involves, in satisfactory quantity and quality, all fundamental ingredients of welfare.

Although my argument goes through on any plausible conception of welfare, for the sake of concreteness, let's assume *the objective list theory*.⁷ A major

⁷ This label is due to Parfit (1984). See Lin (2022) for a recent overview of philosophical literature on welfare.

attraction of this view is that it recognises a whole array of basic prudential goods, such as personal achievements, loving relationships, social contributions, social recognition, pleasure or satisfaction with one's life, and possibly others. Accordingly, a complete life is conceptualised as a life which involves all of these fundamental ingredients of welfare in satisfactory quality and quantity.⁸

With these preliminary considerations in place, the first premise of my argument takes the following shape: justice requires that each person be provided with sufficient opportunities to have a complete life. The next section explains the extent to which our society falls short of this ideal.

3. Lopsided lives

Certain goods are of special importance to our opportunities in general, and to the opportunity to have a complete life in particular. Freedom is one obvious example: a person who cannot act freely will not be able to engage in loving relationships or personal projects, which could culminate in achievements, or at least their opportunities to attain these and other prudential goods will be very diminished.

Health has this profile as well. As Norman Daniels (1985; 2001; 2008) has emphasised, health is a precondition for attaining what we, as a matter of justice,

⁸ It is worth noting that 'complete life' is a higher standard than 'life worth living'. Moreover, on the objective list theory, it is possible for a life to have very high welfare but not be complete. This is so, for example, when a life is rich in pleasures but largely deprived of achievements and relationships.

should be able to attain. To take an extreme example, a person who is not capable of moving on their own or who lacks the ability to make and carry out plans will find it extremely difficult to attain certain basic prudential goods, such as achievements, relationships, or social contributions. But even less severe conditions—such as chronic pain, depression, or impaired mobility—can present serious challenges in this respect.

What is often overlooked, however, is that not just *the state of health* is important. An adequate *healthspan*—the number of years we are in physical, cognitive, and reproductive health—is crucial as well. After all, maintaining projects which culminate in achievements or social recognition requires a considerable amount of time, and so does finding a romantic partner and maintaining a valuable relationship. Consequently, people with short healthspans will often lack sufficient opportunities to have a complete life.

This observation has an influential precedent in the bioethics literature. According to John Harris (1985), there is some length of life (say, 70 or 80 years) which is *normally* sufficient for a complete life, and this age serves as a threshold that we can use to allocate scarce healthcare resources. On this view, we should save a 20-year old rather than an 80-year old because the former person has not yet

received their fair share of opportunities, or ‘playtime’—at least when other things are equal.⁹ This is known as *the Fair Innings View*.

The Fair Innings View can serve as a useful heuristic for allocating limited healthcare resources, but it clearly oversimplifies the relationship between healthspan and opportunities. There are many people for whom even a normal healthspan does not present sufficient opportunities to have a complete life. This is particularly true of people who face various forms of social disadvantage, such as poverty, discrimination, disability, and exclusion. Each of these conditions makes it more difficult for a person to achieve things, form loving relationships, receive recognition, make social contributions, and experience pleasure in satisfactory quantity and quality. As a result, many people’s lives are, at best, *lopsided*: rich in one type of prudential good, but largely deprived of others.¹⁰ To appreciate this situation, consider the key prudential goods constitutive of a complete life one by one.

Start with *achievements*. An achievement is a product of a personal project or some other sustained, goal-oriented effort.¹¹ Such efforts can be undertaken within an educational context, as a part of one’s job, or in one’s disposable time. However, many people lack sufficient opportunities to have significant achievements in these settings. To begin with, in many countries, higher education is unaffordable, has a

⁹ For example, when the 20-year old and the 80-year old can be expected to live for as many years and enjoy equally high quality of life.

¹⁰ This label is due to Pummer (2017).

¹¹ See Bradford (2015) for a recent book-length treatment of the nature of achievements.

high alternative cost, or otherwise presents various barriers to entry for the disadvantaged parts of the population. Similar considerations apply to labour markets. For example, an Amazon warehouse employee whose primary task is to reach for an item in one container and place it in another has few meaningful opportunities for accomplishments, and occupations which are more rewarding in this respect are scarce. Finally, cultivating personal projects—be it running a marathon or learning an instrument—often requires significant amounts of time, attention, energy, or wealth, and those in socially disadvantaged positions often cannot afford that.¹²

Turn now to *relationships*. Many people's opportunities to form and maintain romantic and parent-child relationships are significantly restricted in virtue of the fact that women's reproductive capabilities—understood in terms of fertility, offspring health, offspring retention, and maternal health—decline rapidly in their thirties.¹³ This is so in at least two ways. For one thing, many people are forced to choose between parenthood and other important facets of life, such as pursuing a professional career or a major personal project, because of the difficulty involved in balancing these roles.¹⁴ This is especially true of people who cannot afford childcare

¹² For an illuminating discussion of how the changing nature of employment exacerbates the impact of social disadvantage on people's lives, see Boushey (2016).

¹³ Between 2017 and 2019, only 3.5% of all births in the United States were from women past the age of 40 years (Martin et al. 2021).

¹⁴ Indeed, increases in women's educational enrolment and labour force participation are understood to be key factors contributing to the decline in fertility rates and the postponement of childbearing (Mills et al. 2011).

or count on the support of their relatives. And even those who have the necessary financial means and a support network might have to give up parenthood, for example, because pregnancy and childbirth make it impossible for them to compete in sports at the highest level. For another, since the desire to have children is almost universal, many people find themselves under considerable pressure to find a romantic partner relatively early in their life, before their reproductive health deteriorates. As a result, they can rush into things and settle for less than what's best for them.

Next, consider *social contributions*. Some opportunities for making significant social contributions can arise within the regular context of employment: for example, a microbiologist might develop a new vaccine for a common disease or a firefighter might prevent a fire in one building from spreading across the neighbourhood. Alternatively, one might seek out such opportunities in one's disposable time by volunteering, donating a portion of their income to support effective charities, or being an active member of their local community. As in the case of achievements and relationships, not everyone is in a good position to do that. Careers which allow people to make meaningful social contributions are scarce and may require higher education or specialised training. And the contributions made outside of one's employment often require a substantial investment of time, attention, energy, or wealth that many people simply cannot afford.

Further, turn to *social recognition*. It is often mediated by other basic prudential goods: achievements, relationships, and social contributions. For example, our society might appreciate the achievement of an athlete who was the first to scale a high peak, the dedication of a local community leader, or the supportiveness of a family member. However, this ‘second-order’ nature of social recognition means that people who are deprived of sufficient opportunities to attain these other goods will also lack sufficient opportunities to receive social recognition. To make matters worse, people who face forms of social disadvantage are sometimes forced by their circumstances to take jobs or engage in activities which not only do not give them a chance to receive social recognition but can even undermine it (e.g., criminal activity and perhaps sex work).

Finally, consider *pleasure* (or satisfaction with one’s life). Our pleasant experiences tend to have a derivative character, much like social recognition. This is so in the sense that our achievements, relationships, social contributions, and even social recognition are all important sources of pleasure in our lives. Thus, inadequate opportunities to attain the former goods will often result in insufficient opportunities to find pleasure as well. The other important cradle of pleasant experiences are things we do in our disposable time: travelling, reading, exercising, watching plays and sports events, or cooking. Unfortunately, immersive participation in these activities often requires significant amounts of time or money,

so those who are disadvantaged will often lack genuine opportunities to engage in pleasurable pursuits on a daily basis.

Taken together, these considerations point to the following conclusion. As is well-known, forms of social disadvantage—such as poverty, discrimination, disability, and exclusion—undermine our opportunities to attain each of the basic prudential goods. However, against the backdrop of the present-day healthspan, these forms of disadvantage have an even more profound effect. They make the basic prudential goods *exclusionary*, in the sense that many people are forced to choose which of them to pursue. This is a key reason why many people currently lack sufficient opportunities to have a complete life.

Of course, there is room for reasonable disagreement about where exactly the threshold for sufficiency lies, and thus how many people are below this threshold. But for now let's put this issue aside and consider whether the status quo would be substantially improved by giving people access to technology that would increase their healthspans by at least 20 years. The next section argues that it would, in at least four ways.

4. The promise of healthspan extension

Healthspan extension technology has the potential to transform many aspects of our society and daily life, ranging from political and economic arrangements to the nature of work and personal relationships. The ensuing discussion focuses on four

effects that are particularly plausible and would have the biggest positive impact on people's opportunities to have a complete life. Needless to say, each of these mechanisms is to be understood in *ceteris paribus* terms; we are examining one variable at a time, while holding others constant. This is a natural starting point for an analysis of any complex social phenomenon.

First and foremost, healthspan extension would substantially improve people's opportunities to have a complete life simply in virtue of giving them more time to attain the relevant goods. We can call this *the Temporal Empowerment Effect*.

There are, in fact, two distinct aspects of temporal empowerment. In the first instance, healthspan extension would enhance people's ability to attain multiple prudential goods *sequentially*, one after the other.

One reason for this has to do with reproductive health. If people had the option of having children in their forties and fifties without incurring significant health risks and financial costs, they would be in a better position to take on endeavours that culminate in significant achievements and social contributions. This is because they would no longer be forced to choose between, or try to balance, demanding careers or personal projects and starting a family. Similar considerations apply to relationships. If the 'deadline' for having children were to be extended by twenty or more years, many people would be better placed to find a true match.

There would be analogous benefits in virtue of the extended window of physical and cognitive health. If people were equally able to succeed in demanding

careers, educational settings, and strenuous personal projects later in life, they could pursue them one after the other, or after they had already raised children. As a result, they would not be compelled to try to attain all of these goods at the same time, which is something that many people, especially those facing social disadvantage, are unable to do.

That said, there is another—perhaps less obvious—aspect of temporal empowerment. Healthspan extension would also enhance people’s ability to pursue multiple goods of life *simultaneously*. The main factor here is that extending the prime of their lives holds the promise of reducing the temporal burden of employment. As things stand, relatively long periods of childhood and older age—both of their own and of their dependents—force people to squeeze out as much income as possible from their productive years. This is especially true of those who suffer from forms of social disadvantage or live in a country without an adequate social security system.

If people’s healthspan were to increase substantially, and with it the number of years they could spend in employment, they could afford to spend fewer hours on any given day and in any given year working, as well as preparing for and recovering from work.¹⁵ This is because, even if the absolute length of childhood and older age (the nominal stages of life when they are more dependent on others

¹⁵ For an insightful discussion of the problem of the increasing temporal colonization of work, see Danaher (2019).

or our past selves) would not change, their relative length would shrink. Consequently, people would have more time—and in some cases, finally enough time—for personal projects, socially beneficial pursuits, and pleasurable activities.

Second, healthspan extension would likely enhance people's ability to overcome socio-economic disadvantages, and thereby access a wider range of opportunities. We can call this *the Upward Mobility Effect*. To illustrate this, consider higher education and specialised training. Both are prime contexts for achievements and are often necessary to access especially rewarding careers and opportunities for making significant contributions to the society. Unfortunately, in many places, these forms of tuition are expensive, have a prohibitively high opportunity cost, or otherwise exclude various parts of the population. Healthspan extension has the potential to reduce these barriers to entry. For example, people who do not have the means to enrol in their preferred education programme, or move to a location where it is offered, would now have more time to first build up the necessary wealth, as well as security that comes with work experience. Moreover, the prospect of a longer healthspan would make the decision to pursue educational opportunities later in life less risky, as the number of years during which one could reap their benefits would increase. Crucially, these kinds of benefits would likely be passed down the chain of generations. After all, improving a single person's opportunities to attain the goods constitutive of a complete life would often also improve the opportunities

of their offspring. In that way, healthspan extension could have a lasting impact on upward mobility.

Third, what goes for social disadvantage also applies to bad luck and bad decisions. If people's healthspans were longer, many of their misfortunes and mistakes would, in one way, weigh less heavily on their lives. A traffic accident resulting in an injury requiring years of rehabilitation would no longer put an end to a promising cycling career. A decade wasted in an abusive relationship would not jeopardise one's dream of starting a family. And choosing not to apply to graduate school after college would not prevent one from embarking on a long and successful academic career later in life. We can call this *the Safety Net Effect*.

Fourth, healthspan extension would improve people's opportunities to have a complete life because, in the process of living a longer life, they would likely gain a better understanding of themselves and what's worth doing. We can refer to this as *the Wisdom of Age Effect*.

There are two main issues here. The first is that, when we are young, we are still learning about our capabilities and the world around us. In particular, we are yet to fully appreciate what possible engagements the world can offer us, how much effort they require, what we are good at, what brings us joy, and how much we can fit in a day, a year, or a lifetime. It is only as we grow older and gain more experience that we develop this kind of insight. If people's healthspans were longer, they would have more time both to gather this information and to put it to use. This would put

them in a better position to take advantage of their talents and unique circumstances to achieve things, find meaningful relationships, and make lasting contributions.

The second issue has to do with the order in which we tend to pursue things. While the objective list theory identifies social contribution as one of the fundamental prudential goods, not everyone recognises its importance from the outset. For many people, achievements, relationships, and pleasure take priority. It is only after they have attained these other goods that they turn to endeavours aimed at contributing to the wider society. We see that most clearly in the case of some billionaires who, having spent the first fifty or sixty years of their life chasing profits and status, eventually turned to philanthropic efforts. But this trajectory is not exclusive to the rich. Many people seek out ways to contribute to the wider society only after they ‘made it’ in other domains of life. Insofar as healthspan extension would help them do the latter, it would also enhance their opportunities to make a social contribution.

The discussion of these four effects—the Temporal Empowerment Effect, the Upward Mobility Effect, the Safety Net Effect, and the Wisdom of Age Effect—has focused on achievements, relationships, and social contribution. But it should be easy to see that their influence extends to the other two basic goods of life mentioned earlier: social recognition and pleasure. Because these goods have a largely derivative character (what we are appreciated for and what we take pleasure

in are often our achievements, relationships, and social contributions), healthspan extension would also improve people's opportunities to attain them.

Before concluding this section, it is important to acknowledge that any technological revolution brings with it a host of unintended and unforeseeable consequences. So it is entirely possible that healthspan extension could also diminish people's opportunities to have a complete life in certain ways. Still, the four effects described above seem strong enough to warrant the conclusion that healthspan extension *would*, and not just *could*, substantially improve the status quo.

If that's right, we have all the necessary pieces to complete the overarching argument of this paper. Justice requires that people be provided with sufficient opportunities to have a complete life (§2), many people currently lack such opportunities (§3), and healthspan extension would substantially improve the status quo (§4), so justice requires that, if possible, we develop healthspan extension technology and make it widely available. The remainder of the paper addresses three objections to this argument and some practical considerations.

5. Three objections

The first objection targets the idea that those who fail to have a complete life are typically in this position because they lack sufficient opportunities to attain the core prudential goods in sufficient quantity and quality. Instead, the thought goes, they fail to have a complete life because they waste time. Thus, extending people's

healthspans by twenty or more years would make no meaningful difference to their opportunities. In the literature, this kind of concern, which we can call *the Wastefulness Objection*, goes back all the way to Seneca, who wrote:

It is not that we have a short time to live, but that we waste a lot of it. Life is long enough, and a sufficiently generous amount has been given to us for the highest achievements, if it were all well-invested. (2004, p. 1)¹⁶

What should we make of this concern? To begin with, I think that those who share Seneca's sentiment overstate how much of our lives is actually wasted. We certainly spend a significant portion of our lives idle. But idleness serves an important role. Most strikingly, an athlete can train only so many hours each day because it is essential to give the body time to recover. Likewise, a novelist can write only so many words on a given occasion before their prose becomes insufferable. And the same goes for many other kinds of engagements. If we are to make the most of our physical, creative, social, and rational capacities, we simply cannot spend every waking hour in active pursuit of the core goods of life.

Moreover, while it is true that some people fall short of having a complete life because they waste opportunities presented to them, this does not apply to

¹⁶ See also Jonas (1985) and Kass (2004) for this objection.

everyone. For reasons listed earlier, it seems impossible to deny that, in virtue of social disadvantage, many people are deprived of the relevant opportunities.

Finally, suppose that some people would, in fact, waste the entirety of their additional healthspans. (This seems unlikely if they are afforded only a few decades rather than full-blown immortality, but it's worth considering nonetheless.¹⁷) That would not be a major strike against my argument, either. The conception of justice employed in this argument has as its currency not 'outcomes' but rather 'opportunities'. That is, we are primarily concerned with giving everyone sufficient opportunities to have a complete life, and not with everyone actually having a complete life. So, extending people's healthspan would improve the status quo with respect to justice even if many people ended up wasting their additional time.

The second objection concerns the notion of the complete life. One might worry that healthspan extension would not improve people's opportunities to have a complete life because what counts as a complete life is not fixed but rather relative to one's healthspan. That is, compared to a person with a normal healthspan, a person with an extended healthspan would need to have correspondingly greater achievements, relationships, social contributions, social recognition, and pleasure in order to count as having a complete life. And, the objection goes, extending a person's healthspan would afford them at most a proportional increase in their

¹⁷ For an influential discussion of the effect that immortality could have on our motivation and ability to pursue certain goods of life, see Williams (1973).

opportunities. So if that person lacked sufficient opportunities to have a complete life in the absence of healthspan extension, they would not have them in its presence as well. We can call this *the Shifting Goalposts Objection*.

Both premises of this objection are dubious. First, on reflection, the idea that what counts as a complete life is indexed to one's healthspan is not tenable. We can see that most clearly when we consider lives with healthspans which are shorter than normal. It is a great tragedy when a person has very few achievements, relationships, social contributions, etc., even when they get to live for only ten or twenty years, rather than eighty. Indeed, this very line of thought plays an important role in motivating the popular Fair Innings View about the allocation of healthcare resources mentioned earlier.¹⁸

Second, there is some reason to believe that a substantial enough increase of a person's healthspan would afford them a greater-than-proportional increase in their opportunities to have a complete life. The key idea here is that certain opportunities have cumulative effects on our life. For example, having access to higher education and specialised training gives us greater opportunities for accomplishments and appreciation not only in the educational context, but also further down the road in our professional and personal life. Likewise, if the

¹⁸ Of course, it could be that morality is asymmetric in this respect and only increases—and not decreases—of healthspan change the requirements for a complete life. But the burden of showing that this is the case lies with the objector.

arguments of the previous section are sound, we should expect healthspan extension to have a positive impact on social mobility further down the chain of generations. Thus, even if it were true that what counts as a complete life is relative to one's healthspan, healthspan extension could still significantly improve people's opportunities to have a complete life.

The third objection takes a different direction. It seems that the present-day, limited healthspan is not *in itself* an impediment to having a complete life. Instead, it merely amplifies the force of other social problems, such as poverty, discrimination, disability, and exclusion. After all, it is plausible that at least some people with the present-day healthspan who do not face social disadvantage have sufficient opportunities to have a complete life. For this reason, the objection goes, it is a mistake to focus our efforts on extending the human healthspan rather than on ameliorating these more fundamental problems.

Moreover, the objector might add, the availability and widespread use of healthspan extension technology could make it more difficult to address these societal problems. This is because this technology would make the severity of social disadvantages harder to appreciate, and, as a result, our society would be less motivated to combat them. In this respect, healthspan extension technology could be likened to a painkiller which can make it seem that a disease which is causing the pain is less important to treat than it actually is. We can refer to this as *the Real Issue Objection*.

This objection raises an important concern. There is no denying that poverty, discrimination, disability, and exclusion are important issues which need to be addressed head-on, and we should not lose sight of them in the pursuit of technological advancement. However, the Real Issue Objection does not carry the day. For one thing, it is far from clear that a definitive solution to these problems can be found and implemented in the foreseeable future. In the meantime, we should settle for the second-best option, which is to alleviate the effects that social disadvantage has on people's opportunities to have a complete life. The four effects identified in the previous section—Temporal Empowerment, Upward Mobility, Safety Net, and Wisdom of Age—suggest that healthspan extension has the potential to do just that.¹⁹

Moreover, the second part of the Real Issue Objection unjustifiably instrumentalises people's misfortune. While it may be true that our society would be more motivated to address poverty, discrimination, and exclusion if these problems had more severe impacts on our lives, it is unacceptable to let people suffer avoidable consequences just so that they can serve as a reminder of the existence of the underlying problems. In that respect, it would be like refusing to treat cancer

¹⁹ Of course, that is not to say that we should divert all funding away from fighting poverty, discrimination, and exclusion, and into healthspan extension research. There are many possible ways of improving people's opportunities to have a complete life and to determine which one we should pursue we should assess their relative effectiveness. My aim in this paper is to show that healthspan extension merits a place on the list and serious consideration.

patients using the imperfect forms of treatment available nowadays so that our society would accelerate the efforts to find a genuine cure.

6. Unconditional access

Finally, let's examine a practical matter. On the plausible assumption that its cost would be non-negligible, should healthspan extension be available to everyone or just to those who currently lack sufficient opportunities to have a complete life?

It might seem that the presented argument can only support the latter option. After all, the weak sufficientarian principle of distributive justice assumed earlier seems to generate no reason to provide people who already have sufficient opportunities to have a complete life with means of improving these opportunities even further.

However, that's not quite right. Even on this minimal conception, reasons of justice support making healthspan extension available to everyone and unconditionally. To see that, consider the steps that would be required to offer healthspan extension on a means-tested basis and the effects that this regime would have on the society.

To begin with, there would have to be a system for assessing whether a particular person already has sufficient opportunities to have a complete life. That is already difficult at the theoretical level: ethical theory can rarely provide precise guidance in these matters. And it sounds almost impossible at the practical level,

considering how much information about an individual would have to be collected and analysed.

Moreover, even if it were possible to efficiently collect and process such data, a means-tested system would likely be less effective at alleviating the problem at hand. This is because it would divide the society between those who need help and those who can manage on their own, and thereby stigmatise the recipients of healthspan extension. Moreover, the intrusive nature of means-testing would disincentivise people from applying for healthspan extension. In light of these issues, we should expect healthspan extension to reach fewer people under a means-tested scheme compared to an unconditional scheme.

In this latter respect, the argument for unconditional healthspan extension is similar to a common argument for Universal Basic Income. As proponents of this idea often emphasise, there is nothing humiliating or intrusive about receiving a basic income granted to all members of the society, and thus UBI is likely to be more effective at alleviating poverty than unemployment benefits and other means-tested income support programmes.²⁰

7. Concluding remarks

In this paper, I have argued that we have a justice-based reason to develop healthspan extension technology and make it unconditionally available to everyone.

²⁰ See, for example, Van Parijs and Vanderborght (2017).

This is because justice requires that each person be provided with sufficient opportunities to have a complete life—which involves achievements, relationships, social contributions, social recognition, and pleasure in satisfactory quantity and quality—and healthspan extension would substantially improve the status quo in this respect.

In closing, let me emphasise that this discussion alone does not imply that we ought to, *all things considered*, develop healthspan extension technology and make it unconditionally available to everyone. While important, considerations of justice having to do with opportunities to have a complete life are just one element of a broader moral landscape. There are other normative considerations that we also need to take into account.²¹ But that must be left for another occasion.

²¹ For a comprehensive overview and discussion of objections to life extension technology found in the literature, see Davis (2018).

Bibliography

- Andersen, S. L., Sebastiani, P., Dworkis, D. A., Feldman, L., & Perls, T. T. (2012). Health span approximates life span among many supercentenarians: compression of morbidity at the approximate limit of life span. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, 67(4): 395–405.
- Aaronson, S. (2016). Ghost in the Quantum Turing Machine. In S. B. Cooper and A. Hodges, eds. *The Once and Future Turing: Computing the World*. Cambridge: Cambridge University Press.
- Adams, R. M. (1999). *Finite and Infinite Goods: A Framework for Ethics*. New York: Oxford University Press.
- Arendt, H. (1963). *Between Past and Future*. Cleveland: Meridian.
- Arneson, R. (1989). Equality and Equal Opportunity for Welfare. *Philosophical Studies*, 56 (1): 77-93.
- Arrhenius, G. (2008). Life Extension Versus Replacement. *Journal of Applied Philosophy*, 25 (3): 211-227.
- Bader, R. (2022). The Asymmetry. In J. McMahan, T. Campbell, J. Goodrich, & K. Ramakrishnan (eds.). *Ethics and Existence: The Legacy of Derek Parfit*. New York: Oxford University Press.
- Barnes, E. (2016). *The Minority Body: A Theory of Disability*. Oxford: Oxford University Press.
- Bigelow, J., Campbell, J., & Pargetter, R. (1990). Death and Well-being. *Pacific Philosophical Quarterly*, 71(4): 119-140.
- Blackorby, C., Bossert, W., and Donaldson, D. (2005). *Population Issues in Social-Choice Theory, Welfare Economics and Ethics*. New York: Cambridge University Press.
- Bognar, G. Fair Innings. *Bioethics*, 29 (4): 251-261.
- Bourguignon, F. & Morrisson, C. (2002). Inequality Among World Citizens: 1820-1992. *The American Economic Review*, 92(4): 727-744.
- Boushey, H. (2016). *Finding Time: The Economics of Work-Life Conflict*. Harvard University Press.
- Bradford, G. (2015). *Achievement*. Oxford: Oxford University Press.
- Bradley, B. (2009). *Well-being and Death*. Oxford: Oxford University Press.

- Brentano, F. (1973/2009). *The Foundation and Construction of Ethics*, edited and translated by Elizabeth Hughes Schneewind. New York: Routledge.
- Broome, J. (1993). Goodness is Reducible to Betterness: The Evil of Death is the Value of Life. In P. Koslowski and Y. Shionoya (eds.) *The Good and the Economical: Ethical Choices in Economics and Management*. Berlin: Springer.
- Broome, J. (2004). *Weighing Lives*. Oxford: Oxford University Press.
- Bury, J. B. (1921). *The Idea of Progress: An Inquiry into Its Origin and Growth*. London: Macmillan.
- Casal, P. (2007). Why Sufficiency is Not Enough. *Ethics*, 117: 296-326.
- Chalmers, D. (2010). The Singularity: A Philosophical Analysis. *Journal of Consciousness Studies*, 17 (9-10): 7-65.
- Christian, B. and Griffiths, T. (2016). *Algorithms to Live By: The Computer Science of Human Decisions*. London: William Collins.
- Cohen, G. A. (1989). On The Currency of Egalitarian Justice. *Ethics*, 99 (4): 906-944.
- Cohen, G. A. (2012). Rescuing Conservatism: A Defense of Existing Value (All Souls Version). In his *Finding Oneself in the Other*, edited by Michael Otsuka. Princeton: Princeton University Press.
- Cohen, J. (2015). Death-defying experiments: pushing the limits of life span in animals could someday help lengthen our own. *Science*, 350 (6265): 1186-1187.
- Correia, S., Dickinson, A., & Clayton, N. (2007). Western Scrub Jays Anticipate Future Needs Independently Of Their Current Motivational State. *Current Biology*, 17 (10): 856-861.
- Crisp, R. (2003). Equality, Priority, and Compassion. *Ethics*, 113: 745-746.
- Danaher, J. (2019). *Automation and Utopia*. Harvard University Press.
- Daniels, N. (1985). *Just Health Care*. Cambridge University Press.
- Daniels, N. (2001). Justice, Health, and Healthcare. *The American Journal of Bioethics*, 1 (2): 2-16.
- Daniels, N. (2007). *Just Health: Meeting Health Needs Fairly*. Cambridge University Press.
- Davis, J. (2018). *The New Methuselahs: The Ethics of Life Extension*. MIT Press.
- Delon, N. (2016). The Replaceability Argument in the Ethics of Animal Husbandry. In: Thompson, P., Kaplan, D. (eds) *Encyclopedia of Food and Agricultural Ethics*. Springer, Dordrecht.

- Dickinson, A. (2011). Goal-Directed Behavior and Future Planning in Animals. In R. Menzel and J. Fischer, eds. *Animal Thinking: Contemporary Issues in Comparative Cognitive*. Cambridge: MIT Press.
- Dorsey, D. (2015). The Significance of a Life's Shape. *Ethics*, 125(2): 303-330.
- Doyle, D. J. (2018). *What Does it Mean to be Human? Life, Death, Personhood and the Transhumanist Movement*. Berlin: Springer.
- Feldman, F. (1991). Some Puzzles About the Evil of Death. *The Philosophical Review*, 50(2): 205-227.
- Feldman, F. (2006). *Pleasure and the Good Life: On the Nature, Varieties, and Plausibility of Hedonism*. New York: Oxford University Press.
- Fletcher, G. (2016). Objective list theories. In G. Fletcher. *The Routledge Handbook of Philosophy of Well-being*. New York: Routledge.
- Frankfurt, H. (1987). Equality as a Moral Ideal. *Ethics*, 98: 21-43.
- Frick, J. (2017). On the Survival of Humanity. *Canadian Journal of Philosophy*, 47(2-3): 344-367.
- Frick, J. (2020). Conditional Reasons and the Procreation Asymmetry. *Philosophical Perspectives*, 34: 53-87.
- Frick, J. and Lerner, A. (ms). Speciesism and Acceptable Lives.
- Glasgow, J. (2013). The Shape of a Life and the Value of Loss and Gain. *Philosophical Studies*, 162(3): 665-682.
- Greaves, H. (2017). Population Axiology. *Philosophy Compass*, 12(11): 1-15.
- Greaves, H. & MacAskill, W. (2021). The Case for Strong Longtermism. *GPI Working Paper* No. 5-2021. Available at: <https://globalprioritiesinstitute.org/wp-content/uploads/The-Case-for-Strong-Longtermism-GPI-Working-Paper-June-2021-2-2.pdf>
- Gregory, A. (2016). Hedonism. In G. Fletcher. *The Routledge Handbook of Philosophy of Well-being*. New York: Routledge.
- Gruen, L. (2021). The Moral Status of Animals. In: E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2021 Edition.
- Harman, E. (2009). "I'll be glad I did it" reasoning and the significance of future desires. *Philosophical Perspectives*, 23: 177-99.
- Harris, J. (1985). *The Value of Life: An Introduction to Medical Ethics*. Routledge.

- Heathwood, C. (2016). Desire-fulfillment theory. In G. Fletcher. *The Routledge Handbook of Philosophy of Well-being*. New York: Routledge.
- Hesiod. (2008). *Works and Days*, translated by M. L. West. New York: Oxford University Press.
- Hirose, I. (2015). Intra-personal aggregation. In his *Moral Aggregation*. Oxford: Oxford University Press.
- Hurka, T. (2016). Love and Reasons: The Many Relationships. In Esther Kroeker and Katrien Schaubroeck, eds. *Love, Reasons, and Morality*. New York: Routledge.
- James, P. D. (1992). *The Children of Men*. London: Faber and Faber.
- Jonas, H. (1985). *The Imperative of Responsibility: In Search of an Ethics for the Technological Age*. Chicago University Press.
- Kagan, S. (2019). *How to Count Animals, more or less*. Oxford: Oxford University Press.
- Kamm, F. (1998). *Morality, Mortality Vol. 1: Death and Whom to Save From It*. New York: Oxford University Press.
- Kamm, F. (2003). Rescuing Ivan Ilyich: How we live and how we die. *Ethics*, 113(2): 202-233.
- Kant, I. (1793/1996). 'On the Common Saying: That May Be Correct in Theory, But Is of No Use in Practice' in his *Practical Philosophy*, translated by Mary Gregor. Cambridge: Cambridge University Press.
- Kass, L. R. (2004). L'Chaim and Its Limits: Why Not Immortality? In Stephen G. Post and Robert H. Binstock (eds.), *The Fountain of Youth: Cultural, Scientific, and Ethical Perspectives on a Biomedical Goal*. Oxford University Press. 304–320.
- Kauppinen, A. (2012). Meaningfulness and Time. *Philosophy and Phenomenological Research*, 84(2): 345-377.
- Kauppinen, A. (2015). The Narrative Calculus. In Mark Timmons (ed.) *Oxford Studies in Normative Ethics*, vol. 5: 196-220.
- Kavka, G. (1978). The Futurity Problem. In R. Sikora & B. Barry, *Obligations to Future Generations*. White Horse Press.
- Kennedy, B. K., Berger, S. L., Brunet, A., Campisi, J., Cuervo, A. M., Epel, E. S., Franceschi, C., Lithgow, G. J., Morimoto, R. I., Pessin, J. E., Rando, T. A., Richardson, A., Schadt, E. E., Wyss-Coray, T., & Sierra, F. (2014). Geroscience: linking aging to chronic disease. *Cell*, 159 (4): 709–713.

- Kolodny, N. (2003). Love as Valuing a Relationship. *The Philosophical Review*, 112(2): 135-189.
- Korsgaard, C. M. (1983). Two Distinctions in Goodness. *The Philosophical Review*, 92(2): 169-195.
- Lange, M. M. (2011). 'Progress'. *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta. Available at: <https://plato.stanford.edu/entries/progress/>
- Lenman, J. (2002). On Becoming Extinct. *Pacific Philosophical Quarterly*, 83: 253-269.
- Lewis, C. S. (2001). *A Grief Observed*. New York: HarperOne.
- Lewis, D. (1986). *On the Plurality of Worlds*. Oxford: Blackwell Publishers.
- Lin, E. (2018). Welfare Invariabilism. *Ethics*, 128(2): 320-345.
- Lin, E. (2022). Well-Being, part 2: Theories of Well-being. *Philosophy Compass*, 17 (2): 1-23.
- MacAskill, W. (2022). *What We Owe the Future*. New York: Basic Books.
- Martin, JA., Hamilton, BE., Osterman MJK., & Driscoll, AK. (2021). Births: final data for 2019. *National Health Statistics Report*, 70(2).
- Marusic, B. (2018). Do Reasons Expire? An Essay on Grief. *Philosophers' Imprint*, 18(25): 1-21.
- McMahan, J. (1981). Problems of Population Theory. *Ethics*, 92(1): 66-127.
- McMahan, J. (1996). Cognitive Disability, Misfortune, and Justice. *Philosophy and Public Affairs*, 25(1): 3-35.
- McMahan, J. (2002). *The Ethics of Killing*. New York: Oxford University Press.
- McMahan, J. (2008). Eating Animals the Nice Way. *Daedalus*, Winter 2008: 1-11.
- McMahan, J. (2013). Causing People to Exist and Saving People's Lives. *The Journal of Ethics*, (17)5: 5-35.
- Mills, M. Rindfuss RR., McDonald, P. & Te Velde, E. (2011). Why do people postpone parenthood? Reasons and social policy incentives. *Human Reproduction Update*, 17(6): 848-860.
- Moore, G. E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press.
- Nagel, T. (1970). *The Possibility of Altruism*. Princeton: Princeton University Press.
- Nagel, T. (1979). Death. In his *Mortal Questions*. Cambridge: Cambridge University Press.

- Nebel, J. (2015). Status Quo Bias, Rationality, and Conservatism about Value. *Ethics*, 125(1): 449-476.
- Nebel, J. (2022). Conservatisms about the Valuable. *Australasian Journal of Philosophy*, 100(1): 180-194.
- Newman, J. C., Milman, S., Hashmi, S. K., Austad, S. N., Kirkland, J. L., Halter, J. B., & Barzilai, N. (2016). Strategies and Challenges in Clinical Trials Targeting Human Aging. *The journals of gerontology. Series A, Biological sciences and medical sciences*, 71 (11), 1424–1434. <https://doi.org/10.1093/gerona/glw149>
- Nisbet, R. (1980). *History of the Idea of Progress*. New York: Basic Books.
- Olson, E. T. (2022). Personal Identity. In: E. N. Zalta, *The Stanford Encyclopedia of Philosophy*, Summer 2022 Edition.
- Ord, T. (2020). *The Precipice: Existential Risk and the Future of Humanity*. New York: Hachette Books.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Parfit, D. (1997). Equality and Priority. *Ratio*, 10 (3): 202-221.
- Parfit, D. (2016). *On What Matters, vol. 3*. Oxford: Oxford University Press.
- Partridge, L., Fuentealba, M., and Kennedy, B. K. (2020). The quest to slow ageing through drug discovery. *Nature Reviews Drug Discovery*, 19 (8): 513-532.
- Portmore, D. (2007). Welfare, achievement, and self-sacrifice. *Journal of Ethics and Social Philosophy*, 2(2): 1-28.
- Pritchard, D. (2005). *Epistemic Luck*. New York: Oxford University Press.
- Pummer, T. (2017). Lopsided Lives. In Mark Timmons (ed.), *Oxford Studies in Normative Ethics*, vol. 7. Oxford: Oxford University Press.
- Scheffler, S. (1982). *The Rejection of Consequentialism*. Oxford: Clarendon Press.
- Scheffler, S. (2007). Immigration and the Significance of Culture. *Philosophy & Public Affairs*, 35(2): 93-125.
- Scheffler, S. (2010). *Equality and Tradition: Selected Essays*. New York: Oxford University Press.
- Scheffler, S. (2013). *Death and the Afterlife*, edited by Niko Kolodny. New York: Oxford University Press.
- Scheffler, S. (2018). *Why Worry About Future Generations?* New York: Oxford University Press.

- Sen, A. (1973). *On Economic Inequality*. Oxford University Press.
- Seneca. (2005). *On the Shortness of Life*. Translated and edited by C. D. N. Costa. Penguin Books.
- Setiya, K. (2016). Retrospection. *Philosophers' Imprint*, 16(15): 1-15.
- Singer, P. (1975). *Animal Liberation: A New Ethics for Our Treatment of Animals*. HarperCollins.
- Slote, M. (1983). Goods and Lives. In his *Goods and Virtue*. Oxford: Clarendon Press.
- Smuts, A. (2017). *Welfare, Meaning, and Worth*. London and New York: Routledge.
- Sosa, E. (1999). How Must Knowledge Be Modally Related to What Is Known? *Philosophical Topics*, 26(1/2): 373-384.
- Temkin, L. (1993). *Inequality*. New York: Oxford University Press.
- Temkin, L. (2003). Equality, Priority, or What? *Economics and Philosophy*, 19: 61-88.
- Temkin, L. (2012). *Rethinking the Good*. New York: Oxford University Press.
- Timmer, D. (forthcoming). Justice, Thresholds, and the Three Claims of Sufficiencyarianism. *The Journal of Political Philosophy*.
- Valentynne, P. (2005). Of Mice and Men: Equality and Animals. *The Journal of Ethics*, 9(3/4): 403-433.
- Van Parijs, P. and Vanderborght, Y. (2017). *Basic Income: A Radical Proposal for a Free Society and a Sane Economy*. Harvard University Press.
- Velleman, D. (1991). Well-being and Time. *Pacific Philosophical Quarterly*, 72(1): 48-77.
- Velleman, D. (1999). A Right to Self-Termination? *Ethics*, 109(3): 606-628.
- Wallace, R. Jay. (2013). *The View from Here: On Affirmation, Attachment, and the Limits of Regret*. New York: Oxford University Press.
- Williams, B. (1973). The Makropulos Case: Reflections on the Tedium of Immortality. In his *Problems of the Self*. Cambridge University Press.
- Williams, B. (2006). The Human Prejudice. In his *Philosophy as a Humanistic Discipline*, edited by A. W. Moore. Princeton University Press.
- Williamson, T. (2000). *Knowledge and its Limits*. Oxford: Oxford University Press.
- Wolf, S. (2010). *Meaning in Life and Why it Matters*. Princeton: Princeton University Press.

World Bank. (2020). Poverty and Shared Prosperity 2020: Reversals of Fortune.
World Bank. <https://doi.org/10.1596/978-1-4648-1602-4>.