# Wild Animal Suffering and Human Responsibility: Essays on the Ethics of Beneficent Intervention in Nature

A thesis submitted for the degree of Doctor in Philosophy

Word count: 64,288

Gary David O'Brien

Blackfriars College, University of Oxford

# Abstract

Over the past five decades, animal ethicists have made tremendous progress in articulating the moral importance of the wellbeing of non-human animals. Their work has naturally focused on those domesticated animals we use for food, research, etc. Comparatively little has been written about the moral importance of the wellbeing of wild animals. This is unfortunate, as wild animals far outnumber human beings and their domestic animals, and they suffer from a great number of harms such as predation, parasitism, starvation, and disease. In fact, because most animals reproduce in vast numbers while their populations remain relatively stable, we can infer that the vast majority of them die before reaching reproductive age (Ng, 1995; Horta, 2010; Faria 2023a). The typical life of a wild animal is to be born alongside dozens, hundreds, or thousands of siblings, to receive no parental care, and to struggle for survival in a hostile world, before succumbing to predation or starvation only a few hours or days after birth.

This is the problem of wild animal suffering (WAS). The central question is whether humans should try to intervene in nature to reduce the suffering of wild animals, and, if so, what interventions should be performed. Positions on the intervention question range from complete prohibition (Regan, 2004; Palmer, 2010; Korsgaard, 2018) right up to massive genetic modifications of wild animals (Pearce, 1995; Johannsen, 2021), deliberately making certain animals go extinct (McMahan, 2010, 2015; Bramble, 2021), and the deliberate destruction of wild animal habitats (Tomasik, 2016).

In this integrated thesis, I assume an interventionist position. I explore the implications of recognizing the scale and severity of WAS, and the moral complexities that arise once we begin to systematically intervene in nature on a large scale. In paper one I critique Milburn's (2022) version of the relational objection to intervention in nature, and I argue that human beings have been morally entangled with wild animals since prehistory. In paper two I argue that the non-identity problem affects the strength and kind of our moral reasons to intervene in nature. In paper three, I argue that WAS is a predictable outcome of the evolutionary process, and this gives us a strong reason not to spread animal life to other planets. In paper four I argue that the basic argument for longtermism applies to animals just as much as it does to human beings, that longtermists ought to take animals much more seriously than they do, and that animal advocates ought to be concerned about the long-term future. In the fifth paper I argue that Totalism has deeply counterintuitive implications when we include animals in the populations to be evaluated.

# Acknowledgements

This has been a long road, and I am very grateful to the many people who helped me along it. First, my supervisors Guy Kahane and Jeff McMahan, who worked tirelessly and generously to turn the crooked timber of my thoughts into something that is (hopefully) worth reading. Both have been my supervisors throughout not only the DPhil, but the BPhil too, and have helped me develop every step of the way. I owe a great deal to them. I would also like to thank Oscar Horta, whose work inspired my own, and who generously gave his time to read and discuss much of my work with me.

For helpful comments and discussion of various parts of this thesis I am also grateful to Rhys Borchert, Mattia Cecchinato, Lorenzo Elijah, Adam Gibbons, Kyle Johannsen, Kida Lin, Beatrice Marchegiani, Josh Milburn, Katie Prosser, Asher Soryl, Rhys Southan, Diego Expósito Teixeira, Yip Fai Tse, and Patrick Williamson. I'm also thankful for comments from anonymous reviewers at the *Journal of Applied Philosophy*, and attendees of the St. Andrews Wild Animal Ethics conference, the Society for the Study of Ethics and Animals, and the Sheffield Animal Studies Research Centre. For partly funding my work I am grateful to the Center on Long-Term Risk, the Royal Institute of Philosophy, and Animal-Ethics. For providing a fun, welcoming, and stimulating work environment I want to thank everyone at the Oxford Uehiro Centre for Practical Ethics.

For supporting me through a long and sometimes difficult process, I am thankful to my family, Martin, Mary, Amy, and Isla and Quentin (my two favourite little animals). I am especially grateful to my mother Una, who supported me throughout my education, from playschool to PhD.

Infine, grazie a Beatrice, l'amore della mia vita, che mi rende felice ogni giorno.

# Contents

# Introduction

## 1 Introduction to Wild Animal Suffering

Human beings are intimately familiar with suffering. From infancy to old age we are vulnerable to a wide range of kinds and degrees of suffering. There is suffering caused by the lack of some physical necessity like food or water, by disease and sickness of the body, by injury or accident, and by physical conflict. We suffer from negative emotions, from the loss of loved ones, from the frustration of our goals and desires, from humiliation, guilt, and shame, from despair and from psychological conditions like depression and schizophrenia. Inevitably, we suffer as our bodies age and die. Much labour and ingenuity has been expended in mitigating suffering and in the attempt to shield ourselves from the worst of it, and this species-wide project has been remarkably successful. We have taught ourselves to farm food, to provide ourselves with abundant fresh water and to build elaborate shelter from the elements. We have made incredible advances in medical technology, and in the creation of labour-saving devices to make our lives easier and more pleasant. We are even learning, albeit slowly, how to interact with each other in peaceful ways, and to treat each other justly. There is a long way to go before we can free ourselves totally from all useless forms of extreme suffering, but a brief glance at the history and prehistory of our species should give us cause for optimism.[1]

Of course, humans are not the only beings who suffer. Other animals are also vulnerable to many of the same causes of suffering as we are. It is a grim irony that as our technology has developed, and we have freed ourselves more and more from the yoke of suffering, we have simultaneously devised and implemented new ways of inflicting suffering

---

[1] The case for an optimistic interpretation of human history, and optimism about our future, is made by Pinker (2011) and Ord (2020). In his ambitious 'Hedonistic Imperative' (1995), Pearce outlines a manifesto for the abolition of suffering from all sentient beings through the use of advanced technology.

on other animals for our own benefit. We have created a vast system of industrialized animal agriculture that confines and kills tens of billions of animals each year, in ways in which, were they done to humans, we would consider to be the very basest kind of brutality and barbarism. The number of animals involved, and the extremes of suffering that they endure, make it overwhelmingly likely that the suffering of domesticated and exploited animals far outweighs the suffering of human beings.[2] Even here there is some room for optimism – in the last few decades, significant numbers of people have woken up to the injustice of our treatment of our fellow creatures. We have begun, slowly, to take the suffering and wellbeing of other animals seriously. As more and more people turn away from using animals, we can begin to hope that our exploitation and brutalization of other animals will one day come to an end.[3]

Even if animal use were utterly abolished, however, the majority of the suffering in the world would remain. Most of the suffering in the world is neither endured, nor caused, by human beings. At the time of writing the global human population is around 8 billion, and we kill around 60-70 billion land animals for food each year, and 1-3 trillion wild fish. Wild animals, however, far outnumber domesticated animals and humans. Tomasik (2009a) surveys the existing scientific literature on wild animal populations and produces the following estimates for their total numbers: mammals 100 billion to 1 trillion; birds 100-400 billion; reptiles 100 billion to 100 trillion; amphibians 100 billion to 100 trillion; fish 10 trillion to 1 quadrillion.[4] Of course, there is a huge amount of uncertainty in these estimates,

---

[2] According to data compiled by Faunalytics (Sanders, 2018), more than 70 billion land animals are killed for food annually, and Fishcount (2019c) estimate that between 1 and 3 trillion wild fish are caught each year. See Animal-Ethics (2023a) on the brutal practices employed in animal use industries.
[3] Singer's (1975) *Animal Liberation* has been seen by many as the book that brought animal ethics into the mainstream. For suggestive statistics about the growth in popularity of veganism see Minassian (2022) and Anthony (2021).
[4] Throughout this thesis I will almost exclusively discuss vertebrate animals. This is for epistemic reasons – we can be very sure that most vertebrates are sentient, and that they can suffer in morally relevant ways. It is less clear if many invertebrates can also suffer (though some, such as cephalopods, very likely can). The case for the

but even taking the lowest estimate for each category it is clear that wild animals far outnumber domesticated animals and humans. These wild animals are vulnerable to many of the same causes of suffering as humans and domesticated animals are. Their suffering, for the most part, is the result of purely natural processes independent of human agency. It is undeniable that there is a vast amount of suffering in nature, and, as we will see below, likely that the vast majority of wild animals on Earth have lives that are not worth living. This is the problem of wild animal suffering (WAS).

I take it as axiomatic that suffering is bad, and, in general, if we are able to reduce it without sacrificing too much, we ought to. Faria (2016, 2023a) builds her case for beneficent intervention in nature to reduce WAS on these two assumptions, and argues that they apply to animals as well as humans. If we accept that sentient nonhuman animals are morally considerable, then these simple assumptions have important implications for our moral reasons to help non-human animals. If it is true that many animals are suffering in nature, and if it is possible for us to do something to significantly reduce that suffering without sacrificing too much, then we ought to do so. This may seem highly counterintuitive to many people, and yet it follows from two relatively uncontroversial moral principles, and basic empirical facts about the lives of wild animals.

This integrated thesis is an exploration of what recognition of the importance of WAS entails for human beings. I do not argue the case for intervening in nature to reduce WAS, as I take this to have been convincingly demonstrated already.[5] Rather, I explore some of the implications of recognizing that there is so much suffering in nature, and the moral

---

importance of WAS is strong enough if we consider only vertebrates. If it turns out, as it well may, that many invertebrates also suffer in morally significant ways, WAS becomes orders of magnitude more important.

[5] The case for intervention has been most systematically made by Faria (2016, 2023a). It has also been made by Ng (1995), Pearce (1995, 2009), Cowen (2003), Tomasik (2009b), Horta (2010, 2017), McMahan (2010, 2016), Moen (2016), Johannsen (2017, 2020, 2021) and others.

complexities that arise when we begin to intervene in nature on a large scale to reduce WAS. This includes questions about the conditions under which we are, or could become, responsible for the apparently natural harms endured by wild animals; how the non-identity problem affects our reasons to intervene in nature on a large scale; whether we should spread animal life beyond Earth; why we should think about the wellbeing of animals over the very long-term future; and how the inclusion of animals affects our population ethical thinking. In the rest of this chapter I will briefly outline the main causes of WAS, and the arguments for and against intervention in nature to reduce that suffering. Then I will outline the content of the papers that follow. Though all the papers are independent of each other, they are united by the central theme. Taken together, I hope that they will increase our understanding of what we owe to the wild nonhuman animals with whom we share this world.

## 2 Causes of WAS

In this section I will briefly describe some of the causes of WAS. Like domesticated animals, wild animals often suffer because of humans. Sometimes they suffer directly, due to fishing, hunting, and destruction of habitats. Other times, they suffer indirectly, from harms such as pollution and climate change. As important as these anthropogenic harms are, in this thesis I will mostly put them aside, and focus on 'naturogenic' harms, that is, those harms that do not involve human agency either directly or indirectly. As we will see in several of the papers in this thesis, sometimes the lines between anthropogenic and naturogenic harms can become blurred, and naturogenic harms can even become 'transformed' into anthropogenic ones.[6]

---

[6] See especially paper 2, on how naturogenic harms can be 'transformed' into anthropogenic ones, and paper 1 for some examples of ways in which human beings may have already become responsible for some 'natural' harms.

Nonetheless, it is still a useful distinction for categorization of the harms endured by wild animals.

Naturogenic harms are practically important, as they are the largest source of harms to nonhuman animals. As we will see below, most animals die either from predation or from starvation shortly after birth. Given the sheer numbers of wild animals who die in these natural ways, it is likely that most suffering in our biosphere is naturogenic. If we wish to significantly reduce the suffering of sentient beings on our planet, we must eventually contend with these naturogenic sources of harm. Furthermore, naturogenic harms have been around a lot longer than anthropogenic ones, and, if humanity ever goes extinct, we can expect that naturogenic harms will continue for a long time. Compared to anthropogenic harms, naturogenic ones are potentially extremely persistent – while anthropogenic harms require the continued existence of human beings, naturogenic harms can continue for a very long time, so long as animal life continues. If we are concerned about the wellbeing of animals over the long-term future, we have reason to better understand naturogenic harms.

Naturogenic harms are also interesting theoretically. Though there are complexities, the anthropogenic harms we inflict on wild animals are, like the harms we inflict on domesticated animals, things we could simply refrain from doing. At least in paradigm cases (factory farming, recreational hunting, etc.) it is clear that we are inflicting serious harms for trivial reasons, and there is a very strong moral case for us to stop doing so. The case of naturogenic harms is less clear, morally speaking. We are not responsible for naturogenic harms, nor can we solve the problem of WAS by simply ceasing to engage in harmful activities. Rather, we must intervene in ancient natural systems which we had no part in creating, and in which some of us find a kind of sacred, inviolable value. Intervention to reduce WAS also involves difficult trade-offs, when the interests of some animals conflict

with those of others, as in the case of predation. The moral complexities of intervening in nature to reduce naturogenic harms raise interesting questions which I will discuss throughout this thesis.

We can divide the proximate causes of WAS intro three broad categories – antagonism, scarcity, and misfortune. In practice it will not always be clear to which category a given harm belongs, but it is nonetheless a useful framework for categorizing harms and understanding their ultimate causes.[7]

(1) Antagonism

In ecology, antagonism is an association between organisms in which one benefits at the expense of another.[8] This is because the interests of one animal are in conflict with those of another. Examples of antagonism in nature are predation, parasitism, parasitoidism, conflicts over territory or mates, sibling rivalry, and sexual conflict.[9] Predation and parasitism are extremely common in nature, and cause a huge amount of suffering for wild animals.

(2) Scarcity

Scarcity is when there are insufficient resources in an environment for all animals to survive. Scarcity is, unfortunately, endemic in nature, as most animal species reproduce far beyond the carrying capacity of their environments, thus dooming many of them to death by starvation, thirst, or exposure through lack of shelter from the elements.

---

[7] In paper 3 (section 2.4) I go into greater detail on antagonism and scarcity and argue that they are necessary features of evolution by natural selection, and so the suffering caused by them is predictable.
[8] I have phrased this definition in terms of individual organisms and their interests, as I am interested in the wellbeing of animals, and it is, of course, individuals, not species or populations, that can have positive or negative wellbeing. Eaton (2008), in the Encyclopedia of Ecology, defines antagonism in terms of populations, not individuals, and fitness rather than interests: "Antagonism is an interaction between two populations that positively affects the fitness of one population and negatively affects the fitness of the other population".
[9] See the Animal Ethics (2023c) pages on Antagonism in nature: Interspecific Conflict, Intraspecific Fights, and Sexual Conflict for examples and discussion.

## (3) Misfortune

Under misfortune I include things like animals becoming injured in accidents, dying in natural disasters, becoming trapped, or succumbing to disease.[10] Many of these causes of suffering are not exclusively the result of misfortune, however, as scarcity and antagonism make animals more prone to certain harms. For example, malnutrition, a result of scarcity, may make an animal more vulnerable to disease. Misfortune can also make animals more vulnerable to other harms. For example, if an animal becomes injured through accident, they are more likely to be eaten by a predator, etc.

Specific examples of instances in which animals suffer from each of these causes are easily found, and I will not include them here.[11] In any case, a simple catalogue of harms endured by wild animals may shatter the illusion that life in the wild is idyllic, but it is insufficient to show how serious a problem WAS actually is. After all, a similar catalogue of harms endured by human beings could easily be produced, and yet most of us believe that human life tends to be very good, despite the extensive list of harms to which we are vulnerable.[12] It could also be the case that, despite their vulnerability to harms, most wild animals manage to live good lives. What is required in order to know the seriousness of WAS is the extent of suffering in nature, what proportion of animals live bad lives, and whether there is more suffering or positive wellbeing in nature as a whole.

---

[10] Of course, many diseases are caused by bacteria or viruses, and so these might be better classed as instances of antagonism. Bacteria and viruses are presumably insentient, and so they have no interests which can come into conflict with the interests of sentient animals. Since the 'antagonistic' interactions between sentient animals and insentient bacteria and viruses do not involve conflicting interests, I have classified disease as a case of misfortune rather than antagonism. Nothing important hangs on the classification, however.
[11] See Chapter 3 of Faria (2023a), Animal Ethics (2023c), part 1 of Animal Ethics (2020), and Tomasik (2009b).
[12] Though see Benatar (2006), Zapffe (1933), and Schopenhauer (1851) for defences of a pessimistic appraisal of human life.

Attempts to answer this question have appealed to population dynamics, and their effects on the wellbeing of wild animals.[13] There are, very broadly speaking, two reproductive strategies adopted by animal species on Earth, and which strategy a species has evolved to have makes a huge difference to the average welfare levels of individuals of that species. Some animals, like humans, give birth to a small number of offspring, in whom they invest a great deal of energy, either in the form of parental care or in giving birth to better developed offspring. Consequently, these animals have relatively high survival rates. We can expect animals like this to have fairly decent lives. Though they may suffer in many ways, and are likely to die painfully, they have sufficiently long lives in which they can also enjoy positive experiences, such as socializing, feeding, and mating. This strategy is known as K-strategy, and the animals who utilise it are K-strategists.

The other strategy, adopted by the vast majority of animal species, is to have large numbers of offspring in each of whom parents invest very little energy. Each individual offspring is very unlikely to survive, but, if parents have enough of them, it is likely that at least one per parent will survive to reproduce themselves. This strategy is called r-strategy, and the animals who utilise it are r-strategists.[14] Some animals can have thousands, or, in the case of reptiles, amphibians, and fish, even millions of offspring.[15] Of course it is not the case that all of these animals survive – if they did then the population of animals would increase exponentially every generation. Rather, what happens is that almost all of them die before they reach reproductive age. Often, they will die in very painful ways such as being eaten by predators, or dying of starvation or thirst. Since they live for only a short time, and endure

---

[13] This argument first appeared in Ng (1995) and has since been used by Tomasik (2009b), Horta (2010), Faria (2016, 2023a), Moen (2016), Johannsen (2017, 2021) and others.

[14] See MacArthur and Wilson (1967) and Pianka (1970) for the original formulation of this theory. Though some elements of the theory have been criticized by ecologists, the general distinction between the two reproductive strategies, and their implications for the wellbeing of animals, remain useful. See Reznick et al. (2002) for further discussion of the validity of the r/K distinction.

[15] See Ng (1995) p.270, Horta (2010) p.10, and Johannsen (2017) p.337-338 for examples.

painful deaths, it is reasonable to suppose that their lives are not worth living. That is, that their lives contain more things that are bad for them (such as pain and suffering) than things that are good for them. Since this is the fate of the vast majority of individual r-strategists, and since the vast majority of animals are r-strategists, we can infer that the vast majority of animals don't have lives worth living. We can call this:

> *The Net Bad Lives* claim – For all animals on born on Earth each year, the majority (or, perhaps, the vast majority) have lives that are overall bad for them.

Of those who accept the net bad lives claim, some use it to argue for the stronger:

> *Net Suffering* claim – Nature as a whole contains more experiences of suffering than of happiness or enjoyment.[16]

At first glance, these claims might look to be the same. It seems that, if the vast majority of lives contain more suffering than happiness, then it must also be the case that nature as a whole contains more suffering than wellbeing. It could be the case, however, that the animals who live generally good lives live long enough, and have good enough experiences, that the goodness in their lives outweighs the badness in the lives of short-lived animals. A weaker claim is made by Johannsen (2021). He argues for:

> *Net Failure to Flourish* – The majority of animals born on Earth do not live flourishing lives.

This claim is weaker than the net bad lives claim. Johannsen doesn't claim that most animals have bad lives, in the sense of their lives containing more of what is bad for them

---

[16] This argument appears in Ng (1995), Tomasik (2009b) and Horta (2010, 2015, 2017). I have followed these authors in expressing this claim in hedonic terms, but this isn't essential. A similar claim could plausibly be made using different conceptions of wellbeing, for example, that nature as a whole contains more frustrated than satisfied desires. For critical discussion of the Net Suffering claim see Browning and Veit (2023), Gardner (2022), Groff and Ng (2019) and Plant (2016).

than of what is good for them. Rather, he claims, plausibly, that "a life that's filled with suffering and ends shortly after birth is not a flourishing one" (2021, p.2). Most animals live short lives which end in a painful, or a very painful death, and these deaths occur early in life, before the animal ever has a chance to become comfortable with, or competent to manage, their environments (p.1). Net Failure to Flourish isn't as bad as Net Suffering or Net Bad Lives, but might still generate significant moral reasons to intervene in nature to try to improve the lives of wild animals. Furthermore, as I will argue in one of the following papers, when we create sentient beings, providing them with a reasonable chance of flourishing may be a reasonable threshold for permissible creation.

Finally, and most weakly, as it makes no claims about the net balance of suffering or about bad or non-flourishing lives, we could endorse:

*Mass WAS* – there is a huge amount of suffering in nature.

This doesn't claim that there is net suffering in nature or that most lives are bad, but simply that there is a huge amount of suffering in nature. How much exactly can be left open, but, given the numbers of animals involved, it is reasonable to claim that (i) the amount of WAS exceeds the amount of suffering generated on our factory farms and (ii) this suffering is sufficient to be an extremely important moral issue.

These claims are not always kept apart in the literature, and it is not always clear which claim is being defended. To some extent, this is because, in many contexts, it doesn't matter which of these claims is true. So long as even the weakest of them is true, then we have moral reasons to try to reduce the suffering that we find in nature, even if it is not the case that suffering prevails over happiness, or that bad and/or non-flourishing lives are the majority. For example, after defending the net suffering claim, Horta (2010, p.14-15) acknowledges that the suffering of wild animals would give us reason to intervene in nature

to mitigate it, regardless of whether there is net suffering. Faria (2023a, p.64-65) argues that, even if it turns out that the net suffering claim is false, the normative implications of this would be limited. She points out that, in the case of human beings, we don't require that some group is experiencing net negative lives before we consider it morally obligatory to help them. Similarly, the huge amount of suffering experienced by wild animals gives us sufficient reason to help them, even if, on balance, there is more wellbeing than suffering in nature as a whole.

There are, however, some contexts in which these claims will have different normative implications, and so it is important to be able to distinguish the different claims clearly. For example, if one endorses the net suffering claim, and one is also a utilitarian, then it makes sense to try to reduce wild animal habitats as much as possible, as this will very effectively reduce suffering.[17] Some of the arguments in this thesis need only the Mass WAS claim to be true. Arguments in some of the papers, however, require a stronger claim to be true. In the third paper, for example, I argue on consequentialist grounds that we shouldn't pursue a particular way of spreading animal life beyond the Earth, for the reason that this would result in greater net suffering. This argument depends on either the Net Suffering or the Net Bad Lives claim being true, depending on the form of consequentialism endorsed. When I refer to WAS I will generally assume only that the Mass WAS claim is true, but when I rely on stronger claims I will indicate this.

---

[17] See Tomasik (2016) for promotion of this idea, and Johannsen (2020, p.53-62) for critical discussion.

## 3 The Case for Intervention

The case for beneficent intervention in nature to reduce WAS is very straightforward on Utilitarian grounds. Wild animals suffer, and there seems to be much we can, or could, do to reduce that suffering. If this is true, and the possible interventions available to us are likely to be effective, then there is a very strong case for intervention.[18] Furthermore, Utilitarianism is, in Sebo's (2023b) terms, a highly unrestrictive moral theory, meaning that we ought to do the most good by any means necessary. It also has a monistic theory of value. Taken together, these features of Utilitarianism open up many possibilities for intervention that might not be countenanced by non-consequentialist theories, such as deliberate destruction of wild animal habitats (Tomasik, 2016) or the painless killing of predators (Bramble, 2021).

It is not only utilitarians, however, who must answer the moral demands raised by WAS. As Faria (2023a, p.186) points out, non-consequentialists can also accept the basic normative claim that we ought to help others in need when we can do so without sacrificing too much. That is, they can (and should) endorse a principle of beneficence, that there is a prima facie duty to do good for others.[19] Johannsen (2021, chapter 3) also bases his case for a collective obligation to intervene in nature on a principle of beneficence. Pro-interventionists who are not utilitarians may reject some specific forms of intervention on the grounds that they are contrary to other values, or because they violate the negative rights of animals not to be deliberately harmed.

In the papers in this thesis I remain agnostic about which normative theory is correct. Sometimes I will make arguments that are broadly consequentialist, though I do not make any suggestions for interventions that would violate the basic negative rights of sentient

---

[18] Tomasik (2009b) and Pearce (1995) argue for intervention on explicitly utilitarian grounds. See Hills (2010) for a critique of Utilitarianism on the grounds that its commitments to wild animals make it too demanding.
[19] See for example Ross (1930/2002 p.21).

beings. In other papers I will make arguments that are decidedly non-consequentialist. This is partly due to my own normative uncertainty. I also believe, however, that, most plausible moral theories will converge on some form of interventionism. This is partly because I take it that the principle of beneficence is sufficient to ground a large-scale programme of intervention in nature, and beneficence is likely to be a feature of any sensible moral theory. In some papers I will also appeal to considerations independent of beneficence, such as rectificatory justice (paper 2) and special duties of creators towards their sentient creations (paper 3).

There are many ways we can, or could, intervene in nature to reduce WAS. Which ones we choose depends on the technology and scientific knowledge available to us, on our moral theory, and on which of the claims about the prevalence of WAS we accept. Already, there are some limited ways in which we help wild animals. For example, we go to considerable lengths to rescue whales who have become beached, or trapped beneath ice; we vaccinate wild animals against certain diseases; and we help animals affected by large-scale natural disasters such as forest fires.[20]

These are relatively uncontroversial examples of ways of helping wild animals. They do not depend on any particular moral theory, nor do they require any new technology or scientific knowledge. Other proposals in the literature are both more controversial, and more speculative. These include: the 'policing' of nature to prevent animals from harming each other (Cowen, 2003); eliminating predators (McMahan, 2010, 2016) (Bramble, 2021); genetically altering or 'reprogramming' predators so that they are willing and able to subsist on plants (Pearce, 2009) (Johannsen, 2021); turning the wilderness into a compassionately managed wildlife park (Moen, 2016); eliminating wild animal habitats altogether (Tomasik,

---

[20] See Horta and Teran (2023) and the Animal Ethics (2023b) page on helping animals in the wild.

2016); genetically modifying r-strategist animals so that they act more like K-strategists, that is, by having fewer offspring but taking better care of them (Johannsen, 2017, 2021) or to remove their capacity to suffer (Johannsen, 2021); or fundamentally rewriting the genetic code of all vertebrate animals in order to replace the motivational structure of pleasure/pain with a structure based on gradients of pleasurable states (Pearce, 1995).

At this stage of our technological development and scientific understanding, many of these interventions are technically impossible, and the possible ones might have devastating unforeseen side effects. It may be that the best thing we can do for animals at present is to engage in serious research in order to identify and develop safe and effective ways of intervening in nature to help them. Several authors have argued that we should start a new interdisciplinary field of research called 'Welfare Biology'.[21] This field of research would investigate basic questions about the wellbeing of wild animals, such as, which animals are sentient, how much do animals suffer in various ecosystems, etc., as well as developing effective means of improving wild animal welfare. In this thesis I do not propose any new ways of directly intervening in nature to reduce WAS. I do, however, consider some of the moral complexities that arise when we do intervene in nature (see paper 2). I also suggest some indirect interventions that can be expected to reduce expected WAS over the long-term future (paper 4) and I investigate how population ethical questions might affect our decisions about what sort of population we should aim at as the end result of our interventions (paper 5).

---

[21] See Ng (1995), Faria and Horta (2019), Soryl et al (2021), and Faria (2023a).

**4 Objections to Intervention**

Faria (2016, 2023a) has already provided an extensive taxonomy of the various objections to beneficent intervention in nature, and has responded convincingly to them. Her exhaustive treatment of the many objections to intervention in nature comes to almost ninety pages of her book. Others, such as Johannsen (2021) and Horta (2010, 2017), have also responded to many of the arguments against intervention. In this section then I will very briefly outline only the main objections that have been raised against interventionism, and indicate what I believe to be the best responses to them.

(1) Fallibility

Fallibility objections are those which agree that, while in principle in would be good to intervene in nature to reduce WAS, in practice we cannot succeed in this goal. In fact, it may be that if we attempt to intervene in the large-scale systematic ways that would be required to meaningfully reduce WAS we could end up making things even worse. Singer (1973) held this view, but has since been persuaded that some forms of intervention in nature might be effective at reducing suffering (2023). The most detailed defence of this view is that of Delon and Purves (2018). They argue that the complex, often indeterminate nature of ecosystems means that we have no reason to believe that our interventions would make things better rather than worse. As Faria (2023a, p.169-170) has pointed out, however, it is not in theory impossible for humanity to acquire the requisite knowledge to intervene in nature safely. Suitably designed small-scale pilot projects, such as vaccination of a small number of animals in a geographically confined area could help us develop the knowledge and skills required to intervene safely on larger scales, as well as directly helping the animals targeted by the projects.

Even if Delon and Purves are correct, and it is impossible in principle for humanity to develop the knowledge and technology to intervene in nature without risking unforeseen side effects, it isn't clear that this means we have decisive reason not to intervene. They argue that large-scale interventions with unpredictable results are as likely to increase suffering as to decrease it. Of course, this depends on how bad things are now. Delon and Purves point out that risky interventions in well-functioning systems are more likely to do harm than good. But Earth's biosphere isn't a well-functioning system as regards animal welfare. There is reason to believe that suffering far outweighs wellbeing among Earth's animals, i.e. that the Net Suffering claim is true. If the welfare of wild animals is generally very low or negative, then it isn't obvious that interventions are as likely to increase as to reduce suffering. If things are already about as bad as they can get, then the risk of intervention making things even worse is likely to be lower.

Furthermore, given the high level of suffering in nature, we may be justified in taking risks. As an analogy, imagine two patients suffering from disease X. Patient 1 suffers only very mild effects from the disease, while patient 2 suffers terribly. There is a surgical procedure that can be performed to cure the patients, but it is risky. Depending on the level of risk involved, the doctor may be justified in operating on patient 2 despite the risk, while at the same time not being justified in operating on the relatively healthy patient 1. Patient 2's welfare levels are already so low that the risk posed by the operation is a rational one to take. Patient 1 suffers very little, so the risk may not be worthwhile for him. The same is true of large-scale interventions in the biosphere. If we have reason to believe that suffering is extremely bad, then we are more justified in launching risky interventions than we would be

if the level of suffering was less bad. The threshold for justifying risky interventions depends not only on the level of risk, but on how well or badly things are going in the biosphere.[22]

(2) Animal autonomy/competence

Regan (2004) argued that, even if it is the case that wild animals suffer terribly from naturogenic causes, our duty to them is simply to leave them alone. He suggests two arguments for this conclusion. First, our duty to prevent injustice is much stronger than our duty to prevent mere misfortune. Since animals aren't moral agents, they cannot be guilty of unjustly violating the rights of other animals they harm. This means that our reasons to intervene to prevent animals from harming each other are much weaker than our duties to prevent moral agents from doing so. Secondly, he argues that, despite the hardships they endure, wild animals are competent to manage their own affairs, and neither need nor desire our assistance.

In response to his first point, even if we accept the claim that our reasons to prevent injustice are stronger than our reasons to prevent misfortune, this doesn't imply that we have no reason to prevent misfortune, and each wild animal who suffers needlessly and dies prematurely from naturogenic causes is certainly the victim of a serious misfortune. When humans suffer great harms, we don't stop to ask whether those harms have resulted from injustice or from mere misfortune – we help them. If one were to see a child drowning in a pond, one wouldn't try to discover whether the child had been pushed or had simply fallen in before deciding whether to rescue them. The same ought to be true when animals suffer harms. Regarding Regan's second point, he may well be correct that mature animals are often competent to manage their own affairs. Though their lives are difficult and full of risk, it is reasonable to suppose that they are at least somewhat competent to manage their own lives.

---

[22] See Johannsen (2021, p.61) for a similar argument.

The same is certainly not true of infant animals, however. As we have seen, the majority of WAS seems to come from the painful deaths of r-strategist infants who die in their trillions. It is simply false to say that these animals are competent to manage their own affairs, nor do they have parents who are able to protect them. The animal competence objection fails once we realise that most animals in the world are incompetent infants.

(3) Relationality

Relational objections to intervention are those which argue that we are usually not in the right kind of morally relevant relationships with wild animals to make intervention in nature obligatory, or even permissible (Faria, 2023a, p.126). Palmer (2010) argues that there is a relevant moral difference between most domesticated animals and most wild ones. In the case of domesticated animals we have deliberately caused them to be vulnerable and dependent in various ways. This means that we are 'morally entangled' in their lives, in ways which generate positive duties to help them when they need it. This kind of relationship is generally not present between humans and wild animals, and so we have merely negative duties to them. While there are some cases in which we are morally entangled with wild animals, such as when human caused climate change causes harm to them, in most cases of harm endured by wild animals there is no such relationship, and so we are not obligated to intervene on their behalf.

It has been pointed out by Donaldson and Kymlicka (2011, p.162) that this view has two important flaws. First, it weakens our duty to help human beings in need if those humans are not in any kind of special relationship with us. Second, this objection merely shows that intervention in nature is not obligatory. Opponents of intervention, however, typically want to make the stronger claim, that intervention is impermissible. Donaldson and Kymlicka argue instead that we ought to think of wild animals as forming sovereign political communities.

Wild animals are precisely those who have 'voted with their feet' and rejected human intervention in to their lives. Furthermore, though the communities they have formed seem brutal to us, they provide the context in which their members can flourish, or fail to. If we intervene in the natural cycles of predation, parasitism, and food shortages endemic in nature, we run the risk of undermining the environmental context that these animals require to flourish.

It isn't entirely clear what Donaldson and Kymlicka mean by 'community' in this context. This would most naturally be interpreted to mean a voluntary grouping of animals, such as a pack of wolves or a herd of deer. Alternatively, it could be understood in an ecological sense, in which all animals who interact in a particular ecosystem are part of an ecological community. On either interpretation their sovereignty model runs into difficulties. If we interpret community in the voluntary sense, then, since some groups of animals predate upon others, it is not the case that the prey animals are sovereign – their communities are under attack from without, and it is generally justifiable to intervene to protect one sovereign community from attack by another. If we interpret community in the ecological sense, then it is clear that these communities do not provide their members with what they need to flourish, as these communities are structured in such a way that some members exist only to become food for other members. Coupled with the fact that most members of wild animal communities are either eaten or starve to death shortly after birth, these communities are more like 'failed states' than functional sovereign states, and so humanitarian intervention may be justified (Horta, 2015; Milburn, 2016).

**5 Outline of Papers**

In this section I will briefly describe the contents of the five papers which make up the integrated thesis. For each paper I will describe my aims, the main conclusions reached, and how the paper relates to the overall subject of the thesis. Then, in section 6, I will highlight key themes which run through the papers, and show how they relate to each other. Though the papers are independent of each other, taken together they form a coherent and substantial contribution to the literature on wild animal ethics.

(1) On Our Moral Entanglements with Wild Animals

The first paper investigates a form of the relationality objection to intervention in nature. As we saw in section 4, relational objections are those which state that intervention in nature is either not required, or, more strongly, not permitted, because we do not have the right kinds of morally relevant relationships with wild animals. The most prominent proponents of this objection are Palmer (2010) and Donaldson and Kymlicka (2011). I have chosen to engage with Josh Milburn's (2022) version of the relationality objection, for two reasons. First, unlike the relational accounts of Palmer and Donaldson and Kymlicka, Milburn's account has not yet received much attention in the literature. Second, Milburn's account has features which might illuminate the nature and extent of our moral reasons to help wild animals even if it is not the case that particular kinds of moral entanglements are generally required for intervention in nature. Milburn argues that all humans have positive duties towards all domesticated animals, but we generally only have negative duties towards wild animals. Our positive duties towards domesticated animals are grounded on the actions of our ancestors who domesticated animals. For Milburn, humanity forms a morally relevant collective, and modern humans can have moral responsibilities because of the actions of humans in the distant past. If this is correct, then it also seems that modern humans could have positive

duties towards wild animals if our ancestors acted in ways that make us sufficiently morally entangled with wild animals too.

My aim in this paper is to show that this is in fact the case. That is, if we accept that modern humans can have positive duties towards domesticated animals because of the actions of our ancestors, the same applies to wild animals. I outline Milburn's version of the relational objection. I then argue that our ancestors acted in ways which morally entangled them, and us, in the lives of wild animals. By changing the species composition of the planet, changing ecosystems, taking over and transforming the land, and by introducing species into new environments, historic and prehistoric humans made wild animals significantly vulnerable to new harms and challenges, in ways that make us morally responsible for them. Furthermore, the scale and pervasive nature of these 'morally entangling' actions mean that we have positive duties towards many more wild animals than previously thought. If the arguments in this paper are successful, then, it shows that, even on a relational account, it may be that our positive duties towards wild animals are very extensive. Even on a relational account of our moral duties towards animals, it may be the case that large-scale intervention in nature is not only permissible, but obligatory.

Furthermore, the conclusion of this paper may be relevant for those who already endorse intervention in nature on other grounds. For example, Faria (2016, p.xi; 2023a, p.1) and Johannsen (2021, p.29) both ground our reasons to intervene in nature on beneficence. Johannsen (p.29, p.41-47) explicitly states that, while the principle of beneficence is the best grounding for an interventionist position, considerations of distributive and rectificatory justice are also applicable. Johannsen's position in particular is compatible with the claim that, while beneficence is sufficient to ground an interventionist programme, our reasons to intervene can be even stronger if we take considerations of rectificatory justice into account too. If my arguments in this paper are correct, then we do have such justice-based reasons to

intervene in nature, and not just in exceptional cases such as harms caused to animals by climate change. This strengthens the case for intervention significantly. For those who endorse intervention on utilitarian grounds, however, these considerations may be redundant.

(2) Beneficence, Non-Identity, and Responsibility: How Identity-Affecting Interventions in Nature can Generate Secondary Moral Duties

In the second paper I assess Kyle Johannsen's claim that the non-identity problem is merely a theoretical problem which doesn't make any difference to our reasons to intervene in nature. I argue that this is incorrect. By intervening in nature on a large-scale we change which animals come into existence. Many of those animals will experience more suffering than wellbeing during their lives, and others will inflict significant suffering on other animals either by killing them or by reproducing in very large numbers. Since these bad effects are foreseeable, and since animals are not moral agents and cannot be held responsible for the harms they inflict on each other, I argue that human beings must be responsible for these harms. That is, by causing animals to exist with bad lives, or by causing harmful animals to exist, we can become responsible for the resulting harms.

If this is correct, then by engaging in such identity-affecting actions we can take on secondary moral duties towards and regarding those animals we have caused to exist. These secondary duties can be even more costly than the original interventions in nature to reduce suffering. Furthermore, they may be duties of justice, not merely of beneficence, and so they may be stronger. This counter-intuitive implication follows naturally from some independently plausible claims about moral responsibility for the actions of animals which Johannsen explicitly endorses. If my arguments in this paper are correct, then large-scale beneficent intervention in nature is much more costly than generally thought, to the extent that it may be permissible to refrain from intervening in nature in these ways. However, the

scale and severity of suffering in nature, and the fact that we are already significantly entangled in the lives of wild animals may mean that intervention is still mandatory. The more we intervene, however, the more responsible we become for the overall state of the biosphere. Eventually, it may be the case that 'the wild' no longer exists as something independent of human choice. Management of the biosphere, and the mitigation of animal suffering, then becomes a matter of justice.

(3) Directed Panspermia, Wild Animal Suffering, and the Ethics of World-Creation

The third paper is an investigation of the ethics of spreading life to other worlds by directed panspermia. This process involves seeding suitable planets with microbes, which will, over evolutionary timescales, give rise to a complex biosphere. I argue that the problem of wild animal suffering gives us strong reason not to launch such missions. I provide two independent arguments against spreading life in this way. The first is a consequentialist argument, based on the predictability of mass suffering in unmanaged biospheres. I argue that evolution by natural selection has structural features which make it likely that there will be net suffering on any planet on which sentient life evolves without oversight by intelligent beings. If this is correct, then the mass suffering of animals on our own planet should not be seen as an evolutionary fluke; rather it should be seen as the predictable result of the evolutionary process. If we can expect suffering to prevail on any other world on which life evolves, then we have strong consequentialist reasons not to spread life in this unguided way. According to some versions of consequentialism, this will be true even if only the weaker Net Bad Lives claim is true.

The second argument is a non-consequentialist one which considers whether our role as creators of sentient beings grounds any special duties towards them. I compare creation of sentient beings by directed panspermia to procreation. Using insights from procreative ethics,

I argue that we owe certain duties of care to our sentient creations, in virtue of having created them. At a minimum, we owe it to our sentient creations to ensure that they have lives worth living. It is also plausible that we have a stronger duty, to give our creatures a reasonable chance of flourishing. In the procreative context it is usually possible to fulfil these duties, and, in such cases, procreation is permissible. In the case of directed panspermia, however, it is impossible for us to satisfy our duties to our creations, and so this is an impermissible form of creation.

The paper contributes to the literature on WAS in several ways. First, it makes explicit the case that mass suffering in nature is not a contingent fact about how life has happened to evolve on our planet. Rather, it analyses the root causes of WAS and argues that they are predictable features of evolution by natural selection. Secondly, almost all the research on WAS has been concerned with the question of what we ought to do about the suffering of animals on our own planet. If either, or both, of my arguments in this paper are successful, then we have very strong reasons not to spread life to other worlds by directed panspermia. Even those who oppose intervention in nature can accept that we have reasons not to multiply WAS by spreading it throughout the galaxy. Given the ease with which it could multiply WAS, avoiding the deliberate or accidental spread of microbial life in this way might be an important way of preventing suffering over the long-term future. This is, to my knowledge, one of the first papers to consider how wild animal ethics might be relevant to space ethics, and to consider the impacts of our actions on animals over the very long-term future.

(4) The Case for Animal-Inclusive Longtermism

The fourth paper is an exploration of 'longtermism' and its applicability to animal ethics. Longtermism is the view that the very long-term future is extremely important, and making sure that it goes well is one of the key moral priorities of our time. Longtermists have mostly

ignored nonhuman animals, and animal advocates have, for the most part, ignored the very long-term future. In this paper, I argue that both are mistaken. I describe the basic argument for longtermism and argue that it applies equally well to animals as it does to humans. In particular, I argue that it is reasonable to expect that the animal population will remain larger than the human population over the long-term future, and so it is quite possible that most welfare subjects over the long-term future will be nonhuman animals. Furthermore, there are actions we can take in the near-term future that have significant chances of making the long-term future go better or worse for animals. If the basic argument for longtermism also applies to animals, then longtermists should be much more concerned about animals than they currently are.

Furthermore, people concerned with animal wellbeing might also be able to use the longtermist framework to identify cost-effective interventions that can improve animal wellbeing over the very long term. If animal advocates take the long-term future seriously they might consider taking action now to avoid the spread of WAS (by directed panspermia or terraforming for example), cultivating values that take animal wellbeing seriously, and promoting research into ways of reducing WAS even if they will take decades or centuries to come to fruition. They might also consider the importance of changing our attitudes towards animals now, before humanity begins to colonize space. If we fail to abolish animal exploitation before humanity begins spreading beyond the Earth, we increase the chances that animal exploitation will continue indefinitely, and on a far larger scale than it is practiced today.

The paper contributes to the literature by showing how important the long-term future is for animals. There are many things we can do in the short-term that can make the future for wild and domesticated animals go better or worse. It is quite possible that animals will continue to outnumber humans many times over in the future, and, if it is important to take

steps today to ensure the future goes well for people then it should be similarly important to do so for animals. It highlights ways in which we can begin to make the long-term future go well for animals, and suggests further areas of research. This paper concerns both wild and domesticated animals, and so is slightly wider in scope than the other papers. I do, however, focus more on wild animals than domesticated ones.

(5) How Many Animals Should There Be? Population Ethics, Species, and the Repugnant Conclusion

In the fifth paper I explore how the inclusion of animals affects our thinking about population ethics. This is important for interventionists, as our interventions in nature will inevitably alter the size of animal populations, as well as their species composition and total and average welfare. If we do take on control of the biosphere and its animal population, then we need to have some idea of what kind of population would be desirable. In this paper I restrict myself to considering Totalism, and how successfully it can accommodate nonhuman animals. I pay particular attention to the Repugnant Conclusion. I argue that, while in the human case the Repugnant Conclusion poses a serious axiological problem, in practice it is possible to avoid commitment to unpalatable normative conclusions about what we have reason to do. In the case of animals, however, both the axiological and the normative problems are much more difficult to avoid, and potentially much more troubling.

I argue that if we take Totalism seriously, then we have reason to bring about the population out of the possible options open to us that best maximizes welfare. Since we have finite resources, the population with the highest total welfare must be the one which converts those fixed resources into welfare most efficiently. Once we include animals in the populations assessed, however, we may find that the best population is not one composed of human beings. I argue that a very large population of small animals like mice might have

higher total welfare than a human population with the same resources. This is because mice can attain a level of welfare comparable to that of humans, while being far more resource efficient. Totalism, combined with some basic facts about resource restraints and plausible conjectures about the welfare levels of nonhuman animals, seems to lead us to the conclusion that a population composed in so far as possible of mice would be desirable to bring about. Conversely, if we think that there is something particularly special about human welfare, for example, that we engage in 'the best things in life' or Millian 'higher pleasures', then we might think that the best population is one composed solely of human beings. We would have reason, then, to reduce the nonhuman population as much as possible in order to increase the welfare-maximizing human population.

Since neither course of action – bringing about a world full of mice, or one composed solely of humans – is appealing, it may be that we need to adopt a richer population axiology than Totalism if we are to adequately accommodate nonhuman animals in our population ethical thinking. This is of course important for population ethicists, as their theories must remain incomplete if they cannot include animals in a plausible way. It is also important for animal ethicists, especially, as I have indicated, for interventionists. Population ethics may not be so important for moderate interventionists, who aim only to decrease WAS to a modest degree, for example, by feeding animals in winter or by rescuing trapped or injured animals. For ambitious or systematic interventionists, these questions may be vitally important. If humanity does take complete control of the biosphere, and makes radical changes to its structure and species composition, then we will need an adequate conception of what makes one population better than another. I certainly do not claim to have done this – finding an adequate population axiology is extraordinarily difficult even when we only consider human beings. When we include animals too, it becomes significantly more difficult. I hope, at least,

to have highlighted the importance of, and the difficulty in, finding a population axiology that works for animals as well as humans.

**6 Key Themes and Connections Between Papers**

(1) Responsibility

All the papers in the thesis, with the exception of the first, which, for the sake of argument, assumes a relational account of our duties towards wild animals, assume that human beings have a pervasive, collective duty to help wild animals. This collective responsibility is grounded in beneficence – they are suffering terribly, and we have, or can develop, the capacity to help them, without sacrificing anything too important. We should do so.

Other notions of responsibility for the wellbeing of wild animals appear throughout the papers, however, particularly in the first two papers. One, which recurs in several of the papers, is the idea that human beings can become morally responsible for what are normally considered to be natural harms, such as death by predation or starvation. We do so by enabling animals to inflict those harms on other animals, or, by making animals more vulnerable to those harms through our actions. This responsibility for 'naturogenic' harms is most explicitly set out in the second paper, in which I describe what I call the 'transformation principle'. This is the process by which human beings or other moral agents can become responsible for 'natural' harms by inserting themselves into natural processes, thus turning them from purely natural events to the outcomes of human choices. The most salient examples are rescuing harmful animals such as predators, and releasing them back into the wild, in the knowledge that they will go on to inflict serious harms on other animals. I go on to argue that the same applies to cases in which humans create new animals, or cause them to come into existence, for example, by breeding them or by introducing species to a new ecosystem. More controversially, I go on to argue that we can also transform naturogenic

harms into anthropogenic ones by engaging in large-scale interventions in nature which change which animals come into existence.

Similarly, in the first paper, I argue that we become responsible for the harms inflicted and endured by animals whom we have introduced to new ecosystems. A clear example is the introduction of rabbits, cats, and foxes to the Australian continent. These animals would never have existed had we not introduced their ancestors to this new ecosystem. The animals themselves are not moral agents, and so cannot be responsible for the harmful actions they inflict on others. The predatory animals we have caused to exist on this continent inflict serious harms on other animals, harms which were fully predictable, and so we must take responsibility for those harms. Likewise for the harms introduced animals suffer, such as starvation through overpopulation. Other notions of responsibility to wild animals come up in the first paper. For example, I suggest that, by changing the species composition of the Earth through hunting large herbivores to extinction, we have inadvertently created ecosystems in which there is more suffering, and significantly less positive wellbeing, than there would have been in those ecosystems prior to our interventions. If this is true, then, assuming we can be responsible for the actions of our ancestors as Milburn suggests, we may be responsible for having reduced the average wellbeing in these ecosystems significantly and may have duties to rectify this.

The idea of becoming responsible for 'natural harms' also comes up in the third paper. By seeding a planet with life, we would predictably cause animals to come into existence who would suffer and inflict serious harms on one another. In that paper I also argue that we may have special duties towards our sentient creations. If this is correct, then launching a directed panspermia mission is doubly wrong. It is wrong in the first instance because we knowingly create animals who we can reasonably predict will suffer greatly – in Johannsen's terminology we would be creating both 'harmed' and 'harmful' animals. It is also wrong

because we would be unable to meet the special duties of care that we have towards our creations.

(2) The Moral Importance of the Long-Term Future

Another theme running through several of the papers, especially the third and fourth, is the value of the long-term future, and how important it is that we act today to make it go better for animals, or, at the very least, that we avoid acting in ways which we can expect to make things go worse. This theme is most explicitly addressed in the fourth paper, in which I outline the basic argument for longtermism and argue that it applies at least as well to animals as it does to humans. It is also implicit in the third paper. The fact that the animals we would cause to exist by launching directed panspermia missions won't exist until several hundred million years after the launch of the mission is entirely morally irrelevant. What matters is that these future animals will suffer, and that this is the predictable outcome of our choices today. The second and fifth papers, too, concern the importance of the long-term future, at least tangentially. Both papers assume that humanity will eventually intervene in nature in very significant ways, by altering ecosystems and determining the species composition of the biosphere. Such radical, systematic interventions in the biosphere as genetically altering species require technology and knowledge which are likely at least decades away, and will take a long time to implement successfully.

(3) Population Ethics

Another theme running through several of the papers is the importance of being able to incorporate animals, and especially wild animals, into our population ethical thinking. This is most clearly expressed in the fifth paper, in which I explore the consequences of accepting Totalism while including nonhuman animals in the populations to be assessed. I argue that if

we accept Totalism, and endorse a hedonistic theory of welfare, then, given some plausible assumptions about the degree of welfare nonhuman animals are capable of and the fact that we have limited resources, we ought to prefer a population composed, as far as possible, of whatever species is best for converting limited resources into welfare. I argue that mice would be a good candidate for the most efficient species, fleshing out what such a world would look like, and what it would be lacking.

Population ethical questions also arise in the second paper, in which I discuss how the non-identity problem might affect our reasons to intervene in nature on a large-scale. Essentially, I argue that the No Difference view is false. Even if our interventions in nature make things better impersonally considered, by reducing the amount of suffering in the world, it does make a difference that we achieve this by causing one set of suffering animals to exist rather than another set composed entirely of different individuals. I do not argue that we shouldn't intervene for this reason – it may still be the all-things-considered best thing to do, given the huge amount of suffering averted. However, we may take on special duties towards the harmed animals we would cause to exist in this way, at least to rectify the harms they suffer. The third paper discusses the kinds of moral reasons constraining creators of sentient beings, and endorses the claim that we can harm someone by causing them to exist with a life that will not be worth living.

# On Our Moral Entanglements with Wild Animals

**Abstract:** In *Just Fodder*, Milburn argues for a relational account of our duties to animals. Following Clare Palmer, he argues that, though all animals have negative rights that we have a duty not to violate, we only gain positive obligations towards animals in the contexts of our relationships with them, which can be personal or political. He argues that human beings have collective positive duties towards domesticated animals, in virtue of the kind of relationship between us established by domestication. However, when it comes to wild animals, he argues that we have no such morally relevant relationships, and so we have only negative duties towards them. For the sake of argument, I assume that the relational account is correct. I argue, however, that throughout history, and even prehistory, human beings have morally entangled themselves with wild animals sufficiently that we may in fact have collective positive duties towards many, if not all, wild animals, even on a relational account.

## 1 Introduction

In *Just Fodder*, Milburn investigates the ethics, and politics, of feeding nonhuman animals. Particularly relevant to this thesis is his seventh chapter which concerns our moral duties to wild animals. His central questions regarding them are these: in the case of animals who are starving, do we have a duty to provide them with food? And, in the case of animals preyed upon by predators, are we obligated, or even permitted, to prevent them from becoming food? Following Clare Palmer (2010), he agrees that, though all animals have negative rights that we have a duty not to violate, we only gain positive obligations towards animals in the contexts of our relationships with them. More specifically, we gain positive obligations towards others when we are morally responsible for making them vulnerable to harms, or making them dependent on us in some way. It is clear how such relationships are formed with

domesticated animals. By choosing to bring a companion animal into your home you create a relationship of dependency between her and you. This is because you have closed off any other options that she might have had to take care of herself. She cannot find food for herself, for example, so she depends on you to provide it for her. In Palmer's (p.91-92) terminology these are *external* dependencies. Furthermore, domesticated animals have been bred to have certain *internal* dependencies – after generations of selective breeding they are simply no longer capable of living independently in the wild. Although the individual who adopts a puppy isn't directly responsible for breeding her, by adopting or purchasing the dependent animal he takes on responsibility for her. The important point here is that it is not simply the *interest* the companion animal has in being fed that grounds a duty in the owner to feed her. If an interest in being fed were sufficient to ground a duty to feed, then, Milburn argues, humans would also have duties to feed starving wild animals, whose interest in being fed is no less than that of domesticated animals (p.51). Rather it is the relationship between the dependent animal and the particular human being who has created the relationship of dependency that grounds her right to be fed by that human (p.52-53).

Palmer is primarily concerned with the moral obligations that arise though our *individual* relationships with animals. While she discusses the possibility that humans might have positive obligations towards domesticated animals generally, these obligations are grounded on either the benefits that individuals might incidentally gain from the general practice of animal use, or their contribution to the social attitudes that maintain such institutions (p.106-113). Milburn explicitly argues that our normatively significant relationships can also be collective, political ones, grounded on our membership of morally significant groups such as states, or, perhaps, the human species. He distinguishes between

the vulnerability that particular companion animals[23] have to particular humans, and the deeper sense in which "companions *as a whole* are vulnerable to humans *as a whole*" (p.72, italics Milburn's). This is important, as it allows Milburn to maintain that even domesticated animals without any particular guardian have a right to be fed and protected. This is because domesticated animals, as a group, are dependent on human beings collectively. Human beings brought companion animals into their community by domesticating them, and bred them selectively to have certain features. Companions (and other domesticated animals) generally "have both internal and external vulnerabilities to human communities (or, if preferred, *the* human community)", and "[h]umans, collectively, have the lion's share of responsibility for this" (p.72, italics Milburn's). Milburn draws an analogy between the individual responsibility created by an individual human deciding to bring an animal into his home, and the collective responsibility generated by humans as a whole "bringing a whole population into our society and cutting off alternative possible futures for all of its individual members" (p.72). This collective moral responsibility for the general relationship of vulnerability and dependency between domesticated animals and human beings justifies differential treatment of domesticated animals and free-living ones. A starving domesticated animal is the victim of an injustice if humans fail to help her, because humans collectively are responsible for her dependency and thus for her current plight. No such relationship, Milburn claims, exists between humans and free-living animals. Humans have not created any internal or external dependencies in free-living animals, nor are we usually responsible for their plight. So, while we have strict negative duties not to harm free-living animals, we have no positive duties towards them (p.73).

---

[23] The term Milburn uses in place of 'pets'. What he says in this section though applies equally well to all domesticated animals, not merely those who are used as companions.

In the rest of this paper I will assume that the relational account is correct, and that positive duties only emerge in the context of certain morally charged relationships. I will also accept, as Milburn seems to, that it makes sense to say that humanity as whole can have collective responsibilities toward animals because of the actions of our ancestors. I will argue however that if his account of collective responsibility for historic wrongdoing is correct, then we have a greater degree of moral entanglement with wild animals than is generally assumed, and it is plausible that these entanglements generate collective obligations to assist them. The upshot is that even relational accounts of positive obligations may entail that we have significant duties towards wild animals. Furthermore, those who argue that we have reasons of beneficence to intervene in nature to help animals could bolster their arguments by arguing that we also have collective duties to rectify the wrongs we have historically done to animals. Johannsen (2021) argues that beneficence grounds our duty to help wild animals. He accepts that rectificatory justice may play some smaller role in justifying some interventions, namely those which aim to make right certain wrongs that humans have done to wild animals, for example, by climate change. If my arguments below are correct, then the role played by rectificatory justice may be more significant than this.

## 2 The Anthropocene Challenge

Milburn argues that humans generally aren't entangled with wild animals in the way that we are with domesticated animals. In short, since we haven't created a relationship of dependency between us, or made them vulnerable to harm, and since we aren't responsible for their plight, then we don't have any duties of justice to help them. Though life for wild animals is hard, and there is much we could do to help them, we aren't obligated to do so. He acknowledges however that the claim that we have had little or no impact on the lives of wild animals is open to challenge. It has been proposed that we are currently living in the

'Anthropocene', a distinct geological epoch in which human influence on the natural world is "both pervasive and defining" (p.172). If our impact on the natural world is sufficiently pervasive, then we can no longer assume that we are not entangled with the lives of wild animals in morally relevant ways. If we have created vulnerabilities in wild animals, and if we are responsible for some of the apparently natural harms which they endure, then it may be that we owe them positive duties of assistance, just as we do domesticated animals. He uses climate change as a salient example. It is clear that humans are both morally and causally responsible for climate change. If it is the case that the changes to the planet are having negative effects on wild animals, then we can no longer claim that we have no duties to help them.[24]

Climate change poses severe challenges to wild animals, and our collective responsibility for it certainly entangles us in their lives in morally relevant ways. At the very least, it gives us a duty to help animals to adapt to the changing climate, and perhaps to feed those animals who are no longer capable of feeding themselves because of it. Our duties in this regard though are limited to mitigating or rectifying the negative effects that anthropogenic climate change has on wild animals. So, though we may have duties to assist some wild animals, such as those niche specialists who cannot effectively adapt to climate change on their own, we have no general duty to help wild animals who suffer from harms unrelated to climate change.

Climate change is far from being the only impact that humans have had on the lives of wild animals, however. Below I will suggest some other ways in which pervasive human effects on the planet have made us collectively responsible for the plight of wild animals. These involve significant changes to the landscape and to the species composition of various

---

[24] On our responsibility to help wild animals made vulnerable by climate change see Pepper (2019) and Palmer (2021). It is striking that, though Pepper and Palmer are both generally against intervention in nature, both agree that we have duties to help wild animals impacted by climate change.

ecosystems that human beings have caused from prehistoric times up to the present day. One might object that it is implausible that human beings today bear any responsibility for the actions of their ancestors thousands, or even tens of thousands of years ago. I agree that this is counterintuitive. There seems to be little connection between us and our distant ancestors, and it sounds harsh that we would inherit the guilt for actions they performed. However, if Milburn's arguments about our collective duties to companion animals depend on our ancestors having domesticated them, then it should follow that we also bear responsibility for these other harms caused to wild animals by our ancestors. That is, if modern humans somehow bear collective moral responsibility for the domestication of animals thousands of years ago, then it's not obvious how we could avoid the conclusion that we also bear some responsibility for the various harms that our ancestors inflicted on wild animals.

## 3 Extinction of Large Herbivores

As prehistoric humans spread across the globe, a trail of extinctions of large animals followed them. The 'overkill' hypothesis suggested that as humans arrived in each new part of the world, the extinction of that area's megafauna quickly followed, as humans either hunted them to extinction, or outcompeted them (Martin, 1967). A 2022 study from Tel Aviv suggests that humans have always preferred to hunt the largest animal available to them in any given environment, as it gave the greatest yield of food per unit of effort. When the largest animal was hunted to extinction, humans moved onto the next largest, and so on (Dembitzer et al, 2022). Though it has not been decisively proven, the evidence points towards human responsibility for these extinctions. This wave of extinctions not only caused suffering to the animals our ancestors killed, but also changed the nature of those ecosystems in several ways. First, large herbivores like woolly mammoths, mastodons, and giant ground sloths contributed to maintaining open landscapes, by feeding on sprouting trees and shrubs.

Once they went extinct, this open land became forested (Bakker et al., 2015). Secondly, it is plausible that the extinction of these large herbivores meant that there was more plant material available to be eaten by smaller animals, and so their numbers likely increased.

It is possible that both these changes had an enduring negative impact on the welfare levels of wild animals. First, if one believes that there is net suffering in nature, then, all else being equal, one should prefer ecosystems with lower levels of primary productivity, as this ultimately results in fewer animals, and less suffering. Tomasik (2013) has argued that forests generally have higher primary productivity than open habitats such as grasslands. If this is correct, then the newly forested areas may have had more animal life and hence more suffering than the open habitats maintained by the large herbivores.

Second, philosophers researching welfare biology have speculated about what kinds of ecosystems are more conducive to positive wellbeing for wild animals. Looking at modern day ecosystems, Brian Tomasik (2013), Catia Faria and Oscar Horta (2019), and Luke Hecht (2020) have all argued that, all else being equal, when it comes to wild animal welfare, we should favour ecosystems with large herbivores like elephants over ecosystems without them. The main reason for this is that large herbivores consume huge amounts of plant matter, reducing the amount available to other animals. This prevents the presence of large numbers of small animals who reproduce in vast numbers. Large herbivores, like modern elephants or the extinct woolly mammoth, probably have generally good lives. They have few offspring in whom they invest a great deal of parental care, unlike smaller animals which are generally r-strategists with very high infant mortality rates and low welfare levels. They are also generally relatively long-lived and are safer from predators than smaller animals. As Hecht puts it, "[t]he concentration of biomass into high-welfare individuals is the ideal function of food chains in wild animal welfare" (Hecht, 2020). The extinction of so many large herbivores around the world may have allowed more numerous small *r*-selecting animals to

take their places instead. Since these smaller animals generally have much lower levels of welfare, it is possible that by eliminating large herbivores in many ecosystems across the world, human beings may have inadvertently helped create persistent ecosystems with much more animal suffering than the ones they replaced. If we have replaced higher-welfare ecosystems with lower-welfare ones, then this is a clear way in which we have made ourselves responsible for some of the harms endured by animals. In human terms, it is as if we have eliminated safe neighbourhoods in which inhabitants generally live good lives, and replaced them with more dangerous and unpleasant neighbourhoods in which people's lives are generally less good.

It should be noted that I am *not* claiming that the human-caused extinction of large herbivores was worse for any particular animals, aside from those killed by humans. Such large-scale effects on the biosphere would have altered the behaviour of other animals, including their reproductive behaviour, and so the animals born in these ecosystems were different ones from those who would have been born had humans not altered those ecosystems. Such large-scale changes to the biosphere as the mass extinction of animal species clearly raise the non-identity problem. One might argue that these 'post-extinction animals' have not been harmed by human activity, since, in the absence of that activity, they would never have existed.[25] From this, one might make the further inference that there was nothing wrong about that activity, and so we cannot be responsible for the plight of the animals who exist in those ecosystems today. This would be a mistake. First, the non-identity problem is generally framed as a question about explaining the obligations we have towards beings whom we have caused to exist with lives that are unavoidably flawed in some way, and worse than other lives we could have created, yet still worth living (Roberts, 2022). The difficulty in such cases is in explaining how we could have done something wrong when the

---

[25] 'Post-extinction animals' are analogous to McMahan's (2021) 'Climate-change people'.

alleged victim is, if anything, benefitted by having been caused to exist. In the case of wild r-selecting animals, however, things are different. If the Net Bad Lives claim is true, then it is *not* the case that the animals we have caused to exist have lives worth living. By causing them to exist with lives that are bad for them, we have harmed them, albeit in a noncomparative, or 'existential' sense.[26] Second, even if it were not the case that we harmed the animals that we have caused to exist, we can still be morally responsible for bringing about a worse state of affairs rather than a better one. By causing the less well-off animals to exist rather than the better-off ones, we have made the world impersonally worse, even if no particular individual has been harmed by our act. This may be sufficient to ground a duty for us to make the world better if we can, though it is less clear if it can ground any duties towards the less well-off animals themselves.[27]

It might be objected that this is all highly speculative. The case for human-caused extinction of megafauna is still controversial, and even if it were proven it is also not certain whether ecosystems with large herbivores are generally more conducive to positive animal welfare than ecosystems without them. I agree that this is highly speculative. However, it seems plausible enough that it is worth taking seriously. It may be the case that future research into welfare biology will decisively prove that ecosystems with large herbivores do contribute to higher welfare, and if this is the case then a 'backward-looking' welfare biologist might conclude that the actions of prehistoric humans did indeed have a long-lasting negative effect on the welfare levels of wild animals. In any case, if we are unsure about what effects we have had on animals, and hence are uncertain about our current obligations towards them, it seems fair that we should err on the side of taking too much, rather than too little, responsibility for our actions.

---

[26] For the origin of the terms 'existential harm' and 'existential benefit' see McMahan (2013).
[27] This is especially complicated by the fact that it is a different-number case. For more work exploring the implications of the non-identity problem for animal ethics see O'Brien (2022a), Višak (2018), Purves and Hale (2016), Palmer (2012), McMahan (2008) and paper 2 in this thesis.

**4 Introduction of Species**

As humans spread across the planet, they often brought stocks of domesticated animals and plants with them, and this has changed the species composition of those ecosystems significantly. A particularly extreme case is Australia. Dingoes are considered a native species, but they were introduced to Australia by humans around 8,500 years ago. They are the largest extant terrestrial predator on the continent, and there are between 10 and 50 thousand of them across Australia. Subsequently humans have also introduced red foxes (current population around 7 million), cats (of which there are around 2.8 million feral individuals), European rabbits (of which there are around 200 million), and cane toads (between 200 million and 1.5 billion). Given the vast numbers of these introduced animals, which often outnumber 'native' animals many times over, in what sense can we really say that the Australian wilderness is something natural, rather than the outcome of human choices? It may not be the case that we have created a dependency relationship in these animals, at least in the sense of internal dependency. But it is very clear that these animals wouldn't be there had we not deliberately chosen to bring them to the continent, and so we may well be morally responsible for many of their external dependencies.

First, by introducing rabbits to the continent, we created a new vulnerability in the native animals. European rabbits were introduced to the continent in 1859, when Thomas Austin had 13 wild rabbits sent to him. In only 50 years rabbits had spread across the entire continent, and today there are estimated to be 200 million of them. Rabbits are extremely adaptive and breed in very large numbers. This has allowed them to outcompete native animals, who found themselves unable to find sufficient food, with some species going extinct. It is clear that humans are responsible for the plight of these animals. However, we have also created vulnerability in the rabbits themselves, as they are preyed upon by the

dingoes and red foxes, which we also introduced. Our introduction of cats to the continent has also been a disaster for the native animals – it has been estimated that feral cats kill 1.5 billion native animals in Australia each year. Milburn accepts that humans can become responsible for the predatory actions of animals. For example, by rehabilitating and releasing injured predators we thereby become causally and morally responsible for the killings they go on to perform (p.155).[28] If humans can be individually responsible for predatory killings, then it seems that by collectively introducing predators to new habitats we must also become collectively responsible for those killings, and so we may have duties, grounded in the negative rights of prey animals, to protect them from the actions of the predators we have introduced.

The case of introducing the dingo to Australia also illustrates something important about the moral relationship we have towards wild animals whose lives we have impacted. That is, it needn't be the case that our actions make things worse for wild animals in general for us to bear moral responsibility for the harms that they endure. Before the dingo outcompeted it (and humans helped kill it off), the thylacine was the apex terrestrial predator in Australia. It undoubtedly inflicted a great deal of death and suffering on the native animals of the continent before humans ever set foot there. If human beings had never set foot in Australia, then it seems likely that the thylacine would still be terrorizing its prey today. Humanity however would bear no obligations of justice to protect prey animals from the thylacine, as we would not have been morally or causally responsible for the vulnerability of these animals. By replacing the thylacine with predators we introduced, and indeed by introducing many of the prey animals that these predators prey on, we have entangled our

---

[28] For more on humans becoming morally responsible for what are normally considered natural harms see O'Brien (2022a), Johannsen (2021, p.46), Milburn (2021), and paper 2 in this thesis.

lives with the lives of these 'wild' animals in such a way that we take on duties of justice to protect them.

Finally, though the Australian case is a particularly egregious example of humanity introducing new species, it is far from being an anomaly. Humans have introduced new species all over the world.[29] Harms caused by introduced species include predation, the spread of disease, and the outcompeting of native animals.[30] If we thereby take on positive obligations towards and regarding these introduced species, then it seems that we are much more entangled with the natural world, and with the lives of animals, than we ordinarily think.

**5 Land Use**

Wild animals obviously depend on having viable habitats to survive. Since the agricultural revolution, humanity has changed the landscape of the Earth almost beyond recognition. This is primarily due to deforestation for farmland. Of the Earth's habitable land (land that isn't covered by glaciers, deserts, exposed rock, or beaches) approximately 50% is used for agriculture, and about 37% is forested. This means that many animals have been robbed of the kind of habitat that they need to survive. If we think that animals have a right to their habitats, then this constitutes a rights violation on a massive scale.[31] Having essentially stolen land from the ancestors of modern-day animals, it may be reasonable to think that we have some kind of duty to pay them reparations. It would be difficult to repay wild animals by simply rewilding more land for them to live on. First, that might violate the rights of human beings who now depend on that land. Second, by rewilding we might become responsible for

---

[29] A report by the European Environment Agency (2020) states that there are already over 10,000 introduced species in Europe, and the rate of new introductions is still increasing.
[30] Cats have been introduced by humans to over 180,000 islands globally. In the UK alone they kill between 25 and 29 million birds each year. The chytrid fungus, spread by humans, has dramatically reduced numbers of amphibians worldwide. Ibid.
[31] See Milburn (2020, 2017).

the harms endured by animals on that land, as we would be deliberately introducing animals to lands on which they would predictably come to harm, either from starvation or from predation. But if we cannot rectify these rights violations by rewilding land, what can we do? I propose that the best currency with which to rectify the harms done to animals by driving them from their land is to try to increase the welfare of wild animals suffering from various naturogenic harms.[32]

Furthermore, it is not just the fact that humans have collectively laid claim to most of the land surface of the planet that entangles us with wild animals. Human infrastructure, such as roads, railway lines, shipping lanes, buildings, and so on significantly impact the lives of wild animals, often in fatal ways. It has been estimated that in the UK over 300,000 hedgehogs die on the roads each year, and a 2023 study demonstrates that road mortality is a significant factor in reducing mammalian populations globally.[33] Our infrastructure also contributes to habitat fragmentation, by dividing up large natural areas into smaller, fragmentary habitats, separated by roads and other kinds of infrastructure. This has significant effects on animal behaviour, including foraging and mating behaviours, sometimes also resulting in greater predator-prey overlap when prey animals are less able to avoid their predators.[34]

## 6 Conclusion

Milburn argues that the domestication of animals by our distant ancestors is sufficient to ground collective positive duties towards domesticated animals for modern humans. However, we have no relationship with wild animals, and so we generally have only negative

---

[32] See Jalagania (2021) for a similar argument regarding what we owe animals who have benefitted us in some way.
[33] BBC (2020) 'Hedgehog road deaths in UK 'as high as 335,000' https://www.bbc.co.uk/news/uk-england-nottinghamshire-54524338. Moore et al. (2023).
[34] See Banks et al. (2007) and Schneider (2001).

duties towards them. I have argued that human beings have been entangled with the lives of wild animals for thousands of years. By driving some species to extinction and introducing other species into new lands, and by claiming and transforming most of the planet's habitable land, we have had long-lasting and pervasive effects on the lives of wild animals. These effects are sufficient to ground positive obligations towards them. While this may not be sufficient to show that we have general duties to feed and protect *all* wild animals from harm, it does suggest that we may have many more positive obligations to wild animals than we think. Even if Milburn is correct that in the absence of relationships with them our only obligation to wild animals is to leave them alone, the fact is that we have not, do not, and arguably cannot leave them alone. We have created vulnerability in wild animals and are in many cases either directly or indirectly responsible for the harms that they endure, and as such we may well have positive duties to rectify those harms, and perhaps to feed and protect those animals whom we have made more vulnerable or dependent than they would have been in the absence of our actions. The wild, like our companion animals, may have become 'inextricably tainted with human agency' (Milburn, 2022, p.73). Even on a purely relational account of animal rights then, it seems possible to support a programme of large-scale intervention in the natural world.

Even if one rejects relational accounts of our moral duties and endorses intervention in nature on the grounds of beneficence instead, the fact that we are so deeply entangled in the lives of wild animals may be morally relevant. We might think, as Johannsen (2021) does, that while beneficence is the best justification for intervention in nature, considerations of rectificatory justice may play a supporting role. If the arguments I have made in this paper are successful, then this supporting role may be significantly larger than generally supposed. These arguments, however, depend on a contentious, and undefended, claim about collective moral responsibility based simply on being part of the human community. I do not have the

space to investigate this claim here, nor is it relevant to any of the other papers in this thesis. I will simply report that I do not find it too implausible that we may have at least weak moral reasons to rectify harms done by our ancestors, simply because they are our ancestors. This kind of collective responsibility at the level of the species is somewhat similar to the kind of 'national guilt' that members of some contemporary societies feel regarding the actions of their ancestors. If one is completely unmoved by this idea, consider two planets filled with sentient animals. The number of individuals on each planet is the same, as is the total amount and distribution of suffering. Planet A has always been causally isolated from humanity – the suffering there has not been influenced to any degree by human action, and is purely naturogenic. On Planet B, however, the suffering has been causally affected by our ancestors to the same extent that WAS on our own planet has been. If we can do something to mitigate the suffering on only one of these worlds, which should we choose? If we are completely unmoved by the idea of collective responsibility for the actions of our ancestors, we might think we ought to simply flip a coin. If, on the other hand, we think that there is a moral reason to 'make amends' for the actions of our ancestors, we will choose to help the suffering creatures on Planet B.[35]

---

[35] A version of this paper appeared in *Food Ethics.* See O'Brien (2023).

# Beneficence, Non-Identity, and Responsibility: How Identity-Affecting Interventions in Nature can Generate Secondary Moral Duties

**Abstract:** In chapter 3 of *Wild Animal Ethics* Johannsen argues for a collective obligation based on beneficence to intervene in nature to reduce the suffering of wild animals. In the same chapter he claims that the non-identity problem is merely a "theoretical puzzle" (p.32) which doesn't affect our reasons for intervention. In this paper I argue that the non-identity problem affects both the strength and the nature of our reasons to intervene. By intervening in nature on a large scale we change which animals come into existence. In doing so, we enable harmful animals to inflict harms on other animals, and we put other animals in harm's way. The harms that these animals will inflict and endure are foreseeable. Furthermore, since non-human animals aren't moral agents, harmful animals cannot be morally responsible for their harmful actions. I argue therefore that by causing animals to exist, knowing that they will inflict and suffer harms, we become morally responsible for those harms. By engaging in identity-affecting actions, then, we take on secondary moral duties towards the animals we have thereby caused to exist, and these secondary moral duties may be extremely demanding, even more so than the initial costs of intervention. Finally, these duties are duties of justice rather than duties of beneficence, and as such are more stringent than purely beneficence-based moral reasons. Furthermore, this conclusion flows naturally from several plausible principles which Johannsen explicitly endorses.

## 1 Introduction

Johannsen's "*Wild Animal Ethics: The Moral and Political Problem of Wild Animal Suffering*" (2021) has the well-deserved distinction of being the first book on an important, but thus far neglected moral issue, the problem of wild animal suffering. Johannsen offers a

sophisticated account of the naturogenic causes of WAS and an analysis and critique of the idea that nature has intrinsic value (chapter 2); he engages with the literature on the ecological risks of large-scale intervention in nature and argues that the risks are manageable and acceptable, especially given the scale and severity of the problem (chapter 4), and he makes concrete proposals for radical technological interventions that promise to liberate animals from needless suffering (chapter 5). Another distinctive element of Johannsen's book is his focus on the political elements of the problem of WAS. He argues for a collective obligation to intervene in nature grounded in beneficence (chapter 3). Though he accepts that considerations of distributive and rectificatory justice may be important, he argues that interventionists ought to focus on beneficence-based reasons and downplay the importance of justice-based reasons for intervention on both strategic (p.44) and principled grounds (p.30, 42-43, 46).

In the same chapter in which he argues for the primacy of beneficence-based reasons for intervention, Johannsen claims that the non-identity problem is merely a "theoretical puzzle" (p.32) which doesn't affect our reasons for intervention. In this paper I will argue that the non-identity problem does affect both the nature and the strength of our reasons to intervene. By intervening in nature on a large scale we change which animals come into existence. Some of these animals will be what Johannsen has called 'harmful animals', that is, animals who inflict net harm on other animals, such as predators, or r-strategists who will go on to have large numbers of uncared for offspring. Others will be 'harmed animals', animals who suffer because of the actions of harmful animals (for example, prey animals who will be killed by predators, or r-strategist animals who die shortly after birth). The harms that these animals will inflict and endure are foreseeable – we know in advance that if we cause a wolf to exist she will kill other animals, and we know that if we cause r-strategists to exist the vast majority of them will die young. Furthermore, since non-human animals aren't moral

agents, harmful animals cannot be morally responsible for their harmful actions. I argue therefore that by causing animals to exist, knowing that they will inflict and suffer harms, we become morally responsible for those harms. By engaging in identity-affecting actions, then, we take on secondary moral duties towards the animals we have thereby caused to exist, and these secondary moral duties may be extremely demanding, even more so than the initial costs of intervention. Finally, these duties are duties of justice rather than duties of beneficence, and as such are more stringent than purely beneficence-based moral reasons. Furthermore, this conclusion flows naturally from several plausible principles which Johannsen explicitly endorses.

In section 2 I will flesh out the relevant moral principles I take Johannsen to be committed to. The most important of these, which I call *the transformation principle*, is the principle that we can become responsible for the harmful actions of nonhuman animals if our actions have enabled them to inflict those harms, for example, by rescuing them. In section 3 I will show how the responsibility-generating conditions which make us responsible for the harmful actions of nonhuman animals we have rescued also apply to cases in which we create harmful animals. Furthermore, by creating animals vulnerable to harm we take on some responsibility for their wellbeing. In section 4 I will suggest that we can similarly become responsible for natural harms when our identity-affecting interventions determine which animals come into existence. Finally, I will consider how all of this affects our moral reasons to intervene in nature.


## 2 Transformation by Rescue

When discussing the basis of our moral reasons to intervene in nature, Johannsen considers whether rectificatory justice might provide moral reasons for beneficent intervention (p.30). The idea is that, since human beings have caused, and continue to cause harm to wild animals

(by climate change for example) this might generate collective moral duties to compensate animals for those harms. The rectificatory claim is that, since we are causally and morally responsible for the harms that climate change has inflicted on non-human animals, we have strong moral reasons to compensate animals for these harms. These reasons aren't based on general duties of beneficence; rather they are grounded in considerations of rectificatory justice. Furthermore, it's plausible that duties of rectification are generally stronger than duties of beneficence. That is, our duties to rectify harms for which we are responsible are stronger than our duties of beneficence to prevent equivalent harms for which we aren't responsible.

Though he accepts that considerations of rectificatory justice might play some "highly circumscribed" (p.46) role in determining which interventions ought to be funded, he rejects rectificatory justice as a general basis for intervention for two reasons. First, rectificatory justice only applies to harms caused by human beings. Since human beings aren't responsible for naturogenic harms such as those caused by r-selection and predation (which cause the vast majority of WAS), rectificatory justice cannot ground any moral reason to intervene to prevent these harms from occurring.[36] At best, rectificatory justice can ground duties to engage in limited interventions in nature to correct the harms we have caused, while ignoring those naturogenic harms which cause most WAS. By contrast, appealing to beneficence allows us to justify a general programme of intervention in nature to improve wild animal welfare, regardless of whether the causes of suffering are naturogenic or anthropogenic.

Second, rectificatory justice cannot ground duties to help harmful animals, that is, those who are likely to go on to cause significant further harms to other animals. This is because when we save the lives of harmful animals we become causally responsible for the

---

[36] Though if my arguments in paper 1 are successful, we may be more responsible for apparently naturogenic harms than we typically believe.

harms they inflict. If a human being saves the life of a predator, and that predator goes on to inflict harms on other animals, the human being who knowingly allowed the predator to continue causing such harms must be responsible for them. In Johannsen's words "someone who rehabilitates a predator violates the rights of her prey" (p.46). Let's call this the transformation principle. The transformation principle states that human beings can transform naturogenic harms into anthropogenic ones by enabling harmful animals to harm others, for example, by saving their lives. Though humans aren't normally responsible for the harms that animals inflict on each other, by transforming these harms into anthropogenic ones we become responsible for them.

But how exactly does this transformation work? How is it that merely by saving the life of a predator a human becomes responsible for the future harms she inflicts? After all, we don't generally think that by saving the life of another human we become morally responsible for any future harms that he goes on to inflict. For example, if someone saves the life of a child drowning in a shallow pond, and then that child goes on to commit a string of murders twenty years later we don't attribute any moral responsibility to the rescuer even though the rescue was undeniably a causally necessary precondition for the subsequent murders. How are things different in the case of rescuing harmful animals?

Johannsen seems to have three responsibility-generating conditions in mind which must be met in order to transform the naturogenic harms inflicted by harmful animals into anthropogenic ones by rescuing them:

(i)    *Causal responsibility* – by saving the life of a harmful animal we enable him to perform future harmful actions which he couldn't have performed if we hadn't rescued him. In Johannsen's words "Though human beings aren't normally the cause of harms associated with predation or the r-strategy, we become causally responsible when we save predators' or r-strategists' lives" (P.46)

56

(ii)    *Foreseeability* – we can reasonably predict that the rescued animal will go on to cause harm to others. The foreseeability condition only arises in the case of rescuing harmful animals, that is, those who "will *predictably* go on to cause significant harms, namely r-strategist animals or predators" (p.46, italics mine)

(iii)   *Non moral agency* – the harmful animal isn't a moral agent, and hence cannot be morally responsible for his harmful acts (p.46)

These three principles aren't intended to constitute a full account of when one can become responsible for the actions of another agent. Rather they are intended to serve as an intuitively plausible starting point for determining responsibility in such cases. Furthermore, they do seem to give the intuitively correct answer in a variety of cases. They rightly give the result that the rescuer who saves the drowning child who later becomes a serial killer isn't responsible for the killings. In that case, though the rescuer shares causal responsibility for the murders, it isn't the case that the murders were foreseeable – the rescuer had no way of knowing that the child would go on to become a killer later in life. But it should be clear that all three conditions are met when one saves the life of a harmful animal. If I save the life of a wolf, it is obviously true that I am causally responsible for the harmful actions the wolf goes on to perform, since she couldn't have performed those actions if I hadn't rescued her. It's also perfectly foreseeable that the wolf will go on to kill other animals – this is what wolves are biologically programmed to do.

Finally, most people agree that non-human animals aren't morally responsible for their actions.[37] Rescuing the wolf isn't like rescuing a human being who can be morally responsible for her actions; rather it is like setting in motion some purely mechanistic process

---

[37] There has been some exploration of the idea that nonhuman animals can be morally responsible for some of their actions (see Borchert and Dewey, 2023; Ferrin, 2019; Bekoff and Pierce, 2009). I am not aware of any claims that predators are morally responsible for killing their prey, or that r-strategists are morally responsible for the harms that their offspring will endure.

that will foreseeably result in significant harms to other animals. If anyone is morally responsible for those harms, it must be the rescuer. This may seem counterintuitive. We are not used to thinking that we may be responsible for the actions of wild animals whom we have enabled to cause harms.[38] In the case of domesticated animals, however, the idea of taking responsibility for their actions is much more familiar. If I allow my violent dog off the lead and he harms someone, I have enabled that harm, and I am morally responsible for it. The same is true when we enable wild animals to inflict harms.

The transformation principle doesn't entail a general moral prohibition on rescuing harmful animals, but it does mean that we take on additional moral duties should we decide to do so. If we rescue a harmful animal then we take on secondary moral duties to either prevent, or compensate the victim for, the harms that the harmful animal will go on to inflict. If we rescue a wolf for example then we would take on the secondary moral duties to prevent her from harming other animals, perhaps by keeping her confined and feeding her synthetic proteins. These secondary duties can be understood as additional costs of rescue. In some cases, these additional costs will be high enough that we might no longer have a duty to rescue the harmful animal at all. To see this, imagine two rescue scenarios:

Rescue 1: We see an adult monkey stuck in a lake. If we do nothing she will drown. The cost to us of rescuing her is minimal – it would simply take a little bit of time and effort to drag her from the lake. Afterwards she will go on to live a life worth living without significantly harming any other animals.

Rescue 2: We see an adult wolf stuck in a lake. If we do nothing she will drown. The cost to us of rescuing her is the same as the cost of rescuing the monkey. Afterwards she will go on to live a life as worth living as the life of the monkey. However, as a

---

[38] Though see Milburn's (2022, chapter 6) discussion of wildlife rehabilitation centres, and his claim that humans who rehabilitate a predatory animal and release it back into the wild are responsible for the death and suffering caused by the animal (p.137).

predator she will inflict very serious harms on many other animals during the rest of her life. To prevent these harms, we would have to keep her in captivity and provide synthetic food for her for the rest of her life.

In rescue 1 the cost to us is very clearly outweighed by the benefits to the monkey. We have a moral duty to rescue her. Rescue 2 is very different. Though the initial cost of rescue and the benefits to the wolf are the same as those in rescue 1, the fact that the wolf is a harmful animal generates secondary moral duties for us after we rescue her. Since the costs of fulfilling these duties are very high, we may be justified in letting her drown. It may be a very good thing for you to rescue her and bear these secondary costs; nevertheless, unlike rescue 1, rescue 2 may be supererogatory.

Finally, these secondary moral duties to prevent or rectify the harms caused by the harmful animal one has rescued are grounded in considerations of rectificatory justice, not beneficence, and this makes a difference to their strength. By rescuing a harmful animal we become responsible for the harms that he will go on to inflict on other animals, and so we violate the negative rights of prey animals not to be harmed. The negative right not to be harmed is generally stronger than a claim to beneficence, even if the degree of harm is the same. Our duty to prevent or to rectify these harms then is stronger than our duty to prevent or mitigate harms for which we are not morally responsible.

## 3 Transformation by Creation

Rescue isn't the only way in which we can enable a harmful animal to inflict harms, nor is it the only way in which we can become morally responsible for their harmful actions. We can also do so by creating harmful animals. We can imagine for example that, in order to conserve the species, conservationists decide to breed Siberian Tigers in captivity then release

them into the wild. In doing so, they meet all of the responsibility-generating conditions that apply in rescue cases:

(i)     *Causal responsibility* - by creating and releasing the tigers they enable them to inflict harms on other animals;

(ii)    *Foreseeability* - they can reasonably predict that the tigers, once released, will inflict very serious harms on other animals;

(iii)   *Non-moral agency* - the tigers aren't moral agents.

It seems clear that the conservationists are responsible for the harms that the tigers will inflict on other animals. Worse still, there seems to be no reason to limit the responsibility of the conservationists to only those harms inflicted by the tigers that they are directly responsible for creating and releasing. Rather, their responsibility extends into the future, to the harms that the descendants of the released tigers will inflict on future animals. The tigers bred and released by the conservationists can reasonably be expected to breed in the wild and create new generations of tigers. And, of course, it's equally predictable that those new tigers will also go on to harm other animals, to procreate, and so on. The new generations of tigers aren't moral agents, and so they cannot be morally responsible for the harms they inflict, nor can the original generation of released tigers be held morally responsible for procreating. If the conservationists are responsible for the harmful actions of the original cohort of tigers they bred and released, it seems that they must also be responsible for the harms inflicted by future generations of tigers who would never have existed without the conservationists' intervention. If this is correct, then introducing harmful animals to environments where they can spread and multiply is an extremely weighty moral decision, one which may make one morally responsible for innumerable harms even long after one's death.[39]

---

[39] This raises serious questions about the permissibility of rewilding projects, especially those that aim to reintroduce locally extinct predators back into areas in which they historically lived. More speculatively, as I will discuss in paper 3, it suggests that spreading wild animal life to other planets may not be permissible.

We can also become responsible for the wellbeing of harmed animals in virtue of creating them. If I deliberately choose to create an animal, knowing in advance that he will be vulnerable to harm, I have strong moral reasons to ensure that he will have a reasonable chance of having a life worth living, and this is something I owe to him as a matter of justice in virtue of having created him.[40] This also applies to collective acts of creation. Think for example of the introduction of rabbits to Australia in the 18th century. Humans deliberately introduced breeding populations of rabbits to the country for the purpose of hunting them. It is commonly accepted that humans are morally responsible for the ecological damage that the rabbits went on to cause. It is equally true that humans must be responsible for the harms endured by the rabbits themselves – death by starvation, disease, predation, and extreme weather conditions. Deliberate human action resulted in millions of rabbits coming to exist, rabbits which never would have existed without human intervention. This action was causally necessary for the suffering of the rabbits, it was foreseeable that the rabbits would suffer and neither the rabbits themselves nor their predators are moral agents. If humans are responsible for the wellbeing of these rabbits, then the many harmful actions we take against them for the purpose of ecological conservation are doubly wrong – not only do we inflict harms on innocent animals, but we fail to give them what we owe them as their creators.

## 4 Transformation by Identity-Affecting Actions

In this section I will suggest that some of our identity-affecting interventions are morally similar to the kinds of rescue and creation cases described above. By intervening in nature, we change which animals come into existence. In the case of harmful animals, we

---

[40] The idea that creation generates special duties for the creator to his creation seems to be widely accepted in procreative ethics but has been little discussed when it comes to the creation of animals. See paper 3 in this thesis for an argument in favour of extending this principle to the creation of animals, and Unruh (2021) who argues that causing future generations to exist grounds special duties to future generations for the present generation.

foreseeably enable them to harm others, and hence become responsible for their harmful actions. By causing them to exist we enable them to inflict harms on other animals just as much as we did in the rescue and creation cases. In the case of 'harmed-animals' it was our identity-affecting interventions which brought them into existence and put them into harm's way. I will now explain how changing the victim of a natural harm can transform it into an anthropogenic one.

*Smith and Jones 1*

Smith is about to be fatally struck by lightning, a totally natural harm for which no one is responsible. You can't do anything to prevent the harm. You do, however, have the ability to change who the victim is. By clicking your fingers, you can make Smith and Jones instantly switch places. If you do so, Smith will be whisked away to safety and Jones will be put in harm's way. If you do nothing you aren't responsible for the harm that befalls Smith. If you click your fingers, then you are both causally and morally responsible for the harm that befalls Jones. By deliberately substituting one victim for another you have transformed a natural harm into an anthropogenic one. You have killed Jones, and are legally and morally liable for punishment.

*Smith and Jones 2*

This case is the same as the first one, except that it is Smith's heart condition which makes the lightning strike fatal. Jones has no such condition – if the lightning strikes her it will be very painful, and she will spend several weeks in hospital, but she will survive. You know all of this, and are deciding what to do. If you do nothing, Smith will die. The harm is a fully naturogenic one. You aren't morally responsible for it, and you have no secondary moral duties to compensate Smith's family. If you click your fingers this will bring about a better

outcome, impersonally considered (though of course it is significantly worse for Jones). You will again however have transformed a naturogenic harm into an anthropogenic one, and so you will take on secondary moral duties towards Jones. You have deliberately harmed her, and you have duties based on rectificatory justice to make things right, by paying her medical expenses for example. Clicking your fingers to bring about the lesser of two harms may be the right thing to do, but this fact doesn't excuse you from your secondary duties towards Jones.

One might argue that in this case it is actually Smith who takes on duties to compensate Jones. This is because Smith is the one who has benefitted most from the rescue. I think the objection fails in the case of nonhuman animals, though. Since animals aren't moral agents, they cannot have moral responsibilities to compensate others. If we inflict some harm on an animal in order to prevent greater harms to another, it cannot therefore be the case that the benefitting animal must compensate the harmed one. The moral responsibility to compensate the harmed animal must remain with the human agent. Furthermore, in the identity-affecting cases we are going to consider, the 'benefitted' animals don't exist, as we have caused different animals to come into existence in their place.[41]

In Smith and Jones 2 the victim (Jones) already exists, and the agent takes on secondary moral duties towards the victim by diverting a natural harm from an original victim to another one of his own choosing. Can one also become morally responsible for a 'natural' harm by performing an identity-affecting action? That is, can a natural harm be transformed into an anthropogenic one not by diverting the harm from one existing individual to another, but by changing which victim comes into existence? Consider:

> *Mutagenic Gas* – A couple is about to conceive a child. Unbeknownst to them, they
>
> are surrounded by a naturally occurring gas that will cause serious birth defects in any

---

[41] Thanks to Jeff McMahan for this objection.

child who is conceived while the parents are exposed to it. You know this, but you cannot warn the couple. All you can do is launch a fireworks display, which will distract them for an hour, after which some of the gas will have dispersed, and the child they conceive will have less serious birth defects. This will however be a different child from the one they would have had if you hadn't launched the fireworks. Assume that neither child will have a life worth living, and that the parents will be unable to take care of either child.

In this case, the best outcome impersonally considered is the one in which you launch the fireworks, since in this outcome the child with less serious birth defects will be born, rather than the one with more serious defects. The outcome is bad for the child who is born however, since he doesn't have a life worth living. That child, if he were able to, would complain that your actions have harmed him. By your deliberate action, you caused this child to come into existence, knowing that he would have a bad life. Your action was causally necessary for the child's existence, and you were able to foresee that the child would have a bad life. Is it plausible that you acquire any moral duties towards the child? If you had done nothing, then the resulting child would have been nothing to do with you – you aren't causally implicated in his existence, and he would have come into existence whether you were around or not. The second child could reasonably complain to you that he has a bad life as a result of your actions. I think it is at least plausible that you have some duties towards this second child that you wouldn't have had towards the first one.[42]

---

[42] This may be counterintuitive, but there is some legal precedent. In 2021 Evie Toombes successfully brought a claim on her own behalf for 'wrongful conception and birth'. Her claim was that her mother's doctor negligently failed to advise her to supplement with folic acid before trying to become pregnant. As a result, Evie Toombes was born with significant disabilities. Had the doctor not been negligent, Toombes' mother would have conceived at a different time, and a different, healthy child would have been born instead. The doctor's action was an identity-affecting one, which caused Toombes to exist instead of another, presumably better off child. The case is, in one way, even more remarkable than the thought experiment I outlined – Toombes received compensation despite the fact that she has a life well worth living, and so is in one sense the lucky beneficiary of the doctor's negligent action, rather than its victim. See Barraclough Jones and Wallace (2021) for the facts of the case, and McMahan (unpublished manuscript) for detailed philosophical analysis.

Let's now turn to large-scale identity-affecting interventions in nature. Consider:

*Beneficent Intervention* - There is an island populated by a species of herbivorous r-strategists, the Rs. Rs reproduce in very large numbers, and consequently the vast majority of them die of starvation shortly after birth. Their lives are, consequently, bad for them overall.[43] Each year X number of Rs starve to death. Using gene drives we alter the DNA of the Rs so that they produce half as many offspring as they naturally would, and this means that approximately half as many R infants starve to death each year.

Our intervention has had two important consequences. Firstly, we have changed the identities of the Rs that come into existence. After a few generations, no particular individual in the post-intervention world would have existed had we not intervened (this includes both the harmed-animals, that is, the ones who starve to death, and the harmful-animals, those few Rs who survive long enough to reproduce). Secondly, we have made it the case that many fewer animals suffer and die each year than would have if we hadn't intervened. The result isn't better for any particular animal, but impersonally considered the situation post-intervention is far better than the situation which would have resulted if we had done nothing. Though our action has resulted in there being half as many victims of starvation each year than there would have been, it has also substituted one set of victims for a completely different set. We have made it the case that millions of Rs, who would otherwise never have existed, come into existence only to suffer and die, after living bad lives. It's true that we have reduced the harms that occur on the island. It's also true that we have done so at the cost of diverting purely naturogenic harms so that different victims are harmed instead. Essentially, we have sacrificed the wellbeing of some for the sake of the wellbeing of a greater number of others.

---

[43] This thought experiment doesn't depend on the Net Bad Lives claim being true. Rather, it requires only that the particular animals targeted by the intervention have bad lives. To be generally applicable, though, the argument does require that many animals do have lives that are overall bad for them. If the Net Bad Lives claim is true, the argument will have very wide application.

This may well be the all-things-considered right thing to do, but this doesn't mean that we don't take on secondary moral duties towards those unfortunate creatures we have sacrificed. And those secondary duties may be very onerous indeed since they are based on the more demanding ground of justice rather than that of beneficence. Furthermore, it seems that our action meets all three of the responsibility-generating conditions outlined above:

(i)     *Causal responsibility* - by bringing these particular animals into existence we enable the harmful ones to inflict harms, and we put the harmed-animals into harm's way;

(ii)    *Foreseeability* - we could reasonably predict that our intervention would affect the identities of the animals who come into existence, and that the majority of those animals would suffer and die prematurely;

(iii)   *Non moral agency* - none of the animals are moral agents.

What kind of secondary moral duties might we acquire towards those animals we have caused to come into existence? At the very least we owe it to them to try to prevent them from starving to death, and to make their lives worth living overall. This will require further interventions on the island, and these further interventions will also affect the identities of the individuals who will come to exist, making us responsible for the harms which they inflict and endure too. All of this may be far more difficult and costly than the initial relatively simple intervention to reduce their numbers was.

There are two objections to consider here. Firstly, we can ask whether we actually harmed any of the animals on the island by causing them to exist. If they generally live good lives overall, then it may be the case that we have actually benefitted them by causing them to exist, despite the numerous harms that they endure during the course of their lives. It may be the case that many animals do have lives worth living, assuming they survive infancy. For the r-strategists who die painfully shortly after birth (the vast majority of all animals) I think it's

unlikely that their lives are worth living for them. If their lives aren't worth living, then we have harmed them by causing them to exist. Even if we assume that they do have lives worth living, however, it isn't clear that we don't have any secondary moral duties towards them. Even if we have conferred on them a net benefit (a life worth living) we may still have duties to prevent them from suffering harms, or to compensate them for harms they have suffered as a result of our causing them to exist, especially if the harms they endure are caused by harmful animals we have also caused to exist.

Secondly, we can ask whether it is in general plausible that we can acquire additional moral duties solely in virtue of having performed large-scale identity-affecting actions. After all, social policies implemented by governments also affect the identities of those who will come to exist, but we don't normally think that the resulting people are owed compensation for the harms that they endure as a result of being caused to exist. For example, if the government adopts a wide-ranging clean energy policy to prevent climate change, this will change people's behaviour, ultimately resulting in a population composed entirely of individuals who wouldn't have existed had the government adopted a different policy.[44] We don't think that the government owes compensation, or any special moral duties towards the people it has caused to exist by adopting this policy. Why should the animal case be any different?

I think there are three key differences. In the case of the government's large-scale identity-affecting actions the causal responsibility for the existence of the post-intervention individuals is widely dispersed. While the government's adoption of the clean energy policy is a necessary condition for their existence, there are innumerable other actions by human agents that are also necessary, most obviously the procreative actions of the parents. This might not be the case in interventions in nature – in that case there may be no downstream

---

[44] This example is taken from McMahan (2021).

actions by human agents after intervention that are causally necessary for the existence of the post-intervention animals. Secondly, in the case of intervention in nature it is foreseeable that many of the animals who will exist after the intervention will have bad lives, while most human lives are worth living. Finally, many of the harms that the post-intervention people will suffer will be caused by the deliberate actions of human beings who are moral agents and thus must accept moral responsibility for those harms. In the animal case, the harms they endure aren't caused by moral agents. The moral responsibility for those harms must therefore belong to the human agents whose actions enabled them to occur.

**5 How This Affects our Moral Reasons to Intervene**

I have argued that identity-affecting interventions in nature generate secondary moral duties towards the animals we cause to exist. More specifically, (i) we bear some responsibility for the wellbeing of animals we create, (ii) we become responsible for the harms we foreseeably enable harmful animals to inflict, and (iii) we become responsible for the harms endured by harmed animals when our identity-affecting actions cause them, rather than some other animals, to suffer harms. Furthermore, the duties we acquire to these animals are duties of rectificatory justice, and hence are more stringent than comparable duties of beneficence. Far from being a mere theoretical puzzle, or undermining our reasons to intervene, the non-identity problem actually strengthens our moral duties towards wild animals, and Johannsen's emphasis on beneficence becomes more obviously a strategic move rather than a principled one.

The most obvious way this affects our moral reasons to intervene in nature is that it raises the costs of intervention. As in the rescue case above in which rescuing the drowning wolf generated secondary moral duties which were more costly than the initial rescue, some of our large-scale interventions in nature seem to generate secondary duties which are

extremely costly. This increase in the cost of intervention may make it permissible for us to refrain from intervening in nature at all. On the other hand, we may think that the scale and severity of the suffering endured in the natural world makes intervention mandatory, regardless of the secondary costs that may result.[45] Once we begin systematically intervening in nature, however, it may no longer be permissible for us to stop. The more we intervene in nature, the more we transform naturogenic harms into anthropogenic ones, and the more responsible we become for the overall state of the biosphere, the more duties we take on. Natural harms call for intervention, and intervention calls for more intervention, and a greater assumption of responsibility for the state of the biosphere and the wellbeing of all those individuals who compose it, individuals who, to an ever-growing extent, wouldn't have existed but for our actions. Eventually, there are few truly natural harms, as the biosphere becomes more and more an artefact of human choices rather than something independent of our will. Beneficence makes us intervene in the natural world, and our interventions bring the natural world under human control, and hence into the sphere of justice.[46]

---

[45] Furthermore, if my arguments in paper 1 are correct, then we may already have reasons of rectificatory justice to intervene in nature in addition to our beneficence-based reasons.

[46] A version of this paper appeared in *Philosophia.* See O'Brien (2022a).

# Directed Panspermia, Wild Animal Suffering, and the Ethics of World-Creation

**Abstract:** Directed panspermia is the deliberate seeding of lifeless planets with microbes, in the hope that, over evolutionary timescales, they will give rise to a complex self-sustaining biosphere on the target planet. Due to the immense distances and timescales involved, human beings are unlikely ever to see the fruits of their labours. Such missions must therefore be justified by appeal to values independent of human wellbeing. In this paper I investigate the values that a directed panspermia mission might promote. Paying special attention to the outcome in which sentient animals evolve, I argue that we have strong reasons to believe the value of a mission would be negative. Research on wild animal suffering has shown that there is a huge amount of suffering among wild animals on Earth. I argue that there are structural features of evolution by natural selection which explain the prevalence of suffering on Earth, and make it predictable that suffering would prevail on the target planet too. Finally, using insights from procreative ethics I argue on non-consequentialist grounds that creators have duties to their sentient creations which cannot be met in directed panspermia missions.

## 1 Introduction

Directed panspermia is the deliberate transport of microorganisms through space with the aim of establishing them on previously lifeless planets. Two such projects have recently been proposed by Mautner (2010) and Gros (2016), and both are feasible given current or near future technology. Both projects involve sending unpiloted crafts to distant planets and seeding them with microbial life in the hope of establishing a self-sustaining biosphere. Once the target planet has been seeded with suitable microbes, natural selection will take charge and determine the future of the biosphere without human oversight or control. Given the vast timescales and distances involved it is highly unlikely that human beings will ever see the

fruits of their cosmic labours.[47] If these missions are justified, it must be because they promote values independent of human wellbeing.[48] In this paper I will argue that the problem of Wild Animal Suffering (WAS) gives us a strong reason not to launch a directed panspermia mission.

*1.1 Outline of paper*

The moral problem of WAS has gained significant attention among philosophers in the past decade, and the brutality of life in the wild has led some to argue that we ought to intervene in nature to reduce WAS. Little has been written about what WAS implies for spreading animal life beyond the Earth, however. Tomasik (2014) has speculated that future space colonization efforts, including directed panspermia, might result in the proliferation of WAS. However he doesn't provide any reasons for us to think that just because evolution by natural selection has resulted in a great deal of suffering on Earth the same will happen if we spread life to other planets, nor does he consider any of the positive values that might be furthered by directed panspermia. In section 2 I will sketch the possible outcomes of a directed panspermia mission and assess their value. I will pay special attention to the outcome in which sentient animal life emerges on the target planet, and I will argue that we have strong reason to believe that the value of such an outcome would be negative. I give two arguments for the conclusion that directed panspermia missions will result in a massive amount of WAS. First, a simple inductive argument – since it is likely that evolution by natural selection has resulted in either net suffering or net bad lives in Earth's biosphere, we have reason to believe

---

[47] Gros (2016, p.8) targets extrasolar planets within a radius of 100 light-years from Earth. After seeding it will take hundreds of millions of years for complex life to evolve.
[48] This is not strictly true. A successful mission could increase the wellbeing of some humans if a desire-satisfaction or an objective list theory of welfare that includes achievement is true, especially if it includes posthumous benefits. It remains unlikely, however, that any human being will ever see the results of a mission.

that the same will happen on other planets. This inductive argument may be implicit in Tomasik's paper, but as we are making inferences from only one planet, it is weak evidence.

My second argument is more fundamental, and gives us stronger reason to believe that directed panspermia will result in significant WAS on the target planet. I argue that there are certain structural features of evolution by natural selection which make it very likely that it will result in either net suffering or at least net bad lives wherever sentient life emerges. These structural features are necessary to evolution by natural selection, wherever it occurs. Even if evolution is highly unpredictable and significantly influenced by contingent factors, these necessary features are likely to give rise to suffering wherever sentient life evolves. If this is correct, then, given some plausible assumptions about the likelihood of the various outcomes of a directed panspermia mission, its expected value is negative, and this gives us a consequentialist reason against it.

In section 3 I will investigate the morality of launching directed panspermia missions from a non-consequentialist perspective. Using insights from procreative ethics I will argue that creators have responsibilities towards their sentient creations which cannot be met in directed panspermia missions. By launching a successful mission, we may become responsible for the creation of quadrillions of sentient animals who would never otherwise have existed. If we have even weak duties towards our creations, and the arguments in section 2.4 are correct, then creation by directed panspermia is impermissible. Finally, in section 4, I will consider four objections to my arguments. First, in section 1.2 I will give some definitions and briefly describe the missions proposed by Gros and Mautner.

*1.2 Definitions and technical details of proposed missions*

*Panspermia* is the hypothesis that life exists throughout the universe, and is spread through space naturally, for example by cometary collision. *Directed panspermia* is the deliberate

transport of microorganisms through space with the aim of establishing them on a previously lifeless world. It was proposed by Shklovskii and Sagan (1966) and Crick and Orgel (1973) that life on Earth may have been seeded in this way. *Directed panspermia missions* would be launched with the aim of spreading microbial life to previously uninhabited planets (or other habitable bodies) in the hope that the microbes will survive and eventually give rise to a functional biosphere. We can classify directed panspermia missions as part of the broader category of *life propagation missions*. Other kinds of life propagation missions might involve direct human colonization of Mars, or the establishment of an Earth-like biosphere of terrestrial plants and animals on a space station or a suitably terraformed planet. Life propagation missions could vary widely in their aims, methods, scope, and degree of human control and oversight. Directed panspermia missions involve seeding a planet with primitive life, then allowing it to develop by natural selection without any human control. Such missions can be described as *unguided life propagation missions*, since, though they are originally launched by human beings, the subsequent development of the biosphere is left to natural selection. A *guided life propagation mission* by contrast would be one where the development of life in the new biosphere would be guided by human beings, or perhaps by artificial intelligence.

Though all life propagation missions raise moral concerns, in this paper I will restrict my attention to the specific moral problems raised by unguided life propagation missions, particularly directed panspermia missions. I concentrate on unguided missions since they most clearly raise the problem of spreading WAS to other planets. Guided missions, in which human beings or artificial intelligences are on hand to guide the development of the resulting biospheres, might also involve significant WAS. In principle, however, WAS could be avoided in guided missions if humans are both willing and able to bring this about. I will also refer to guided and unguided biospheres. Unguided biospheres are those in which the

conditions of life are not controlled or guided by any intelligent beings, and, consequently, the fates of any animals in them are determined by the forces of natural selection. This, I will argue, will very likely result in mass suffering. Guided biospheres are those in which the conditions of life are determined or controlled by some intelligent beings. In principle, guided biospheres could be managed in such a way as to dramatically reduce WAS, though, of course, this depends on the motivations and values of the guardians.

*Directed panspermia missions*

Life on Earth seems to have arisen quickly, perhaps only a few hundred million years after the formation of the planet. Complex life however took significantly longer to develop. Photosynthesising bacteria evolved around 3.5 billion years ago, and their waste product, oxygen, allowed for the evolution of the more complex eukaryotic cell around 1.85 billion years ago. The first multicellular organisms appeared around 1.7 billion years ago. The Cambrian Explosion 525 million years ago brought greater complexity and diversity to terrestrial life and saw the emergence of the vertebrates and other major phyla. Gros' Genesis Project aims to seed suitable extrasolar planets with eukaryotic cells, thus 'fast-forwarding' evolution on the planet by several billion years, ultimately culminating in a self-evolving biosphere of complex organisms. The idea is to send small spacecraft to suitable planets outside of the solar system. On arrival in orbit, the onboard AI will scan the planet, determining whether there is already life there. If there is, the mission will not proceed.[49] If there isn't, the AI will assess the precise conditions on the planet, and using its onboard gene laboratory it will synthesize the kinds of micro-organisms best suited to the conditions on the planet. These micro-organisms will be launched at the planet in tiny shielded capsules. This

---

[49] This avoids the moral problem of disrupting native life on the target planet, as does Gros' proposal to target 'transiently habitable' planets which are unlikely to harbour native life, particularly those rich in primordial oxygen which may make it more difficult for life to get started. See Sivula (2022) on the moral problem of contamination of planets with native life.

seeding will last for hundreds of years, with the composition of the micro-organisms changing based on changing conditions on the planet. The ultimate aim of the mission is to bring the planet to a 'pre-Cambrian' state, in which complex life can evolve independently along its own course.

Mautner (2010, 2009) has also proposed that humanity seed the galaxy with life. He proposes that humanity launch swarms of tiny probes at suitable targets in the galaxy. Each probe would have a radius of 20 microns, weigh .1 micrograms and contain 100,000 microorganisms, a mixture of extremophiles, cyanobacteria, and even hardy multicellular organisms like rotifer eggs. These swarms would be launched towards extrasolar planets, but also to accretion disks around new stars where planets may one day form, and even to star forming clouds such as Rho Ophiuchus 520 light-years away. Once captured in orbit around the target, some of the probes will be caught up in meteoric dust that will eventually land on a planet's surface. Others will be captured by comets where they will be frozen and preserved until the comet crashes into a planet, while others will be captured by asteroids where they may be able to survive and multiply. To be successful a probe must be launched accurately at the target, it must decelerate and be captured in the target zone, be delivered to a planet's surface along with meteoric dust or by a comet impact, and the microorganisms must survive and evolve on the target planet. Each probe has only a tiny chance of success, but by launching a sufficient number of probes, success can be assured. Mautner estimates that it would cost a billion dollars to seed dozens of solar systems with life for eons.

This may sound like science fiction, but similar missions are already technologically feasible. In 2019 SpaceIL, a private organisation, launched a lunar mission. Its payload included dehydrated tardigrades, a hardy lifeform capable of surviving in suspended animation for decades before being revived (Virk, 2019). The probe crash-landed, and the absence of liquid water on the moon ensures the tardigrades won't revive even if they

survived the crash. Despite their failure, this shows that seeding bodies in our solar system is a live possibility, and SpaceIL have already announced a second mission to an undisclosed target. Given the proliferation of well-funded non-state space agencies in recent years it may only be a matter of time until one of them seeds a planet with viable life, deliberately or accidentally. Extra-solar missions would be more complicated, but still feasible. Even if it's unlikely that a directed panspermia mission will be launched soon, the huge amount of suffering that could result gives us reason to take the possibility seriously.

## 2 Outcomes and Values

In this section I will describe four possible outcomes of launching a directed panspermia mission and assess their value. I will proceed by assuming that there is a particular planet *P,* the future history of which will be determined by the actions of human beings. By sketching the value profiles of the various possible future histories of *P* I hope to clarify what is at stake in launching a mission.

### *2.1 Possible outcomes of a directed panspermia mission*

Once we have launched the probe, the development of *P's* biosphere is entirely out of our control. The course that evolution takes will depend on multiple factors, and we cannot predict what kind of lifeforms will emerge. Though we cannot make any specific predictions about the resulting biosphere, we can categorise the possible outcomes into four broad categories:

B0 – Total failure of the project. The microbes die and no biosphere is created.

B1 – Simple biosphere established. The microbes take hold, and eventually a self-sustaining biosphere of simple, insentient organisms emerges.

B2 – Complex biosphere established. After several hundred million years of evolution, sentient animal life evolves.

B3 – Intelligent life emerges. One or more species of intelligent animals evolves and goes on to form a technological civilization.

Any assignment of probabilities to these outcomes would be somewhat arbitrary at this point. For now, I will proceed on the assumption that all four outcomes are possible, and that Earth is not an atypical example of the evolutionary process.


*2.2 How valuable would these outcomes be?*

I will investigate the value of these four outcomes in terms of three axiological frameworks. These frameworks are supposed to be indicative of three general theories of value broadly construed – a biocentric theory according to which life itself is valuable in some way, a hedonistic theory according to which only mental states can have positive or negative value, and an objective list theory that includes some characteristically human values.

V1 – Life as such is good. Even simple insentient life has intrinsic value.

V2 – The experiences of sentient beings can have positive/negative value. Suffering is bad. Pleasure, and other positive mental states are good.

V3 – Human-like intelligence allows for the pursuit of things which are of significant positive value, such as knowledge acquisition, achievement, the appreciation and creation of beauty, etc.

These values are arranged in order of the complexity of the biosphere necessary to produce them. V1 can be realised on any planet with life, V2 requires sentient animal life, and V3 requires intelligent beings to emerge. Now we can analyse the value of the possible outcomes of a directed panspermia mission in terms of these values. B0 has no positive value, and no negative value aside from the cost of the mission so I won't consider it further. It is

worth noting, however, that if the cost of launching a mission is high, then the opportunity cost of using the money for the mission, instead of investing it into some other long-term project, may be very high. B1 has some positive value if V1 is true, though the degree of value depends on how we understand the value of life. I will investigate this further below. B2 will probably have both some positive and some negative value if V2 is true. The degree of value will depend on how much wellbeing and suffering there is in the future history of the biosphere. In 2.4 I will argue that we should expect this value to be negative. B3 could have significant positive value if either or both of V2 and V3 are true, though there is a risk of tremendous disvalue too.

*2.3 The value of a simple biosphere (B1)*

Is life intrinsically valuable? Mautner and Gros both appeal to the intrinsic value of life as justification for their proposals, though it is not clear exactly how to make sense of the idea. It's difficult to make precise, and upon reflection it seems that much of the value of life depends on factors not present in B1 considered as an end state of *P*. First, we must exclude any consideration of the instrumental value of life from our evaluation since there will be no one around to utilise it. Secondly, we must ignore the potential that a simple biosphere might have to develop into a complex one. Finally, if life is valuable aesthetically, we must consider whether beauty is valuable in the absence of observers. If not, B1 loses much of its value.

There are several positions in environmental ethics that argue for the extension of moral consideration or intrinsic value to insentient nature. Biocentrism is the view that individual living things have intrinsic value and that we ought to promote and protect their good, ecocentrism extends moral consideration to ecosystems, and naturocentrists argue that

there is value in natural areas unmodified by human beings.[50] None of these positions as traditionally understood is likely to endorse a directed panspermia mission, concerned as they are with the conservation and protection of currently existing organisms and ecosystems rather than their propagation. For B1 to have positive value then we must interpret V1 in a propagationist way, and we must think that life has value independently of its instrumental, potential, or observer-dependent aesthetic value.

Mautner (2009) argues for a propagationist account of biocentrism which he calls panbiotic ethics, the goal of which is to maximize life in all accessible habitats. He argues that the unique complexity of life, the unity of all lifeforms, and the apparent self-propagating purpose of all organisms supports a panbiotic ethic. His central point is that all organisms act as if they have the purpose to survive and reproduce, and that this 'effective observable purpose' is, in all important respects, equivalent to a genuine purpose. Since we humans are also living beings we share this effective purpose. From this Mautner draws the conclusion that the 'human purpose' is to safeguard and propagate life.

One problem with this argument is that all organisms aim at self-propagation, not at the propagation of life generally, and we cannot infer the latter from the former. This problem can be mitigated somewhat by the unity of all terrestrial lifeforms, which share a large number of the same genes. By spreading the genes of distantly related microbes we will also be spreading some of our own genes. This way of understanding Mautner's argument would transform his panbiotic ethic into a kind of pragmatic second-best option – the ideal for each of us would be to spread copies of ourselves throughout the galaxy, but since this is impossible the next best thing is to spread microbes which share a portion of our genetic material. We can strengthen Mautner's argument by appealing to the principle of

---

[50] See Taylor (1986), Agar (1997), Callicott (1989), and Elliot (1982) respectively. For a recent critique of biocentrism see MacClellan (2023).

universalizability. If I believe that my self-propagating purpose is a valuable end, then, on pain of irrationality, I must consider the self-propagation of other lifeforms to be valuable too. In this way we can justify a panbiotic ethic and see the self-propagation of all organisms as a valuable end which human beings, in virtue of their technological prowess, are uniquely placed to achieve.

It's plausible enough that life has value due to some combination of its unique complexity, our affinity for living things, its beauty, and its purposeful nature. How can we quantify this value? One way is to imagine different possible worlds and compare our intuitions about the value of each. Imagine then that we are to choose between these two universes:

> *The Dead Universe* – This universe contains no lifeforms of any kind and no sentient beings, and it never will.

> *The Green Universe* – This universe contains an abundance of complex and diverse insentient lifeforms. There are no sentient beings, and there never will be.

Faced with this choice I think that most people would choose the latter. Now imagine the choice is between the Green Universe and:

> *The Happy Universe* – This universe contains an abundance of complex and diverse lifeforms, including sentient animals living generally happy lives.

Again, this seems like an easy choice. The latter universe is obviously more valuable, as it contains both insentient life and sentient happy animals. But how much of a difference does the presence of happy animals make? To determine this, we should compare the Green Universe with:

> *The Happy Planet* – This universe is totally devoid of life of any kind, except for one planet which is populated by happy animals. Life will never spread beyond this single planet.

My own intuition is that the Happy Planet is vastly more valuable than the Green Universe. Others' intuitions may differ, but I suspect that most people will assign significantly greater value to happy sentient life than to insentient life.

How valuable then is B1? If we assume that biocentrism is true, that B1 would be as ecologically rich as Earth (despite its lack of sentient animals), and that the value of spreading life into new environments is equal to the value of conserving that which already exists, then B1 has significant value. I imagine that few people are thoroughgoing propagationist biocentrists, however. For the rest of us, once the extrinsic values of life are stripped away there is little left. Even if we accept a version of Mautner's panbiotic ethic, the value of B1 will, for most of us, be small relative to the value of a world with sentient happy life. Furthermore, that value will be entirely impersonal – while the biological richness of B1 may be good, it isn't good *for* anyone.

*2.4 The value of a complex biosphere (B2)*

The value of B2 depends on the total balance, and perhaps the distribution, of positive and negative wellbeing of all sentient beings in the future history of *P*. I will argue in this section that we have strong reason to believe this value will be negative. There are two ways of understanding this claim, a weaker and a stronger. The weaker claim is the Net Bad Lives claim applied to *P*, that is, that most lives on *P* will be net negative – most animals will suffer terribly and die young with few positive experiences. Since most such lives are very short however, it may still be the case that the aggregate wellbeing of all animals on *P* will be positive, if the longer-lived animals have sufficiently good lives. The stronger claim is the Net Suffering claim, that is, that the aggregate wellbeing across all lives on *P* will also be negative. I will argue that it is overwhelmingly likely that the weaker claim is true, and plausible that the stronger claim is too. First, using the existing literature on wild animal

suffering, I will argue that it is likely that net wellbeing of wild animals on Earth is negative. If Earth isn't an atypical example of the evolutionary process, we can infer that a similar outcome on $P$ is likely, especially since we will be seeding it with terrestrial microbes. Since we have knowledge of life on only one planet, however, the inferential argument is weak. Secondly, and more fundamentally, I will argue that there are structural features of evolution by natural selection which are likely to result in mass suffering wherever sentient life evolves in an unguided biosphere. These features explain why evolution has created so much suffering on Earth, and allow us to predict that B2 would likewise be filled with suffering.

(1) Wild Animal Suffering on Earth

Recent work on WAS has shown that a naïve, idyllic view of life in the wild is unsustainable.[51] Animals suffer a great number and variety of harms: predation and parasitism, intraspecific competition, sexual conflict, disease, extreme weather conditions, hunger and thirst, physical injuries, and accidents all impose tremendous amounts of suffering on trillions of animals each year.[52] It is beyond doubt that there is much suffering in Earth's biosphere. But is the Net Suffering claim true? Ng (1995), Tomasik (2009), and Horta (2010) have argued convincingly that it is. Their arguments are based on population dynamics, and the impact that reproductive strategies have on the global welfare balance. As we saw in the introduction, there are two main reproductive strategies among Earth's animals – r-selection and K-selection. K-strategists give birth to a small number of offspring in whom they invest a great deal of energy, either in the form of parental care or by giving birth to fewer but better developed offspring, laying nutrient rich eggs, etc.. The vast majority of animals on Earth, however, are r-strategists who provide little or no parental care to their

---

[51] See Tomasik (2009b), Horta (2010), McMahan (2010, 2016), Faria (2016, 2023) and Johannsen (2017, 2021).
[52] See the Animal Ethics (2023c) page on The Situation of Animals in the Wild for detailed accounts.

young.[53] Instead, they reproduce in vast numbers in order that that some of them will survive to adulthood.

Unfortunately r-selection entails very high infant mortality rates, and consequently a great deal of suffering. Sea turtles for example lay an average of 110 eggs in a nest and have between two and eight nests each mating season. If the majority of the hatchlings survived then their population would grow rapidly. This is not what happens. Rather, since the population of sea turtles (and of species generally) remains fairly constant between generations, we can deduce that, on average, only one child per parent will survive and reproduce. The others die shortly after birth, being eaten by predators, dying of dehydration, or starving. Clearly all of these are terribly painful ways to die. Are there any good things in these short lives that might compensate for the badness of their painful deaths? These animals receive no parental care, so they don't experience any emotional connection with their parents. They die before reproductive age, so they don't experience the pleasures of mating. At best perhaps some of them manage to find some food and enjoy the pleasure of eating before their deaths. Considering the terrible suffering they endure while dying it's clear that these animals don't have lives worth living – the costs to them of coming into existence far exceed the benefits. Since most animals are r-strategists this is the fate of the vast majority of animals who come into existence.[54]

One might accept that the vast majority of animals have lives like this, yet deny that they are sentient. There is growing evidence of sentience among vertebrates and some invertebrates, even prior to hatching or just after birth, but there is still a great deal of uncertainty in this area.[55] However, as Horta (2017) and Johannsen (2020) have pointed out, the number of animals who die in these ways is so vast that if even a tiny percentage of them

---

[53] 'Majority of animals' here refers to individual animals, not to species, though it is true of both.
[54] See Horta (2010) and Johannsen (2017) for further examples.
[55] See European Food Safety Authority (2005), Duncan (2006), Mellor and Diesch (2007), and Birch (2018).

are sentient this constitutes a huge amount of suffering. The evidence that suffering prevails on Earth is strong. Can we infer from this that the same will be true on *P*? Tomasik (2014) states that animals on other worlds would "suffer enormously as a result of living in Darwinian ecosystems in which predation, disease, and premature death are endemic". Since we have knowledge of only one biosphere the evidence for this inference is weak, but in the absence of any arguments to the contrary it gives us at least a prima facie reason to believe that suffering will also prevail on other planets.

(2) Structural causes of suffering inherent to natural selection

Another weakness of the inferential argument is that it relies on evolution by natural selection being relatively predictable, but it is controversial just how predictable evolution is.[56] If evolution is highly sensitive to contingent factors then we cannot predict that the results of that process on *P* will be similar to those that occurred on Earth. In this section I will argue that, even if evolution by natural selection is generally unpredictable, those features of it which give rise to WAS are necessary, and hence large-scale suffering is the predictable outcome of the evolutionary process, even if we cannot predict the precise form it will take. The ultimate causes of suffering in our biosphere are scarcity and antagonism. Scarcity is most obvious in r-strategists who reproduce in quantities far above the carrying capacity of their environment, dooming many of them to starvation or dehydration. Antagonism can take many forms including predation, parasitism, and sexual conflict. The root cause of both is misalignment of interests. Scarcity results from the conflicting interests of genes and individual animals. Antagonism results from the conflicting interests of animals, which is itself caused by the conflicting interests of their competing genomes. Though we cannot

---

[56] For discussion of this question see Conway Morris (2003) and Gould (1989).

predict what kinds of animals will evolve on *P*, we can be confident that if sentient beings evolve there will be suffering caused by these misaligned interests.

*Why the interests of animals and genes conflict*

Genes are insentient – they have no interests in the sense of some states of affairs being better or worse for them. We can speak of genes' interests in the looser sense in which maximum self-replication is 'better' for a gene. Sentient animals have interests in pursuing pleasure and avoiding pain, remaining healthy, and continuing to live. The interests of an animal and his genes can coincide when he is living a successful life for an animal of his kind. Mating is good for him because it is pleasurable, and it is good for his genes as their means of replication. Genetic-interests and animal-interests are different, however, and often diverge. The most obvious example is senescence. From a gene's point of view, the body it is in is disposable – what matters to genes is replication. Ageing may be an example of antagonistic pleiotropy.[57] This is when a single gene has multiple effects, some good and some bad. Ageing could be explained in this way if the genes responsible for senescence also have fitness enhancing effects in early life. If this makes the animal more reproductively successful then it serves the interests of its genes, though of course ageing and death aren't in the animal's interests.

This conflict between genetic-interests and animal-interests also explains scarcity and the suffering it causes. As we saw above, the most successful and widespread reproductive strategy on Earth is r-selection, whereby an individual will have thousands or even millions of young, of whom only a tiny fraction will survive. From the point of view of the genome, this strategy makes perfect sense – the more the animal reproduces, the more likely it is that at least some of the offspring will survive to reproduce themselves, thus ensuring the

---

[57] See Williams (1957) and Austad and Hoffmann (2018).

replication of the genes. However, in a world of scarcity and fierce competition for resources, the vast majority of these animals will suffer and die – from the point of view of the individual animals involved, r-selection is a disaster.

But why should we expect evolution to result in r-selection and scarcity? Both are predictable consequences of the maximizing nature of genes. Competing genomes can be understood as unconscious players in a prisoners' dilemma game.[58] All genomes aim at maximizing their own replication. This requires resources, of which there is a limited supply in the environment. Of course, it would be 'better' for competing genomes if they could somehow agree to limit their replication to utilise all the available resources as efficiently as possible, thus maximizing the total amount of gene replication. This is impossible, however, for familiar reasons. If 'co-operative' genomes were to limit their own replication, they would quickly be outcompeted by 'defecting' genomes that increased their own replication beyond the 'agreed' level. Genes that cause organisms to reproduce more than others will tend to proliferate, and it is individual sentient animals who pay the cost of this over-replication through starvation. The maximizing nature of genes determines reproductive strategies and locks animals into an eternal Malthusian trap. There is no reason to think that this dynamic is unique to Earth. Animal-interests and genetic-interests will never converge precisely, and, since genes determine the life prospects of animals, genetic-interests will generally prevail over animal-interests when they conflict.

Scarcity of some sort seems to be inevitable given the maximizing nature of genes and the simple fact that every world has finite resources. Even human beings, members of a species that is certainly on the K end of the r/K selection spectrum, have suffered greatly from scarcity throughout most of their evolutionary history. The more severe kind of scarcity, in which huge numbers of animals starve to death in infancy, seems to depend on the

---

[58] See Axelrod and Hamilton (1981).

presence of r-strategists, however.  Is r-selection an inevitable outcome of natural selection in an unguided biosphere? How likely is it that r-selection will become the dominant reproductive strategy on $P$?

Some animal speies on Earth have evolved to use the K-selection strategy and consequently have much lower infant mortality rates than r-strategists. Could the evolutionary process on $P$ result in a biosphere where most or all animals are K-strategists, thus resulting in a world with much less suffering than our own? I think this is unlikely for several reasons. First, the vast majority of animal species on Earth are r-strategists. According to the International Union for the Conservation of Nature (IUCN, 2014) over 95% of animal species so far described are invertebrates, the vast majority of which are r-strategists. Of the 66,000 described vertebrate species more than 50,000 are fishes, amphibians, and reptiles, among whom parental care is rare, though some of these may still be K-strategists who invest more energy in their offspring without parenting them, e.g. by having fewer but better developed offspring. This prevalence of r-strategists on Earth suggests that K-selection would also be rare on $P$. Secondly r-selection must evolve earlier than K-selection, at least when we are talking about those complex social K-strategists who invest more energy in their offspring by providing them with parental care. These social K-strategists are complex animals who are capable of learning and responding flexibly to their environment, but of course complex animals must evolve from simpler ones who are overwhelmingly r-strategists. Parental care is a later evolutionary innovation adopted by complex animals living in stable environments, and on Earth at least it has been adopted by only a tiny proportion of animals. Finally, K-strategists may depend on the existence of large numbers of r-strategists. K-strategists are typically relatively large and complex compared to r-strategists. The logic of trophic chains suggests that small simple animals must be much more common than the animals higher up the trophic chain that feed on them. Since the bottom levels of trophic chains contain many

more animals than the higher levels, and since these are usually small r-strategists, it seems probable that in any biosphere r-strategists will vastly outnumber K-strategists. Of course, long trophic chains aren't necessary – there are islands on Earth without predators on which the trophic chains are short. Such ecosystems are a tiny minority, however, suggesting that predation and long trophic chains are the norm.[59]

The maximizing nature of genes, the fundamental conflict between animal-interests and genetic-interests, and the fact that environmental resources are limited are necessary facts which will apply to evolution by natural selection whenever and wherever it occurs. r-selection may be a contingent outcome of this process, but the necessary structural features which constrain evolution, and the arguments in the paragraph above, make it a likely and predictable one.

*Why the interests of animals conflict*

Different animals are composed of different teams of genes which are in competition with each other for resources. These genomes have 'chosen' different strategies for replication, and these strategies can conflict. This results in antagonistic relationships at multiple levels: between different individuals of the same species competing for mates and food; between predators and prey; between parasites and hosts; between males and females; even between siblings. Antagonism results in tremendous suffering for animals, and, like scarcity, it is a predictable outcome of evolution by natural selection. Genes must maximise their own replication. Different ecological niches provide different opportunities, and different phenotypes allow genes to adopt different strategies to exploit the resources of their environment. Since these teams of genes aren't identical, they must maximise their own

---

[59] See de Nooijer, Holland, and Penny (2009). The authors ran simulations of simple environments of unicellular organisms under a variety of conditions. They conclude that it is unlikely that there was ever any extended period in Earth's evolutionary history in which predation was absent. In short, it's simply too good a strategy not to emerge.

replication at the expense of other teams of genes. Furthermore, since animals are made of usable proteins and fats, animal bodies become part of the resource base to be exploited.

I am not claiming that precisely the same sorts of conflicts will arise on $P$ as have arisen on Earth. Evolution may be highly influenced by contingent factors, and so we may not be able to predict what sorts of animals will evolve on $P$ nor what specific kinds of interactions they will have with each other. However, given the maximizing nature of genes and the inevitability of conflicts of interest between competing genomes, antagonistic relationships of some sort between sentient animals are very likely to arise, and suffering will result. Antagonism and scarcity, as well as sheer misfortune, are the most likely outcomes of natural selection in an unguided biosphere. The above considerations support both the weak and the strong claim about wellbeing on $P$. If the stronger claim is true, then, on any plausible axiology, the value of B2 is negative. If only the weaker claim is true then the value of B2 may be positive for classical utilitarians, though it will remain negative for suffering-focused axiologies and for any theory that values certain distributions of wellbeing, such as egalitarianism, prioritarianism, or sufficientarianism.


*2.5 The value of an intelligent biosphere (B3)*

'Intelligent life' isn't a well-defined concept. Different animal species possess different kinds and degrees of intelligence, and varying capacities to transmit culture from one generation to the next. It is unclear what counts as intelligent life, and we don't know how likely it is to arise on other planets. I understand intelligence in such a way that on our planet only humans count as intelligent life – the other animals lack language and so cannot transmit culture as faithfully as we can, and they cannot match our ability to manipulate the environment through technology. Perhaps the best definition of intelligence for my purposes in this paper is 'a species capable of understanding and manipulating the world well enough to create vast

amounts of positive or negative value through technology and/or culture'. Since humans are the only species on Earth which evolved this level of intelligence, and since there seems to be such a gulf between our intelligence and that of even the most intelligent of the other animals, I suspect that intelligence of the human level is a rare fluke of evolution, unlikely to arise on seeded planets.

The value of a biosphere that gives rise to intelligent beings has the potential to be very high. We are only in the early stages of our own B3 biosphere, yet the potential of a long human, or post-human future is already dimly visible.[60] It is possible that our descendants will eliminate suffering from Earth, that they will spread throughout the galaxy, reaching heights of knowledge, wisdom, and happiness beyond anything we have ever experienced. And this blissful future has the potential to last for aeons. If a B3 biosphere arises on $P$, we can hope that the intelligent beings there will bring about a similarly wonderful future in their region of the galaxy. In order to do so, many things need to go right. They must develop a technological culture; they must develop an ethical system that requires the abolition of suffering, as well as correctly identifying and promoting that which is of positive value; and they must survive long enough to actually implement that ethical vision. Of course, it's also possible that a B3 type biosphere gives rise to a future of tremendous disvalue, for example by getting locked into a stable but negative value persistent state, such as domination by a global totalitarian regime or a badly misaligned super-intelligent AI.[61] B3 has the potential to be extremely valuable, extremely disvaluable, or anything in between, and we have no way of telling which outcome is most likely. We cannot predict the future of our own society, much less that of other intelligent beings.

---

[60] See for example Bostrom (2008) and Ord, (2020, chapter 8).

[61] A 'persistent state' is a stable state that, if it becomes actual, we can expect to last for a very long time, even indefinitely. On the possibility of being locked into negative value persistent states see Greaves and MacAskill (2021). The fourth paper in this thesis further explores the idea that wild animal suffering is a persistent state with negative value.

*2.6 The expected value of a directed panspermia mission*

I have sketched the possible outcomes of a directed panspermia mission to *P*. The value of these outcomes ranges from extremely negative to unimaginably positive. This shouldn't be surprising – we are talking about creating worlds, and worlds are the kinds of things that can go very well, or very badly. How then can we assess the value of launching a mission? The standard expected value approach involves assigning a numerical value to each outcome which is multiplied by its probability, then adding these values together. This would have to rely on estimates of probabilities in which we can have little confidence. One thing we can be sure of is that higher level biospheres are necessarily less likely to emerge than simpler ones, for the obvious reason that they must emerge from simple biospheres.[62] For the sake of illustrating the possible expected value I will assume that each step up in complexity of the biosphere from B0-B3 has a 50% chance of occurring (i.e. that by launching the mission we have a 50% chance of establishing a simple biosphere, and if a simple biosphere is established there is a 50% chance that it will develop into a complex one and so on). I assume, based on the considerations in 2.4, that if a complex biosphere emerges there is a 20% chance that its value will be positive, and an 80% chance that it will be negative. I assume that if intelligence emerges there is a 50% chance of an extremely good outcome and a 50% chance of an extremely bad one. I've settled on 50% because of the high degree of uncertainty. The intelligent inhabitants of *P* aren't humans. Since we are completely ignorant of their psychology and values, we have no reason to lean one way or the other. Of course, it's possible that the intelligent beings on *P* will bring about a mediocre future, or that their

---

[62] There is a complication here. It's true that complex biospheres are less likely to emerge than simple ones, but that isn't actually the relevant consideration for determining the expected value of launching a mission. What matters is what state most planets are likely to spend most of their time in. It could be that B3 biospheres last long enough compared to other biospheres that the most common end state of a biosphere is B3 even if that is the one that is least likely to initially emerge. That is, B3 biospheres might come into existence less frequently, but then they might last much longer than other kinds of biosphere. In practice, however, I think it is likely that B1, B2, and B3 type biospheres are comparably persistent, and so we can ignore this technical complication.

civilization will collapse early. I ignore these possibilities in my expected value analysis for two reasons. Firstly, when speculating about the future of intelligent life I am more interested in what is distinctive about their possible future, and I take it that what is most distinctive about intelligent life is its capacity to bring about levels of value or disvalue that couldn't arise in a biosphere of simpler animals. Secondly, incorporating mediocre futures into the analysis won't significantly alter the expected value of a mission, since the moderately good and moderately bad futures will effectively cancel each other out. Using these assumptions I came up with the following probabilities. Though I haven't attempted to quantify the value of each, I have assigned between 1 and 3 pluses or minuses to show the estimated relative scale of the values involved.

| End State | Value | Probability |
|---|---|---|
| No biosphere established | -opportunity cost of investment | 50% |
| Simple biosphere established | +value of simple biosphere | 25% |
| Complex biosphere, suffering | --suffering of quadrillions of animals | 10% |
| Utopian world established | +++ value of utopian society | 6.25% |
| Dystopian world established | ---value of dystopian society | 6.25% |
| Complex biosphere, wellbeing | ++wellbeing of quadrillions of animals | 2.5% |

To determine the expected value of a mission we could use this probability table and plug in quantifiable values for each outcome. The expected value will depend on what values and probabilities we assign to each outcome. Based on the probabilities I've chosen the expected value is negative. One could reasonably assign very different probabilities. For example, if one were more confident that intelligent life will create positive value then it might turn out that the value of a long-lasting utopia is so great that it dominates the other

outcomes giving an overall positive expected value. It should be noted that I have probably overestimated the probability of intelligent life emerging, perhaps very significantly. The fact that there are so many habitable planets in our galaxy, many of which are much older than our own planet, suggests that life should have arisen elsewhere by now.[63] The fact that we have seen no sign of any intelligent beings, especially given the speed with which the galaxy could theoretically be colonized, suggests that either life is very rare, intelligence is rare, or both.[64] Given how quickly life emerged on our own planet, it would be surprising if life itself were rare. If simple life is not rare, however, this suggests that the emergence (or persistence) of intelligent life is extremely rare. If this is the case, then the expected value of a mission depends on the expected values of the B0, B1, and B2 biospheres.

The expected value of a directed panspermia mission depends on a great number of empirical assumptions, and the value we place on each outcome. I have argued that we have strong reason to expect it to be negative, primarily because of the high probability that there will be net suffering on any planet on which sentient animal life evolves. Moreover, those who see certain distributions of goods as inherently more valuable than others are unlikely to value the outcomes above, characterised as they are by gross inequalities and the prevalence of lives not worth living. For example, egalitarians, prioritarians, and sufficientarians are unlikely to endorse the creation of a world in which, though there is net positive wellbeing in the aggregate, many beings (or most if the Net Bad Lives claim is true of *P*) have lives that are not worth living. Negative utilitarians, and other consequentialists who give more moral weight to the avoidance of suffering than the promotion of positive welfare are also unlikely to endorse the creation of such a world. Classical utilitarians who are confident that

---

[63] NASA has confirmed that over 5000 exoplanets (planets outside our solar system) have been discovered, and it is expected that many more will be discovered Brennan (2022).
[64] On the speed with which the universe could be colonized by intelligent beings, see Armstrong and Sandberg (2013).

intelligent life will eventually emerge and create enough positive value to outweigh the animal suffering required to get there might expect positive value.

**3 Non-Consequentialist Constraints**

In this section I will look at the arguments against directed panspermia from a non-consequentialist perspective. Though the causal chain is long and the outcome uncertain, if our mission is successful we will have created sentient beings who never would have existed were it not for our creative act. I propose that we look to the literature on procreation for guidance. Obviously there are significant differences between procreation and the non-procreative creation of other sentient beings. However, there are sufficient similarities between them that work in procreative ethics can illuminate our responsibilities towards other beings we create. I will argue that we have responsibilities towards our sentient creations, and that we cannot meet these responsibilities in directed panspermia missions.

*3.1 What can procreative ethics tell us about the ethics of creating animal life?*
Shiffrin (1999) and Velleman (2008) agree that procreation is a morally equivocal act. For Shiffrin procreation is morally hazardous as it involves imposing risks and harms. By bringing a child into existence we may benefit her if we give her a good life, but we also inevitably expose her to harm and risk, in ways we normally think of as morally unacceptable. Velleman takes a slightly different line, arguing that the equivocal nature of procreation comes not from the imposition of a mixed bundle of harms and benefits, but from throwing the child into a predicament, and confronting her with a serious and difficult challenge with very high stakes. If the child fails to rise to the challenges of the life that we have thrust upon her, she will suffer. These factors that make procreation morally equivocal are also present in the case of creating sentient animal life. We expose the animal to harm and

risk by creating her. We place her in a predicament in which she must struggle to flourish in a hostile environment, and if she fails she will suffer and die. These similarities suggest that at least some of the moral constraints on procreation will apply to the creation of animals too. I will argue that directed panspermia missions cannot meet even the minimal conditions of permissible creation.

One might object here that the conditions for permissibly creating animal life are less stringent than those that apply to procreation, and so it would be better to look at our attitudes towards the creation of animal life on Earth. While I agree that our obligations to our children are much stronger than our obligations to other beings we might create, I think that this objection fails. Firstly, the conditions on permissible procreation that I rely on are minimal – a life worth living on the one hand and a reasonable chance of flourishing on the other. While it's very plausible that we owe our children more than we owe other animals we create, it's very implausible that the conditions on permissible creation of sentient animal life would be less than a life worth living, and even this weak constraint is sufficient to make directed panspermia impermissible. Secondly, our intuitions regarding the creation of animal life are seriously skewed by speciesism. We routinely create animals who will have terrible lives on factory farms just so that we can eat them for example. For this reason, I think it is appropriate to use the literature on procreation for guidance.

There is controversy about using the words 'harm' and 'benefit' to describe the good and bad things that we bestow on sentient beings by creating them.[65] I will not attempt to resolve this issue here. I assume that if we create someone, knowing in advance that their life won't be worth living, then we harm them. If we create them knowing there is a high risk that their life won't be worth living, we wrong them by exposing them to that risk. It is important to note that the notion of harm employed here is not the same as that used when we harm an

---

[65] See Harman (2004) and Velleman (2008).

existing individual. To harm someone typically means to make them worse off than they would have been in the absence of the harmful action. Causing someone to exist can neither be better, nor worse for them than non-existence, as we cannot compare the two states. However, as McMahan points out, it is perfectly coherent to say that coming into existence can be good or bad for someone, though not better or worse. Causing someone to exist is good for them if the intrinsically good elements in their life outweigh the intrinsically bad ones. Similarly, being caused to exist is bad for someone if the intrinsically bad elements in their life outweigh the intrinsically good ones. Following McMahan (2013) we can call the harms inflicted and benefits conferred on an individual by causing them to exist 'existential' harms and benefits.

I will make two arguments against directed panspermia, a weaker and a stronger, both based on principles drawn from the literature on procreative ethics. First I will argue that if a life worth living is the minimum criterion for permissible creation, creation by directed panspermia cannot satisfy this condition and hence cannot be justified. Secondly I will argue that by choosing to create sentient life we take on special positive duties towards our creations and that these positive duties cannot be met in directed panspermia missions.

*3.2 The minimal threshold: A life worth living*

There is a range of opinions in the literature about the expected welfare threshold at which procreation becomes impermissible. Feinberg (1986) has argued that a child is wronged by being created only if his life is so bad that non-existence would have been preferable. I will argue that this weaker condition is sufficient to show that creation of sentient animal life by directed panspermia is impermissible:

   (1) A life worth living is the minimal condition for permissible creation of sentient life.

   (2) Typical sentient animal lives in an unguided biosphere aren't worth living.

(3) Directed panspermia missions aim to create sentient animal lives in an unguided biosphere.

∴ directed panspermia isn't a permissible way of creating sentient animal life.


I have defended the second premise in 2.4. One might quibble about my use of the word 'typical'. I use the word here to mean something like 'statistically normal'. Since almost all sentient animals die painful deaths shortly after birth, having had few if any positive experiences, it is fair to describe lives like this as typical. My arguments in 2.4 are meant to show that lives like these are typical not only on Earth, but are likely to be so generally, wherever sentient animal life evolves. The third is true by hypothesis – my argument only targets unguided missions. The first premise is independently plausible in procreation cases. Does it apply to creation too? It might be objected that procreation and creation by directed panspermia aren't analogous and so the life-worth-living permissibility condition (LWL) doesn't apply. Why would this be so? Perhaps our special relationship to our children means that our duties to them are more stringent, and so LWL applies to any children we decide to bear, but not to other sentient beings we choose to create. While it's true that we have special duties to our children, this objection fails. Firstly, LWL is a very weak condition – essentially it tells us not to cause harm. Such a simple negative duty doesn't require any special relationship. Even relational accounts of our moral duties towards animals, such as those of Palmer and Milburn discussed in the first paper, accept that we have negative duties not to harm any animal. It would be odd to accept this in the case of normal harms but reject it in the case of existential harms. Secondly, part of the explanation for the special relationship we have towards our children is the fact that we created them. If our role as creator generates special duties towards our children, then why wouldn't it do so for our other sentient creations?

It might also be objected that creation by directed panspermia is such an unusual form of creation that comparisons with procreation aren't valid. It's true of course that the two modes of creation and their outcomes differ enormously. However, they are morally relevantly similar; in both cases we make a deliberate choice to create vulnerable sentient beings with the capacity to suffer and die. The temporal distance of the outcome is irrelevant.[66] Creatures on *P* will suffer the same regardless of how long it takes for them to evolve – their suffering doesn't matter less simply because it is half a billion years in the future and a hundred light-years distant. In any case, if the temporal distance means that their suffering should weigh less in our decision-making, then presumably the good things in the distant future should weigh less too. If the reasons against launching a mission are weakened by time, so are the reasons in favour of it. Furthermore, as I argued in section 2.4, the suffering of sentient animals in the far future is predictable, so temporal distance does not indirectly reduce our responsibilities towards our creations by rendering the harms unpredictable.

Perhaps the ethics of world-creation cannot be reduced to its effects on individual creatures but must be judged in some more holistic fashion.[67] A holistic defence of world-creation could claim that it is only the biosphere as a whole that has intrinsic value, while its components have merely instrumental value. Such a position would be very close to Leopold's (1949) land ethic, which claims that "A thing is right when it tends to preserve the integrity, stability, and beauty of the biotic community". Such an extreme holism, with its implication that it can be morally right to sacrifice sentient individuals for the sake of the

---

[66] See the fourth paper in this thesis for a defence of the idea that the wellbeing of future animals is as important morally as that of present-day animals.

[67] It might even be the case that our intuitions aren't suited to judging the value of such large-scale acts of creation. World-creation is an act on a scale beyond anything humans have ever done. Our moral intuitions developed in relation to macro-level actions, such as interpersonal violence. Common-sense morality struggles with the morality of 'micro-actions', for example an individual's minuscule contribution to climate change. It may be the case that our intuitions aren't fit for purpose when it comes to such 'super-macro' level acts as world-creation either. I don't have the space to investigate this possibility here; having raised it I must leave it aside.

ecological whole is unattractive in the human case – few of us would think it right to kill innocent human beings even if that were the best way to preserve the 'integrity, stability, and beauty' of the biosphere. If we reject extreme holism in the human case then we shouldn't believe it is acceptable to sacrifice individual sentient animals for the sake of the whole either.[68] Of course, we may be willing to allow a certain number of people and animals to die as an indirect effect of conservation policies. For example, we may think it right to outlaw poaching and the bushmeat trade in order to protect endangered animals, even if we can reasonably expect that this will indirectly result in some human deaths by starvation or poverty. There is, however, a large difference between allowing a number of individuals to die as an indirect effect of policies designed to preserve existing biodiversity on Earth on the one hand, and deliberately creating quadrillions of beings who will predictably suffer and die in order to create more biodiversity in the galaxy on the other. Even if we are willing to accept the former, I think most of us are unlikely to endorse the latter.

A more plausible holistic defence of world-creation might be to accept that individual sentient animals have intrinsic value and that their suffering is prima facie bad, yet to maintain that the creation of an entire world is so valuable that it outweighs the disvalue of their suffering. A proponent of such a view would have to specify exactly what the value of a world consists in and show that this value is significant enough to outweigh the negative value of the suffering required for its creation. The voluminous literature on the problem of evil shows that this is a very difficult task. Since the value of world-creation isn't clear, and the disvalue of suffering is obvious, the burden of proof must be on the defender of world-creation to show that the ethics of world-creation cannot be adequately understood at the atomistic level of its effects on individuals.

---

[68] For an animal rights-based critique of Leopold's holism see Regan (2004).

*3.3 Positive duties to creatures: Beyond a life worth living*

Velleman (2008, p.251-254) argues that the act of procreation generates parental obligations towards the resulting child. This is because the parents are the ones who have 'thrown' the child into the predicament of human existence, and they therefore are the ones who have a responsibility to ensure that the child acquires those capacities it requires in order to flourish in the world. Parents, by their procreative act, have produced a being who is vulnerable to all manner of harms, and who needs a great deal of instruction and care in order to be able to flourish. Since they created the need, they are responsible for meeting it until the child is capable of doing so herself. Our creatures on *P* would also be thrown into a predicament – the predicament of sentient animal life. Like human children they are exposed to risk and harm, and they are confronted with significant challenges for which the price of failure is suffering and death. If we deliberately choose to create them, we have duties of care towards our creatures. I suggest the following threshold:

> *RCF - we have a duty to provide our creations with a reasonable chance of*
> *flourishing.*

RCF is still relatively weak – it doesn't say we need to provide everything our creatures need, or ensure they have wonderful lives. It merely claims that we owe them a reasonable chance of flourishing. The condition is also very general – it will mean different things for different creatures, depending on their capacities. For human children RCF will require education, social activities, a chance to develop skills and so on. For simpler creatures much less will be required, perhaps only a safe environment with adequate resources, and a chance to interact with their fellow creatures. Why should we think that RCF is a reasonable threshold for permissible creation? To me, at least, it seems intuitive that if we choose to create a creature who will have a life that can go well or badly, the minimum we owe them is a decent chance at a flourishing life.

Two arguments can be brought to bear to support this threshold. First, given that there will inevitably be uncertainty about how well a life we create will in fact go, we might think that the cost of this uncertainty ought to be borne by the person who has decided to create the life, rather than by the creature whose life it will be. If we set out to create a life that is just barely worth living, we may in fact end up creating one that is not worth living. A certain 'buffer' should be put in place between a life that is barely liveable and one that is not. The creator ought to aim to confer on their creation a life that is significantly better than just barely worth living. Secondly, it might be the case that the asymmetry between harms and benefits also applies to existential harms and benefits. In the case of existing people, we think that we have stronger reason to avoid inflicting harms than we do to confer benefits. If the same is true of existential harms and benefits, then, in order to adequately compensate for the inevitable harms that will befall our creations, we should only create them when we are confident that the benefits they will enjoy will significantly outweigh the harms. Both these arguments suggest that the threshold for permissible creation must be significantly higher than a barely-worth-living life.

It should be clear that RCF cannot be met in creation by directed panspermia. First, we have little idea what sort of creatures will evolve on $P$, what capacities they will have, and what they will need in order to flourish. Secondly, we know that the typical life in an unguided biosphere provides only a tiny chance of flourishing, characterized as it is by r-selection, scarcity, danger, antagonism, and painful and premature death. And finally, once the mission is launched, we will have no possibility of intervening on behalf of our creatures.


## 4 Objections

I have argued against directed panspermia missions, on both consequentialist and non-consequentialist grounds. In section 2 I argued that the consequences of launching a mission

are likely to be bad, and in section 3 I argued that, regardless of the consequences, as creators we may have duties towards our sentient creations which we cannot fulfil by creating life in this way. In this section I will look at some of the objections that could be, or have been, raised against my position.

*4.1 The duty to preserve life in the universe*

It might be conceded that launching a directed panspermia mission is likely to produce more suffering than positive wellbeing. It might also be conceded that life in an unguided biosphere is unlikely to be worth living, much less to provide a reasonable chance of flourishing, and so creating life in this way is normally impermissible. That is, it might be impermissible to launch directed panspermia missions as a way of promoting or maximizing life. Despite this, however, it may be the case that we have a duty to prevent life in the universe from dying out. Life may be rare; it may even be the case that Earth is the only planet in the universe with life.[69] If life is uniquely valuable, and is confined to only one planet, then we may have a duty to spread life to ensure that it is not destroyed due to some cataclysm on Earth. Given our current level of technology, a directed panspermia mission may be the only practical means of doing so. If this were the case, then our duty to preserve life in the universe might outweigh our reasons not to launch a mission. This objection is compatible with multiple explanations of the value of life. Life could be valuable for its aesthetic properties, its complexity and/or its apparent self-propagating purpose, or because life is a necessary condition for the creation of anything of value. What is important is that life has the right kind of value to ground a duty for us to preserve it. It seems plausible enough that life might have this sort of value, and that an appropriate reaction to that value is

---

[69] Sivula's (2022) 'favourable case' for directed panspermia depends on life being both valuable and rare.

to try to ensure the survival of life, even if we have no duty to maximise or promote life.[70] If

this duty to preserve life is strong enough, it may outweigh our duties not to cause suffering,

and our duties as creators towards our sentient creations.

Even if we have a duty to safeguard life, it's not obvious that we should do so at any

cost. Given the risk of creating huge amounts of suffering entailed by directed panspermia

missions, and the fact that we don't know what sort of life will eventually emerge, it is not

clear that the value of life is sufficient to ground such a duty. For the sake of argument, I will

accept that life has the right kind of value to ground a strong duty to preserve it, that life is

indeed very rare, and that life on Earth is vulnerable to destruction. If this is the case, how do

we balance the duty to preserve life against our duties not to cause suffering, and our duties

as creators? The first thing to note is that there are alternative ways in which we could

preserve life that do not run the risks that directed panspermia missions do. For example, in

the future we could safeguard life by direct human colonization of the solar system, or by

launching generation ships towards other stars. We could even terraform other planets and

take some animal life with us to these new worlds. Even more speculatively, it may be

possible to colonize space by launching ships containing human embryos, and artificial

intelligences which will create a colony on the target planet and use the embryos to create a

human population.[71]

All of these possibilities could avoid the problems I have raised, though they may face

moral problems of their own. If no animals were brought along, then there would be no

problem of WAS, and, possibly, no problem of unfulfillable duties towards sentient creations.

If animals were present, as in the case of terraforming and transporting animals from Earth,

so long as humans are also present, they would be able to guide the development of the

---

[70] Compare Frick (2017) on the final value of humanity.
[71] See Crowl, Hunt, and Hein (2012) and Edwards (2021).
.

biosphere to prevent the worst WAS, if they were motivated to do so. Humans would be on hand to tend to the animals they have created, so it would be possible to meet the duties that creators have to their creations. Whether humans would be appropriately motivated is of course another question – Faria (2023b) suggests that humans might remain indifferent to animal suffering, and so a terraformed planet might contain as much natural suffering as Earth does. If humans were motivated to reduce WAS, and to create biospheres which are not so dominated by suffering as our own world is, it might be possible to spread life in this way without causing huge amounts of WAS.

These alternative ways of spreading life beyond Earth, however, are not currently possible, while directed panspermia missions are. Whether we are permitted to wait for this technology to be developed before we attempt to fulfil our duty to preserve life in the universe must depend on how large the risk is of life on Earth being destroyed during the waiting period, and on how sure we are that life is rare, or unique to Earth. In general, the rarer life is, and the more precarious its existence on Earth, the stronger the duty to launch a safeguarding mission soon. If we think that life on Earth is relatively unlikely to be destroyed during this waiting period, and/or we believe that there is, or will be naturally occurring life elsewhere in the universe, the more permissible it is to wait. If it were considered absolutely necessary to immediately reduce the existential risk to life, we could launch a modified directed panspermia mission to act as a 'dead man's switch'. The probe could be launched immediately, but seeding would only begin if it loses contact with Earth, thus acting as a kind of insurance policy for life. In this way we could safeguard life without causing suffering unnecessarily, though even this may be unjustifiably risky.

*4.2 Designer microbes*

It has been objected by Jeff McMahan (in conversation) and Mautner (personal correspondence) that the ethical problems I have raised could be solved by appropriate use of genetic engineering. The idea is that we could genetically design microbes that would evolve in ways that would result in much lower levels of WAS than we could expect from natural selection operating on unmodified microbes. Essentially, we would be locking out the development of certain traits that are highly conducive to suffering. If this were successful, then directed panspermia missions with suitably designed microbes could eventually give rise to biospheres filled with happy animal life and a minimum of suffering, thus ensuring both net positive wellbeing and the provision of a reasonable chance of flourishing to all our creations.

It is clearly the case that much WAS is ultimately caused by genetic factors. Some proposed solutions to the problem of WAS on Earth involve the use of genetic engineering. For example, Johannsen (2017, p.341-343) has proposed using CRISPR and gene drives to alter the reproductive strategies of r-strategists to reduce the amount of suffering that results from their starvation. He has also proposed (2021, p.74) that it might be possible to reduce or eliminate their capacity to experience suffering. Jeff McMahan (2010, 2016) and David Pearce (2009) have proposed genetically modifying predatory species so that they will no longer need or desire to engage in predatory behaviours that inflict suffering on prey animals. Even more radically, Pearce (1995) has proposed that in the future it may be possible to alter the pleasure-pain motivational structure of sentient animals so that they would be motivated to perform fitness-enhancing actions by a gradient of pleasurable states, rather than the current mixture of pleasurable and painful ones. None of these technologies is fully developed yet, nor is it clear that we have sufficient understanding of the functioning of ecosystems to implement such radical changes without significant unforeseen effects.

However, it seems reasonable to hope that the requisite technology and understanding will be developed in the near future.

It is not clear, however, whether this sort of gene editing would be effective in the case of directed panspermia missions. In the case of existing animals, we can identify which genes control particular traits, and we can modify them to reduce suffering. In the case of a directed panspermia mission, we are talking about editing or designing microbes in such a way that the genetic causes of mass suffering will never even arise in the course of evolution. This seems significantly more difficult. The history of life on Earth shows that the capacity for suffering, r-selection, predation, etc. are highly adaptive traits. We can expect natural selection to exert significant pressure on organisms to evolve traits such as these, and it's not clear that there is anything we could do to prevent these, or other similarly suffering-causing traits, from emerging over hundreds of millions of years of evolution. If it is possible, we can expect the technology to take a long time to develop. It may turn out to be impossible in principle to design microbes in such a way that they can eventually evolve into more complex animals, but without the possibility of developing the traits associated with mass suffering. In any case, if it does turn out to be possible to create these 'designer microbes' which can be expected to evolve in such benign ways, then I would have no serious objections to a mission which utilises only these microbes. Such a mission would count as a guided life propagation mission, though the 'guidance' of evolution would be built into the genetic code of the organisms themselves, rather than being external to them, as in the case of human or AI guardians. My concern is about the harms that will result from unguided missions in which future life is subjected to natural selection without beneficent control over that process. If we can spread happy life throughout the cosmos by directed panspermia, then arguably we ought to do so (perhaps making sure to target only planets that don't already have life so that we don't disrupt existing alien biospheres).

*4.3 Regretting the emergence of life on Earth*

Guy Kahane (in conversation) has objected that my view seems to imply that we ought to regret that life ever emerged on Earth. The idea is that, since on my view evolution leads to net negative outcomes in terms of suffering/positive wellbeing, we have reason not to launch a mission. If it is the case that we should not launch a mission because the outcome is bad, and it is also the case that the outcome of the evolutionary process on our own planet has been bad, then it seems that we have reason to regret that life ever emerged on our own planet. If this is sufficiently counterintuitive, then it ought to count against the conclusion I have drawn.

Kahane (2023) agrees that evolution by natural selection may have caused more suffering than positive wellbeing on Earth, at least so far. Given how many times over we humans are outnumbered by other animals, and how recently in the history of life humans appeared on Earth, it may be that up till now suffering has far exceeded positive wellbeing on our planet. Kahane (p.252) calls this an 'axiological deficit' and suggests that it is far from obvious that humans have, or ever will overturn it. That is, despite all humanity's achievements thus far, we are still 'in the red' axiologically speaking. Even if we highly value the non-hedonic goods which humanity is capable of, and if we adopt a very optimistic perspective on our long-term future, things will need to go "incredibly well, for a long time" for us to overturn the axiological deficit we inherited (p.257). Kahane doesn't give a clear answer on whether we ought to regret that the evolutionary process got started on Earth. However, he suggests that, if sentience is a necessary condition for value, and if human beings produce an especially important kind of value which would not have been realistically possible but for the grisly evolutionary process that led to us, there may be a case for not regretting that that process took place, but for viewing it as an instrumentally necessary evil for the creation of that value.

There are two things to say about this objection. The first is to point out that it is consistent to claim that it is impermissible to deliberately begin an evolutionary process which we have good reason to expect to result in more suffering than positive wellbeing, while also not regretting that this process in fact happened on Earth. There is a degree of personal prerogative to regret, and it is not required of us that we regret that the best state of affairs possible wasn't realised, especially when we ourselves and everything that we personally value was only possible on the impersonally less good course of history. As Kahane himself rhetorically asks, "Must we regret that our parents didn't have children better than us?" (p.268). However, while this may be consistent, it may also place implausibly lax limits on the permissibility of regret.[72] If the hundreds of millions of years of natural suffering on our planet is really as bad as I claim it is, then it would seem somewhat monstrous to prefer the course of history that includes it, just because it also includes us, especially if we fail to have a long and excellent future. It may be better, then, to bite the bullet here, and simply accept that, if humanity will have a short or mediocre future, then we *should* regret that life ever arose on our planet, even if that means regretting a necessary precondition for our own existence and the existence of everything we value. To prefer the course of history which includes our own existence would be to prefer one utterly dominated by suffering.

However, if we are confident that humanity will last long enough, and develop well enough, morally and technologically, to create a long and flourishing future, not only for ourselves, but for all sentient beings, then we may prefer the course of history which led to us, despite the mass of suffering this route through history entailed. If we spread happy life throughout the cosmos there could be astronomically vast numbers of happy lives lived,

---

[72] On our moral reasons to regret the occurrence of historical atrocities that were necessary for our own existence, see Smilansky (2013) and Kahane (2019).

across millions of worlds over billions of years. None of this would erase the half a billion or so years of suffering on our planet, but it would make it a small part of a much greater whole which is overall very good. To the extent that this optimism is rational, then it is also rational not to regret that life started on Earth. This optimism is compatible with the claim that it would be impermissible to start an unguided evolutionary process on another world. To do so would simply mean multiplying the suffering necessary to produce the good outcome. We might think that the 'suffering-cost' of Darwinian-evolution on one planet for half a billion years is worth it if it means spreading happy life throughout the universe. Multiplying the cost by spreading microbes to distant worlds now runs the risk of creating millions more years of WAS without making the good outcome significantly more likely.

I claim not that this response to the objection is decisive, merely that it is rationally defensible. My response relies on an uneasy mixture of pessimism and optimism: pessimism about the prospects of evolution by natural selection producing happy animals or more intelligent beings like us, and optimism about humanity's prospects for securing a good future. Pessimism about humanity's future is perhaps as rational as optimism, and if we are pessimistic about the future then we ought to believe that launching a directed panspermia mission is impermissible *and* that we ought to regret the emergence of life on Earth.[73] Even if we are optimistic about humanity's future, we might reasonably believe in the permissibility of launching missions. For example, a thoroughgoing optimist might not only be confident that humanity will have a good future, but they may be optimistic about the possibility of unguided evolution on another planet resulting in the existence of another intelligent species who will also create a good future in their region of the cosmos. To the thoroughgoing pessimist, I can only point out the huge improvements humanity has made over the last few hundred years. To the thoroughgoing optimist, I can only ask Fermi's question: if evolution is

---

[73] See Crisp (2022) for a defence of the rationality of adopting a pessimistic view of the future.

likely to produce intelligent beings capable of spreading happy life throughout the galaxy, then 'where is everyone?'.

*4.4 Salting the cosmos*

My consequentialist argument against the permissibility of launching directed panspermia missions is based on the badness of suffering and the predictability of mass suffering in any unguided biosphere of sentient beings subject to evolution by natural selection. This argument is open to an intuitively troubling objection. If we have a moral duty not to spread microbial life throughout the universe because the most likely result is net suffering, then what does this imply about our duties should we discover microbial life on another planet? The logic of my argument suggests that we ought to sterilize the planet to ensure that the microbes will not eventually evolve into sentient beings who will suffer net negative welfare in an unguided biosphere. However, it is hard to believe that we would be justified in sterilizing a planet with native microbial life to prevent WAS in its future. This doesn't sound like an act motivated by benevolent concern for the suffering of future beings – it sounds like cosmic vandalism. If we have a strong intuition that sterilizing planets with the intention of preventing suffering in the future is wrong, and my argument suggests that we have strong reason to do so, then it may be that there is something wrong with the argument.[74]

The normal way to proceed when a counterintuitive implication of one's argument or theory is proposed is to either bite the bullet, or find a way of showing that, contrary to appearances, one is not committed to the implication. One approach of the latter sort is to see if better describing the thought experiment will yield other possibilities besides sterilization. For example, we might ask how it is the case that we are capable of (i) finding native microbial life on a planet and (ii) effectively sterilizing it, and yet at the same time we aren't

---

[74] Versions of this objection were put to me in conversation by Jeff McMahan and Beatrice Marchegiani.

capable of simply standing by and guiding the evolution of life on that world so that it doesn't result in massive WAS. Guiding the evolution of these native microbes so that a world of massive WAS is avoided would be the best scenario. It would allow us both to prevent suffering and to preserve a uniquely valuable biosphere of beings with an origin and evolutionary history entirely independent of our own. Let's say it is impossible then – perhaps the technology required to detect and sterilize microbial life on distant worlds is much less complicated than the tech required to benevolently guide evolution. Though, even if that is true, given how long it takes for simple microbes to evolve into sentient animals it isn't clear why we couldn't simply wait until we have the requisite technology. The microbes aren't impatient. If the reason we can't stand by and wait is that we know (somehow) that we ourselves are going extinct soon, then the moral reason to preserve life in the universe becomes operative anyway – if we would be justified in starting a new evolutionary process to avoid the extinction of life, then, a fortiori it seems that we must also be permitted to simply refrain from sterilizing another world for the same purpose.

Let's assume there is some other reason why we are able to detect and sterilize life on another planet but are not able to stand by and guide its evolution. Ought we then to sterilize it? It is possible to simply bite the bullet, but doing so feels deeply uncomfortable. It would be worth investigating just what it is about this choice that makes it seem so repugnant. It could be a general respect for life, even microbial life, that is generating the intuition. This seems an unlikely explanation, however. On Earth, we destroy microbes daily, with little or no compunction. The fact that the microbes are on another planet shouldn't give them special status. We would not feel much reluctance to sterilise a planet that we had inadvertently spread terrestrial microbes to during an exploratory or industrial mission for example. A more plausible explanation for the intuition is the independent origin of life that these native microbes would represent. It might seem monstrous to destroy microbes that had an

independent origin, while it seems fine to destroy terrestrial microbes. Still, we could imagine that before leaving the planet we take extensive samples of all the lifeforms that exist there, so that we can study and preserve them at our leisure. In that scenario, on leaving the planet, we would have the choice to sterilize the planet, or leave it to its own fate. In that case, sterilization still seems wrong, but the explanation cannot be that we are destroying the only example of life with an origin independent of that of terrestrial life. Perhaps the wrong here is in cutting short an independent evolutionary story which is only just beginning. Standing by and guiding the evolution of life on the planet would of course change the story – and remove much of its grisliness – but it wouldn't end it. The wrongdoing would not be in destruction of rare microbes, or in removing the possibility for future wellbeing of creatures that would eventually evolve. Rather, it would be an impersonal kind of wrongdoing in cutting short an independent evolutionary trajectory which has already begun.[75]

Let's assume that there is something wrong about sterilization, and this explanation of the wrong is a reasonable one. Let's also continue to assume that simply standing by and guiding evolution on the planet is not a possibility, nor is life in the universe in danger of extinction. Our only options then are (1) sterilize the microbes, preventing mass WAS in the future, but cutting short an independent evolutionary story or (2) do nothing, and allow mass WAS in the future, while also allowing the evolutionary story to play out. With this clearer description of the situation, sterilization still seems wrong, yet my argument suggests it is right. It would be desirable to be able to point to some moral difference between launching a directed panspermia mission and sterilizing planets which means that the former is impermissible, and the latter is not obligatory. Two differences come to mind, though it is not clear how much weight they can bear. The first is to appeal to the costs involved in the two

---

[75] Compare with Dworkin's (1993) idea of species as sacred in virtue of the historical process of 'natural investment' that led to their existence, and how it is a shame for species to go extinct because of human action.

courses of action. Launching a mission would be costly, but refraining from doing so would not be. At most, we would have to bear some frustration at not being able to proceed with a project which we find valuable, but this is a relatively minor cost compared to the amount of suffering that launching a mission might cause. Sterilizing a planet on the other hand might be costly, and, in general, the more costly a course of action, the less we are obligated to perform it. Saving a life for free might be obligatory, while saving one at the cost of losing a limb might be supererogatory. If sterilizing a planet is sufficiently costly, then we might be under no obligation to do so, even if it would prevent a great deal of suffering. This argument might go some way towards weakening the objection, but it is not enough to overcome it entirely. First, the cost of sterilizing a planet might be relatively meagre, especially compared to the amount of suffering that this would prevent. If that were the case, it might turn out that sterilization remains obligatory. Second, and perhaps more damagingly, this response merely attempts to sidestep the objection by showing that we might not be obliged to bear the cost of sterilizing planets. Doing so, however, would still be a good thing, albeit perhaps a good thing we don't have a duty to do.

The second response to the objection relies on the doing/allowing distinction. By launching a mission we would be acting in a way that would predictably lead to significant harms, whereas if we refrain from sterilizing a planet with native microbial life we would merely be allowing harm to happen. Furthermore, if we recognise the wrongness of cutting short an independent evolutionary story, then sterilization would be an impersonally wrong action. We might think that, if it is generally worse to commit harmful actions than to merely allow equivalent harms to occur, then launching a directed panspermia mission would be impermissible as it causes harm. Failing to sterilize a planet might be permissible, as it merely allows an equivalent harm to occur. It is not clear, however, if the doing/allowing distinction can bear the weight I am putting on it here. First, consequentialists will not

113

recognise the moral importance of the distinction at all, and so for them this response is a non-starter. Secondly, for those who do recognise the distinction, it isn't clear that it would apply in this case. That is, though the reason to prevent the harm in question (mass WAS) is weaker than the reason not to cause it, the scale of the harm in question is so great that it may still be obligatory to act to prevent it. This may be true even if one recognises the wrongness of cutting short the evolutionary story of the microbes, and if there are significant costs to doing so.

This objection seems to be the most damaging to the arguments I have presented in this paper. It is hard to deny the intuition that we would do something wrong by sterilizing a planet, though it is not entirely clear what the wrongness actually consists in, or how significant it is. The responses I have suggested here may allow us to avoid the necessity of sterilizing a planet, but they do so only indirectly – by suggesting alternative options that are preferable when they are possible, by pointing out the asymmetry in costs of the actions, and by bringing up the doing/allowing distinction. One final consideration is that my arguments for the impermissibility of launching directed panspermia missions were of two kinds, one of which emphasized the bad consequences, the other of which concerned our duties as creators. I offered these not as alternatives only one of which need be true, but as two independent and plausible arguments for the same conclusion. The latter argument of course only applies to creatures we actually create – we can have no obligations as creators to beings whom we had no part in creating. If this latter argument is successful, it means that all else being equal it is worse, perhaps significantly worse, to cause suffering by launching a directed panspermia mission than to merely allow it by failing to sterilize a planet. I recognise, however, that at least in some cases, my argument commits me to the unpalatable conclusion that we ought to sterilize other planets to avoid mass WAS in the future.

## 5 Conclusion

Directed panspermia is a technologically feasible and relatively cheap way of spreading life throughout the galaxy. I have argued that we have strong reason to expect the outcome to be bad, primarily because of the predictability of net WAS in any unguided biosphere in which sentient animal life evolves. Furthermore, I have suggested that creators have duties towards their creatures which cannot be met in unguided missions. Even if we think it necessary to safeguard life against the risk of extinction, there are ways of doing so without creating unnecessary suffering. I conclude that such missions are therefore impermissible. I acknowledge, however, that my view may in some circumstances entail regretting that life emerged on Earth, and endorsing the sterilization of planets with microbial life.[76]

---

[76] A shorter version of this paper appeared in the *Journal of Applied Philosophy.* See O'Brien (2022b).

# The Case for Animal-Inclusive Longtermism

**Abstract:** Longtermism is the view that positively influencing the long-term future is one of the key moral priorities of our time. Longtermists generally focus on humans, and neglect animals. This is a mistake. In this paper I will show that the basic argument for longtermism applies to animals at least as well as it does to humans, and that the reasons longtermists have given for ignoring animals do not withstand scrutiny. Because of their numbers, their capacity for suffering, and our ability to influence their futures, animals ought to be a central concern of longtermists. Furthermore, I will suggest that longtermism is a fruitful framework for thinking about the wellbeing of animals, as it helps us to identify actions we can take now that have a reasonable chance of improving the wellbeing of animals over the very long term.

## 1 Introduction

Longtermism is the view that positively influencing the long-term future is one of the key moral priorities of our time.[77] Since the future has the potential to be truly vast, both in duration and the number of individuals who will ever live, it is plausible that the long-term future might be extremely valuable, or extremely disvaluable. If we care about impartially doing good, then we should be especially concerned to ensure that the long-term future goes well, assuming it is within our power to do so. Most longtermists focus on humans, and largely ignore animals. This is a mistake. In this paper I will show that the basic argument for longtermism applies to animals at least as well as it does to humans, and that the reasons longtermists have given for ignoring animals do not stand up to scrutiny. I will argue that, because of their numbers, their capacity for suffering, and our ability to influence their futures, animals ought to be a central concern of longtermists. Furthermore, I will suggest

---

[77] MacAskill (2022) and Greaves and MacAskill (2021).

that longtermism is a fruitful framework for thinking about the wellbeing of animals, as it helps us to identify effective actions we can take in the near future that have a reasonable chance of improving the wellbeing of animals over the very long term.

In section 2 I will lay out the basic argument for longtermism and consider some of the reasons longtermists have neglected animals. In sections 3 and 4 I will show that the basic argument for longtermism goes through for animals and that we can use the longtermist framework to identify interventions that have a reasonable chance of making the long-term future go better for animals. More specifically, I will argue that (1) now or in the near-term future humans can act in ways that will predictably increase or decrease the scale and duration of wild animal suffering in the long term and (2) we are in an especially influential time for locking in values that can be expected to be good or bad for domesticated animals in the long term. Finally in section 5 I will suggest some longtermist interventions for animals that might be more effective than short-term alternatives and will suggest areas for further research.

For simplicity, I'll assume a hedonistic theory of animal wellbeing, though nothing I say will be incompatible with the view that there are also important non-hedonic elements in animal wellbeing. I will assume that all vertebrates have the capacity for sentience, and hence for positive and negative welfare. I will assume the increasingly accepted view that the majority of animals in the wild live bad lives, that is, their lives contain more suffering than positive wellbeing.[78] On the view I will assume, the typical wild animal life is short, contains little or no positive wellbeing, and ends in a painful death. This is because of structural features of ecosystems, such as antagonism, scarcity, and reproductive strategies, which are consequences of evolution by natural selection. Since wild animals outnumber domesticated animals many times over, and since wild animal suffering seems to be much more persistent

---

[78] See Ng (1995), Tomasik (2009b), Johannsen (2021), and Horta (2022).

than the suffering of domesticated animals, I will focus more on wild animals than on domesticated ones. However, I will also consider the effects of human practices on the wellbeing of domesticated animals over the long term.

## 2 The Longtermist Framework

Arguments for longtermism have generally focused on the potential value of the long-term future, though some have also given arguments grounded in our duties to previous generations.[79] Though the future-focused arguments differ in their precise forms, all accept some version of these three central premises: the axiological premise that the intrinsic value of good and bad things is unaffected by their location in time, the first empirical premise that the future has the potential to be astronomically large, and the second empirical premise that there are actions we can take in the present and near future that have a significant chance of making the long-term future go better or worse. Putting these premises together with a principle of beneficence we can construct:

*The Basic Argument for Longtermism*

(P1) *Temporal Neutrality*: the intrinsic value of good and bad things is unaffected by their temporal location. An hour of intense suffering is equally intrinsically bad whether it happens tomorrow or in a billion years, and the same is true of the value of good things. We cannot discount the intrinsic value of good and bad things solely because they are in the future.[80]

(P2) *Big Future*: the long-term future has the potential to be astronomically large, both in duration and the number of individuals who may ever live. Since individual lives can be very

---

[79] See MacAskill (2022), Greaves and MacAskill (2021), Whittlestone (2017) and Beckstead (2013) for future-focused arguments and see Ord (2020 chapter 2) for arguments grounded in the past.

[80] For acceptance of *Temporal Neutrality* see MacAskill (p.10), Greaves and MacAskill (p.18), Ord (p.52), Whittlestone, and Beckstead (p.18).

valuable or disvaluable, the long-term future has the potential to be very valuable or disvaluable.[81]

(P3) *Beneficence Principle*: in general, when we can do so without too high a cost to ourselves, we ought to help others and promote positive value in effective ways.

(P4) *Causal Efficacy*: there are actions we can take in the present and near future that have a significant chance of making the long-term future go better or worse.[82]

(C) *The Longtermist Thesis*: we ought to take actions which we can reasonably expect will make the long-term future go well when those actions are more effective than the short-term alternatives.

One might accept the first three premises but reject *Causal Efficacy*. One could either claim that there are no actions that can be expected to have a positive or negative effect on the long-term future, or that we simply cannot identify such actions. Tarsney (2022, p.7) has called the latter the epistemic challenge to longtermism. It may be that given the difficulties in predicting the long-term outcomes of our actions the best we can do is try to make the near-term future go well. Longtermism might then be an interesting claim about what we ought to value, but it would tell us nothing important about what we ought to do.

Longtermists have developed several conceptual tools which in principle allow them to identify actions that can be expected to have a positive effect on the long-term future, thus securing *Causal Efficacy*. The first is the concept of a persistent state. A persistent state is a

---

[81] See MacAskill (p.12-21), Greaves and MacAskill (p.6-9), Ord (chapter 1), Whittlesone, Bostrom (2013, p.18-19), and Beckstead (p.56-59).
[82] See MacAskill (p.21-28), Greaves and MacAskill (p.9-16), Ord (chapters 3-5, 7), Whittlestone, Beckstead (p.3-8).

state of the universe which, if made actual, can be expected to last for a very long time, perhaps indefinitely (Greaves and MacAskill, p.10). A clear example of a persistent state is a state in which humanity has become extinct. Once humans go extinct that state of the universe will persist indefinitely.[83] Another example is the state in which humanity has come under the domination of a global totalitarian regime the primary goal of which is self-perpetuation (Ord, p.150-152). Greaves and MacAskill use the concept of a persistent state to identify possible actions we can take in the present which will have positive effects on the long-term future despite the general difficulty in predicting such distant outcomes of our actions. They argue that (i) if there are a number of persistent states which could be actualised in the future, (ii) they differ in value, and (iii) there is something we can do now which makes it more likely that the more valuable state will be actualised, then we have identified an action which can be expected to have a significant positive effect on the long-term future. And it appears that all three conditions are met in reality – there are persistent states which differ in value, and humanity can influence the chances of our ending up in these states. Human extinction for example would be persistent, much less valuable than most of the states in which humanity survives, and human beings today have some influence on the probability of our entering that state.

Related to the concept of a persistent state is the concept of value lock-in. A value lock-in is a situation in which the values of humanity (or some significant subset of our values) become so entrenched that they become difficult or impossible to change (MacAskill, 2022 chapter 4). This is important because our values are a key determinant of what the future will look like, and if we lock-in a set of suboptimal values then the future may end up being much worse than it could have been. Ord (p.153-155) gives several examples of value lock-ins which could plausibly make the long-term future go worse. These include humanity

---

[83] Though it's possible that some similarly valuable species would evolve to replace us.

as a whole renouncing all further technological development, permanently failing to recognise some form of harm or injustice and hence perpetuating it blindly, or permanently locking in a single fundamentalist religion. Another way in which value lock-in might give rise to a sub-optimal future would be if a super-intelligent AI took control of the world and enacted whatever values its designers had programmed into it (Ord, endnote 126).

Prominent longtermists have generally concentrated on humans and have had comparatively little to say about nonhuman animals. MacAskill (p.208-213) briefly discusses the welfare of present-day animals but says little about their future. Greaves and MacAskill (2021) don't mention animals as a possible locus of longtermist concern at all. Bostrom briefly mentions the possibility that human beings might have a responsibility to help suffering animals in natural environments (Bostrom, 2013, footnote 23), and Ord (p.212-213) argues that, if humanity survives long enough, we can prevent the natural extinction of animal species over evolutionary timescales, and even transport terrestrial animal species to other planets to prevent their inevitable extinction when the Earth becomes uninhabitable. Beckstead (p.10-11) explicitly argues on longtermist grounds that it is relatively unimportant to reduce the suffering of animals on factory farms.

This neglect of animals could have serious consequences for their wellbeing over the long term. Beckstead's argument is particularly concerning. He argues that reducing the suffering of animals on factory farms is relatively unimportant compared to helping human beings, because animals, unlike humans, cannot be expected to engage in activities that will improve the long-term future (p.22-23). The assumption here is that by benefitting currently existing people we both benefit them directly and make it more likely that they will engage in socially useful activities thus generating indirect benefits for more people in the future. These indirect benefits may compound over time, thus outweighing the initial benefit to the original people. Benefitting animals doesn't provide such indirect benefits. Helping humans then is

like investing into capital, with the expectation that the investment will pay off in the long run. Helping animals is like a moral cul-de-sac – whatever resources we invest into making animal lives go better will not be returned, as animals are incapable of contributing meaningfully to the value of the long-term future. This argument ignores the fact that helping animals isn't necessarily only of value in the short term. Effective ways of helping animals might have long-term positive effects too. Beckstead gives the example of giving pain relief to animals on factory farms. He may be right that this is of only short-term efficacy, helping only those animals we give the pain relief to, without any positive indirect effects. Abolishing factory farming altogether or replacing animal agriculture with synthetic meats is an entirely different matter. Stopping factory farming would not only help animals in the short term, but also prevent the suffering of many animals into the long-term future.

Beckstead (and perhaps other longtermists) may be assuming that animal agriculture is likely to be supplanted by synthetic animal products or plant-based agriculture in the near term, and so efforts by activists to abolish it can, at best, slightly speed up this inevitable transition. If this were true, then activism might not be an effective longtermist intervention. However, the abolition of animal agriculture isn't inevitable. Though we may have the technology to do so soon, such a change depends not only on our technology, but on our values, and as MacAskill points out in his discussion of the abolition of slavery (p.62-70), we cannot assume that such moral progress is inevitable. It is possible that humanity might soon perfect the technology to create 'animal' products synthetically, yet due to suboptimal values they might continue to engage in animal agriculture indefinitely. For example, if humanity maintains speciesist values, and if consumers prefer 'natural' or 'authentic' animal products to synthetic ones, then we may continue factory farming animals indefinitely despite having the technology to do otherwise. It is also possible that we will invent new ways to exploit animals. For example, developments in the technology of xenotransplantation may make the

breeding of genetically modified animals for organ transplants commonplace. It is also already possible to create genetically modified immunodeficient animals for use in medical research. These are animals in which the immune system has been deliberately impaired so that they are more susceptible to various diseases, infections, or tumour growth, and hence more useful for the scientists researching these diseases.[84] As our gene-editing technologies become more powerful it may be possible to create many more kinds of suffering animals for use in research. Finally, it's also possible that our moral values regarding animals get worse rather than better. If that happens, then we might exploit animals in ways that are unthinkable today.[85] We cannot simply assume that moral progress in our treatment of animals is inevitable; rather we must ensure that humanity makes this transition by promoting values that are more friendly to animals. Furthermore, as I will argue in section 4, we may be in a particularly important time in which to try to shift human values regarding animals. If this is correct, then morally informed activism for animals may be an extremely important longtermist concern.

Ord recognises that we may have some duties towards animals, but seems to think of these as duties to animal species rather than to individual animals, and consequently his suggestions for how to help them are potentially harmful. Ord thinks that it would be a good thing for humanity to relocate animal species to other planets to ensure their survival once the Earth becomes uninhabitable. If wild animals generally lived good lives, then spreading them beyond Earth might be a good thing. However, in recent years the problem of wild animal suffering has been increasingly recognized in the animal ethics literature. It's possible that most wild animals live bad lives, and if this is the case then spreading wildlife to other

---

[84] See Hastings and Bala (2005)

[85] Think for example of the kinds of animal 'blood-sports' widely enjoyed in medieval and early modern times.

planets might cause a huge amount of suffering for a very long time. I will discuss the problem of WAS from a longtermist perspective in section 3.

In sections 3 and 4 I will argue that the basic argument for longtermism applies to animals. In section 3 I will argue that we ought to be concerned about how the long-term future goes for animals by showing that *Temporal Neutrality* and *Big Future* apply to them. In addition, I will suggest two further considerations which make it particularly important to think about the long term for animals, based on our future capacities to help wild animals and the higher stakes for animals than for humans. In section 4 I will argue for *Causal Efficacy*. Using the longtermist concepts of persistent states and value lock-in I will argue that there are actions we can take in the near term which have a reasonable chance of making the long-term future go better for animals. If these arguments are successful, then the longtermist thesis applies to animals.[86]

## 3 Temporal Neutrality and Big Future

It is difficult to see how *Temporal Neutrality* could apply to the wellbeing of humans but fail to apply to that of animals. For this to be the case there would have to be some relevant difference between human and animal wellbeing that makes it rational to think that future human wellbeing is as valuable as present-day human wellbeing but the same doesn't hold of animal wellbeing. It is difficult to imagine a candidate difference. As we saw above, Beckstead argues that, from a longtermist perspective, it might be less important to confer benefits on currently existing animals than on currently existing people, as the latter, but not the former, can contribute to improving the long-term future. This is a difference between benefitting humans and animals, but it isn't the right sort of difference to make *Temporal Neutrality* apply in one case and not the other. This is because Beckstead's argument concerns

---

[86] I will assume that the beneficence principle applies to animals, as I have done throughout this thesis.

the comparative instrumental benefits of conferring intrinsic benefits on people and on animals, but what we would need is an argument that allows us to discount the value of future intrinsic benefits for animals but not for humans. Since such a difference isn't forthcoming, I will assume that *Temporal Neutrality* applies equally to animals, and that we ought to be as concerned about future animal wellbeing as present-day animal wellbeing.

In fact, it might be the case that *Temporal Neutrality* applies even more strongly to animals than it does to people. One might think, for example, that some of the strength of our moral obligations towards other people alive today comes from the fact that we are enmeshed in special relations of various kinds. Some people alive today are our family and friends, some are fellow citizens, some are engaged in the same projects we are. Some of these special relations can be expected to disappear or weaken over time, thus reducing the strength of our obligations towards future people to whom we are not so specially related. It is possible, however, that in many cases our duties to animals, especially wild animals, are not based on any kind of special relationships, and so their strength does not diminish over time. If this is right, then while our obligations to future people might diminish somewhat as the relations between us weaken, the same is not true of animals.[87]

It also seems likely that the *Big Future* premise applies at least as well to animals as it does to humans. At present there are around 8 billion human beings on the planet, but we are outnumbered by other sentient vertebrate animals of whom there are at the very least 10 trillion ($10^{13}$) and perhaps more than a quadrillion ($10^{15}$) (Tomasik, 2009a). On conservative estimates, then, animals outnumber human beings by between a thousand and a hundred thousand to one, but the true population of animals may be orders of magnitude higher. Bristlemouth fish alone may number in the quadrillions (Woods Hole Oceanographic

---

[87] See Mogensen (2022) for an exploration of the idea that we might be able to justify an intergenerational time preference by appealing to agent-relative moral reasons based on partiality between generations.

Institution, 2019). This means that most sentient individuals are nonhuman animals. If we assume that animal and human populations will remain at the same ratio in the long term, then the long-term future has the potential to contain far more animal suffering and wellbeing than human suffering and wellbeing. If we are impartially concerned with maximizing wellbeing and minimizing suffering, then we ought to be very concerned about the long-term future of nonhuman animals on the assumption that they will continue to constitute the vast majority of welfare subjects. If *Big Future* inclines us to take longtermism for humans seriously, then it should also incline us to take longtermism for animals seriously.

It might be objected here that it is unlikely that animal populations will remain at their current levels. Animal populations might decline drastically and remain at low levels for the rest of our planet's period of habitability. In fact, it seems that wild animal populations have decreased dramatically in the past fifty years, as human development has reduced habitats for many species, and this process might continue until relatively few wild animals remain. A 2020 report by the WWF suggests that there has been a 68% decline in the average populations of monitored vertebrate species since 1970. It must be borne in mind, however, that this report is concerned specifically with biodiversity, not with animal populations generally, and the findings of the report are compatible with global animal populations staying the same or increasing. For example, the report mentions invasive species as a major cause of biodiversity loss (p.22, 54) but if invasive species are replacing native ones this need not cause a net decrease in animal populations. Similarly, reductions in the numbers of large animals can be more than compensated for by increases in the numbers of small animals. Other causes of biodiversity loss more clearly seem to indicate general population declines, however. Future humans might also eliminate domesticated animal populations, for example by switching from animal agriculture to plant-based or synthetic alternatives.  If this happens

then the *Big Future* premise isn't true of animals, and so the longtermist thesis doesn't apply to them.

In response, we cannot simply assume that animal numbers will drop significantly and remain low indefinitely. This will depend in large part on the values and actions of human beings over the coming centuries, and these are yet to be determined. Even if one were somehow certain that humanity will eventually drive all the other animals to near extinction, in order to do so humanity itself must survive long enough. Ord (2020) puts the risk of human extinction within the next century at about 17%. Newberry (2021, p.4) cites several experts on existential risk who also put the risk at around 20% in the next 100 years. If one accepts that there is a significant risk of human extinction in the next 100 years, then in order to maintain that there is a high probability that humans will drive animal populations to near extinction one would also have to believe that they will do so relatively soon, and that animal populations would never rebound after human extinction. Secondly, as we will see in section 4, there are some ways in which humanity could ensure a long future for animals even if animals do eventually go extinct on Earth.[88] Thirdly, if one believes that animals generally live good lives, or that they could live good lives with human assistance, then one ought to be concerned about animal extinction too, as this might reduce total wellbeing in the long term.[89]

A more fundamental point is that longtermists generally think about the size of the future in terms of expected population, rather than in terms of the most likely outcome or the median across various possible outcomes.[90] In essence this means that they generate high and low estimates for possible future populations, then weigh them according to their probabilities thereby establishing a figure for the expected population. They find that, since human populations have the potential to be truly vast, the resulting figure is also vast, even

---

[88] We investigated one such way, directed panspermia, in paper 3.
[89] Of course, this depends on whether animals are replaced with humans, and whether humans or animals are better 'carriers' of positive welfare. I discuss this question in paper 5.
[90] See Newberry (p.2) and Greaves and MacAskill (p.7).

when they account for the possibility of very low populations or total extinction. As Greaves and MacAskill say:

> Even a 50% credence that the number of future beings will be zero would decrease the expected number by only a factor of two. In contrast, a credence as small as 1% that the future will contain, for example, 1 trillion beings per century for 100 million years… increases the expected number by a factor of 100.

The same consideration applies to the animal population – since it has the potential to be very large, we would need to have an unreasonably high degree of confidence that their population will decline to zero or near-zero in order to endorse a low figure for their expected population. But, if a high expected human population is sufficient to ground longtermist concern for human beings, then a high expected population of animals ought to be sufficient to ground longtermist concern for them too.

There is one final consideration relevant to the application of the *Big Future* premise to animals. Longtermists often bolster their case by arguing that human beings might one day colonise other planets, or that we might one day create human minds on computers, all of which would multiply the possible future population enormously.[91] If we consider such huge expansions to the human population likely, but think it very unlikely that future animal populations will be expanded in the same way, then we might be justified in thinking that the vast majority of future sentient beings will be humans (or digital human minds) and so it is much more important to ensure the long-term future goes well for them than for animals.[92] That is, though the absolute number of animals in the future remains large, it would be

---

[91] See Bostrom (2003, p.2-3), Greaves and MacAskill (p.6-9), and Newberry (p.5-10). See Chalmers (2022) on the possibility, and moral importance, of digital humans.
[92] Rowe (2020) argues that human beings colonizing space but leaving animals behind on Earth is the only realistic scenario in which human beings dominate our welfare considerations.

dwarfed by the human population, and thus rendered irrelevant to our longtermist decision-making.

This is of course highly speculative. There is no reason in principle, however, why these technologies couldn't also multiply animal populations. Tomasik (2014) speculates about the possibility that future human beings might choose to create digital worlds replete with sentient digital animals, perhaps in natural settings similar to those on Earth in which they might suffer greatly. Future humans might also choose to create digital minds in order to maximize hedonic wellbeing, and since animal minds are so much simpler than human minds they might do so by creating digital animal minds rather than digital human minds. A human brain contains around 86 billion neurons, while a mouse brain contains a mere 70 million. Assuming that the difficulty of creating a digital mind increases proportionally with the number of neurons required, then a utility-maximizing would-be digital mind creator might be advised to create digital mice rather than digital people. In section 4 I will consider the possibility that humans might one day spread wild and/or domesticated animals to other worlds. Here I will only say that this possibility is not negligible, and in fact there are some ways in which it is much easier to spread animal life than human life, and so we cannot discount the possibility that animal populations will also be multiplied across physical and digital space. It is possible therefore that animal populations will be multiplied by these technologies just as human populations might be. Even if we think this unlikely, so long as we don't think it impossible, the huge numbers involved serve to increase the expected future animal population significantly.

Of course, it is extremely difficult to make accurate predictions about far future population sizes, for humans or for animals. There are so many known and unknown factors that could cause the population to increase dramatically, fall to zero, or anything in between. This uncertainty may count against the plausibility of longtermism generally but is *not* an

argument against including animals in the longtermist framework. To show that *Big Future* is true of humans but not of animals it is not enough to point to the uncertainty about future population sizes. Rather, it must be shown that there is an asymmetry between humans and animals which means that, while the future human population is large in expectation, the future animal population is not. I have argued above that there is no such asymmetry – if we believe that the expected human population is large, we should think the same about the expected animal population. Furthermore, there are some factors which give us reason to expect a large animal population rather than a large human one. For example, humans are just one species among many. We are an exceptional species, and it is likely that we have the means to survive threats to our existence which other animals do not, and so perhaps we are less likely to go extinct than other species. However, there are millions of species on the planet. Though our likelihood of going extinct might be lower than that of any other species, it is unlikely that our chances of survival are greater than those of all other species on Earth combined. It may be the case then that humans go extinct while other animals live on indefinitely, thus ensuring *Big Future* for animals, but not for us. If it is unwarranted to be confident about the size of the future human population, it is doubly so to be confident that our population will increase and the population of animals will shrink or vanish.

Finally, there are two considerations that indicate that longtermism is particularly important for animals. The first concerns our current capacity to help wild animals, which is far more limited than it is likely to be in the future if we act today to ensure we expand that capacity. It is presently possible to help wild animals in limited ways by vaccinating them against some diseases, leaving out food supplements in times of extreme scarcity, and rescuing and rehabilitating them when they are affected by natural catastrophes. The main causes of wild animal suffering, however, are systemic features of the natural world, primarily reproductive strategies, endemic scarcity, and various forms of antagonism,

including predation, which we currently have little ability to influence. In order to seriously improve the lives of wild animals we need to develop our scientific knowledge of the causes of suffering and wellbeing in diverse species of wild animals, we need to understand how the various components of the biosphere interact so that we can better predict the results of our interventions, and we need to develop the technology to intervene safely and effectively. These developments may take decades or centuries to come to fruition. Those who are concerned with animal wellbeing might therefore think that it is important to begin the research work that will allow us to significantly help wild animals in the future.

The second consideration concerns the relative stakes for humans and animals in the future. The bad futures for humans considered by longtermists often involve premature extinction, or failure to reach full technological maturity. Futures like these are only 'bad' compared to the positive futures which we could have, but they are not intrinsically bad. Of course, there are possible futures in which human beings suffer significantly, for example if a sufficiently bad global totalitarian regime took over, or if a super-intelligent AI came to positively value human suffering, but these futures are relatively unlikely. The most likely futures open to animals on the other hand are intrinsically bad, in the sense that they involve large amounts of suffering for large numbers of animals over very long periods. Unlike humans, who have the intelligence and technological skill to solve the problems facing them, animals seem almost helpless against the natural forces that cause them to suffer. We might say that the default state of animals is to suffer. Since the most probable long-term futures are extremely bad for animals, we have additional reason to take measures today to try to change our trajectory to avoid those futures.

**4 Causal Efficacy**

In this section I will argue that *Causal Efficacy* applies to animals. By utilizing the longtermist concepts of persistent states and value lock-in, it is possible to identify some actions we can take today that we can reasonably expect to make the long-term future go better or worse for animals. First, I describe wild animal suffering (WAS) as a negative-value persistent state which, absent human intervention, we can expect to last for a very long time. I also describe some ways in which human agency could increase the scale and persistence of WAS. Second, I argue that there is a danger that humanity might lock in values which will make the long-term future much worse for animals by spreading and perpetuating animal exploitation, and by making future people less concerned about mitigating WAS. Furthermore, if it is the case that humanity will soon expand into space, then this is a very strategically important time at which to avoid locking in bad values.

*WAS as a negative-value persistent state*

On Earth trillions of animals suffer and die painfully each year as a result of purely natural processes, while a much smaller number seem to live good lives. It's likely that the vast majority of lives lived by wild animals today are bad (the Net Bad Lives claim), and possible that aggregate welfare across all lives on Earth is also negative (Net Suffering claim). Furthermore, this doesn't seem to be an accidental feature of life on our planet; rather it is caused by structural features of evolution by natural selection, and so WAS is likely to continue to dominate animal wellbeing for as long as life on Earth continues to be governed by those processes. In fact, since these natural harms are the predictable result of natural selection, it is reasonable to think that similar outcomes will obtain on any planet on which sentient life evolves.[93] It is difficult to quantify just how bad WAS on Earth is, but we can

---

[93] See paper 3 and O'Brien (2022b).

make some reasonable estimates. Extremely conservatively we can say that ten trillion animals die each year after enduring only 1 minute of net suffering. This would amount to approximately 18.2 million years of experienced suffering each year across all animals on earth. If this doesn't sound that bad, remember that the degree of suffering we are talking about is high – for example, being eaten alive or dying of starvation. This estimate is highly conservative in many ways. First, it takes the lowest estimate of wild animal populations. Second, it falsely assumes that the number of deaths per year is equal to the number of animals alive at any given time, but since the vast majority of animals are small short-lived ones who breed frequently and in vast numbers, the number of deaths per year must be much higher than this. Third, it assumes only one minute of net suffering. This may be true of animals such as small fish who are quickly eaten by predators immediately after hatching, but it must be remembered that many animals also starve, or die from disease, and even those who are killed by predators must sometimes endure more than a minute of suffering from being pursued and eaten more slowly. A reasonable higher estimate would be that there are a quadrillion deaths per year. If we assume again only one minute of net suffering each, that comes to 1.82 billion years of experienced suffering per year.

If we accept these claims, it's clear that WAS is a negative-value state. It is also a highly persistent one. Sentient animal life has existed on our planet for around half a billion years, surviving numerous mass extinction events, and life itself has survived for around four billion years. While individual animal species come and go, sentient animal life is much more robust. Given that sentient animal life has already endured so long it's highly likely that it will endure as long as Earth is habitable, another billion years or so. And since WAS isn't an accidental feature of life on our planet, but emerges predictably from necessary structural

features of natural selection,[94] it is highly likely that as long as sentient life endures, suffering will dominate wellbeing, unless humans intervene.

There are actions human beings can take in the short term that alter both the scale of WAS and its persistence. If, for example, humanity goes extinct, this could both increase the scale of WAS by freeing up more land for animals thus increasing their population, and its persistence by removing the only technologically capable animal on Earth that might one day be capable of significantly reducing WAS.[95] Somewhat surprisingly then, humanity's continued existence might be a good thing for wild animals.[96] It is also possible that humanity will increase the scale of WAS by spreading animal life beyond Earth. This could be done directly, by bringing wild animals with us to establish a biosphere when we colonize other planets. We could also do this indirectly, by launching directed panspermia missions aiming to spread microbial life throughout the galaxy in the hope that some of the microbes will survive and evolve on the target planets, eventually giving rise to complex biospheres with sentient animals. Such missions are possible using current-day or near-future technology, and they would be a relatively cheap way of spreading life throughout the galaxy.[97] It's also possible that we will spread life accidentally, for example by failing to decontaminate spacecraft originating from Earth, thus introducing terrestrial microbes to other planetary bodies where they might eventually evolve into more complex sentient life. It is very difficult to completely decontaminate spacecraft, and essentially impossible to eliminate the millions of microbes that live on, and in, human astronauts. Some scientists have already suggested that contamination of other planets is inevitable, and this should be seen not as regrettable but

---

[94] Paper 3, section 2.4.
[95] Though human extinction would also end our exploitation of animals and dramatically reduce the possibility of life spreading beyond Earth.
[96] Bruers (2019) also makes this point.
[97] See paper 3 in this thesis.

as a good thing, with microbes forming a first wave of colonization of other worlds (Lopez et al., 2019).

Such actions would not only increase the scale of WAS by multiplying it across other planets, but would also make it more persistent. Though life has survived on Earth for four billion years it is still possible that some global cataclysm might eliminate life on the planet. By spreading life to other worlds, we would be increasing the chances that sentient life, and hence WAS, will persist for a very long time. Furthermore, even if no extinction event occurs, our planet has a finite period of habitability. Other planets may have longer periods of habitability than our own. It is also possible that we could spread life to planets which orbit stars like our own, but at an earlier stage in their development. WAS on Earth is bad, but human action could result in WAS that is vastly greater in scale and much more persistent. Even without human extinction, or spreading animal life beyond Earth, there are more prosaic ways in which we can make WAS on Earth worse. For example, if we drive more large herbivores to extinction, they may be replaced by smaller animals, who, because of their reproductive strategy and greater vulnerability to predators, may have lower levels of average welfare.[98] Sebo (2023a) raises the possibility that climate change might result in a world with a higher proportion of small to large animals. If small animals have net negative lives, then this will be a change for the worse for animals. Furthermore, since evolution works slowly, we can expect that this 'new world' of small animals will persist for a long time.

The fact that humanity can make the long-term future worse for animals is sufficient to secure *Causal Efficacy*. It is also possible, however, for us to take actions that we can reasonably expect to make the long-term future go better for animals. We know the causes of WAS, and in principle it is possible for us to systematically intervene in the natural world to mitigate the suffering endured by animals, or even to enhance animals so that they are more

---

[98] See paper 1 for an argument that humanity may already be responsible for changes like this.

likely to live happy lives. We already intervene in nature on an ad hoc basis to help wild animals, albeit for anthropocentric reasons. The ultimate causes of WAS, though, are genetic – most animals have evolved to have huge numbers of offspring, the vast majority of whom will die soon after birth, and many other animals have evolved to prey upon or parasitize other animals. To eliminate WAS, then, would likely require genetic editing of wild animals. It would also require extensive understanding of the functioning of ecosystems so that our interventions don't have unintended bad consequences. The precise nature of the interventions required must be determined by scientists, and the technology and understanding required for successful interventions are likely to be decades away. We cannot know at this point whether this will be successful. It might be the case that the best we can hope to do is to mitigate WAS somewhat. On the other hand, it might be the case that future science and technology will enable us not only to greatly reduce WAS, but to ensure that most animals have good lives. When we remember just how bad WAS is, and just how long it could last, however, it should be clear that even a slim chance of success has tremendous expected value.

*Value lock-in*

Animals, whether wild or domesticated, are almost entirely at the mercy of human beings – whatever values we have will in large part determine their futures. The values held by most human beings today are disastrous for animals. Most humans are speciesists – they think it is acceptable to sacrifice the most important interests of animals in order to satisfy the most trivial interests of humans. This is most obvious when we consider the animals we exploit for food. If humanity maintains these anti-animal values this will likely have very bad consequences for animals, as we will continue to dominate and exploit them indefinitely.

Anything we can do to undermine speciesism, and to shift humanity to more animal-friendly values can be expected to improve the long-term future for animals.

Furthermore, we may be in an especially critical period regarding the impact on the future that our attitudes to animals could have. It's not unreasonable to think that humanity might begin colonizing space within the next century.[99] If we maintain those speciesist values which underlie our mass exploitation of animals, it is possible that we may take domesticated animals with us and multiply our exploitation of animals across other worlds. Scientists have already made proposals for establishing systems of animal agriculture off Earth, and as our spacefaring technology develops, more such proposals may be made.[100] If there is no shift in our values regarding other animals, it seems likely that we will continue to exploit them in space if it is technologically and economically feasible to do so. On the other hand, if we change our values regarding animals before we begin colonizing other worlds, then it is unlikely that we would take animals with us to exploit, given the technical challenges and costs involved. The possibility of human expansion into space brings a greater sense of urgency to abolishing our practices of animal exploitation. If we expand into space and colonize other planets before eliminating industrial animal exploitation, then colonists are likely to bring some domesticated animals with them. Once colonies are established on other worlds, cultural differentiation may be inevitable, so that even if we subsequently renounce speciesism on Earth, colonists on other worlds may retain speciesist attitudes.[101] If that happens, then further colonization originating from these worlds may spread animal exploitation throughout the galaxy, even if the population of Earth renounces animal exploitation.

---

[99] See MacAskill (p.93, 99, 189) on space settlement as a possible cause of value lock-in.
[100] See the International Space University (2020) report on lunar insect farming and Przybyla (2021) for a proposal to establish a system of fish farming on the moon.
[101] See Torres (2018) on the likelihood that colonization of space will result in radical differences between the home world and the colonies, due to evolutionary changes, deliberate genetic modifications to better adapt colonists to their new homes, and cultural differentiation.

There is an opportunity then to lock in anti-speciesist values here and now, thus significantly reducing the probability of spreading animal exploitation to other planets. If we abolish animal exploitation before we begin colonizing space this will reduce the possibility of spreading animal exploitation beyond Earth in two ways. First, all colonists leaving Earth would start out with anti-speciesist values, and all our colonies would have to revert to speciesist values independently. Secondly, colonies would have no stocks of domesticated animals as we wouldn't bring any with us from Earth, and so obtaining animals to exploit on colony worlds would be much more difficult. It's possible that humans will begin colonizing other worlds later this century. Given this possibility, and the strategic advantage of locking in anti-speciesist values before expansion begins, this may be the most important time to promote anti-speciesist values. After colonization it may be too late to prevent the spread of animal exploitation throughout the galaxy. The expected value of efforts to shift human values away from speciesism may be much higher now than they will be after space colonization has begun. We are at a crucial point in the history of humanity, and the values we settle on may determine the fates of countless animals.

**5 Longtermist Interventions for Animals**

In this section I will suggest some interventions that present-day or near-future humans can take in order to increase the likelihood of the long-term future going well for animals. These suggestions are tentative, however – much more empirical research is required to determine how effective they would be. Nevertheless, they are sufficient to demonstrate that longtermism may be a fruitful framework for thinking about animal wellbeing. Some of these interventions have the potential to be highly effective, perhaps more so than the best short-term interventions, and people concerned with reducing the harms suffered by animals should consider them. Furthermore, some of these interventions are scarcely even conceivable on a

more short-term focused framework. Though these suggestions for improving the long-term future for animals are highly speculative, the same is true of most longtermist suggestions for improving the human future.

*Reducing the likelihood of spreading WAS*

One of the easiest ways in which humanity could make the future worse for animals is by spreading wild animal life to other planets, where natural selection will almost inevitably result in mass suffering. As we saw in section 4, this could be done deliberately by direct colonization or by launching directed panspermia missions, or accidentally, by contaminating other planets with microbial life. To see how bad this could be, let's imagine that we spread life to just one other planet, and that sentient life will survive there for a billion years. Taking my most conservative estimate for the net amount of WAS on Earth each year and multiplying that by a billion years we get 18.2 quadrillion years of animal suffering for one planet. If there is some action we can take that reduces the chances of spreading wild animal life to one other planet by just 0.1% that action has an expected value equivalent to preventing 18.2 trillion years of suffering. I have calculated elsewhere that the annual suffering of animals we exploit for food comes to about 16 billion years every year.[102] Reducing the probability of spreading WAS to one other planet by 0.1% has an expected value equivalent to stopping all suffering caused by exploiting animals for food on Earth for over a thousand years. One could of course question my calculations here. However, it should be borne in mind that I have been extremely conservative in my estimates of the amount of WAS experienced annually on Earth, and in my presumption that whatever action we take to prevent the spread of WAS would only prevent its spread to one planet.

---

[102] unpublished manuscript.

Taking the figures I have given at face value, it is at least plausible that such interventions might be more effective than most short-term ones. Of course, we won't know how effective they might be until we know exactly what form they will take, how much they will cost, and how the cost/benefit ratio compares to short-term interventions on behalf of animals, all of which must be the subject of empirical work. Tentatively I suggest the following:

(1) strict legal controls over space technology, enforced by a suitable international body which prohibits spreading nonhuman life to other worlds, whether directly or via indirect methods such as directed panspermia, and which enforces;

(2) strict decontamination procedures for spacecraft originating from Earth, whether they are scientific, exploratory, or industrial, thus lowering the risk of accidental contamination.

*Reducing WAS on Earth*

It may be possible for humanity to develop the science and technology required to safely intervene in nature to remove the causes of WAS and thereby reduce the suffering of wild animals, perhaps to the point at which most animals live net positive lives. If such interventions were successful and long lasting, they could make the long-term future far better for animals. It's difficult to say exactly how good such a future might be, but I will attempt a very rough estimate. Again, taking my conservative estimate of the amount of net WAS on Earth (18.2 million years per year) we can imagine that a successful intervention programme would eliminate this suffering, and replace it with the same amount of mild positive wellbeing, giving us a net of 18.2 million years of mild wellbeing distributed across all wild animals on Earth each year. If we imagine this state enduring for the remainder of

Earth's period of habitability, this comes to 18.2 quadrillion years of wild animal wellbeing. Even a tiny chance of success would have a very high expected value in terms of reducing suffering and promoting positive wellbeing for vast numbers of animals.

Reducing WAS on Earth is likely to be much more difficult and costly than preventing the spread of WAS to other planets, and at this point it is not possible to say how cost-effective such interventions would be compared to other long- or short-term interventions on behalf of animals. There are some considerations which do favour an interventionist programme, however. First, though likely to be expensive, most of the costs will be at the research and development stage – the interventions themselves are likely to be automated processes rather than a continual hands-on process of 'policing' nature. For example, gene drives can be used to spread desirable traits through the animal population, and once developed and released into the target population, the rest of the process is automatic (Johannsen, 2021, p.40-41). If we consider the costs of research and development to be fixed costs which will continue to benefit wild animals indefinitely, then such research may in fact prove highly cost-effective. Secondly, at least some interventions in nature are highly persistent in the face of natural changes, so they would remain effective for a very long time even if human beings were no longer around to maintain them (Liedholm, 2019, p.10-11). Finally, if humanity ever does manage to intervene in nature so effectively that most animals in fact have good lives, there is no reason in principle why we couldn't then utilize this knowledge and technology to spread happy animal life to other planets, thus multiplying the positive welfare effects across many worlds. I suggest:

(3) investing in research and the development of the technology required to safely help wild animals on a large scale;

(4) supporting efforts to reduce existential risk to humanity, since without human beings there is little prospect of WAS being eliminated.

*Shifting to more animal-friendly values*

It is likely that speciesism underlies our willingness to exploit nonhuman animals. It is unclear to what extent speciesism can be combated, and what the best methods of doing so are. The rise of veganism gives grounds for optimism though. Furthermore, research has shown that speciesism is related to other prejudices like racism and sexism; if the latter have declined over time then it might be possible to reduce the former too (Caviola et al., 2019). Finally, several studies have shown that speciesist attitudes emerge between the ages of 11 and 18, and that they are socially acquired, and thus it seems that it ought to be possible to educate children differently so that they adopt more animal-friendly values instead.[103] Therefore I suggest:

(5) promoting anti-speciesism and/or moral circle expansion (Anthis and Paez, 2021).

Finally, it is possible that speciesism also underlies humanity's general lack of concern for WAS and resistance to the idea of intervening in nature on a large scale to reduce it, though I am not aware of any research on this question. In any case, I suspect that speciesism is only part of the explanation for our indifference to WAS. It is also likely that most people are simply ignorant of the scale and severity of WAS. I suggest:

---

[103] Wilks et al. (2021), McGuire et al. (2022).

(6) education on WAS, and attempting to move the general public away from valuing nature in a purely holistic fashion and towards an appreciation of the value of individual sentient animals.

I do not pretend to have shown that these interventions would be effective - much more research is required to determine their feasibility and cost-effectiveness. However, they may be a promising place to begin research into longtermist interventions on behalf of animals. Furthermore, I hope that these suggestions are plausible enough that animal advocates who aren't longtermists may at least consider the utility of adopting a longtermist framework when thinking about how to promote animal welfare, both because they have the potential to be effective, and because some of the proposed interventions are scarcely even conceivable within a short-termist framework.

## 6 Conclusion

I have argued that the longtermists' neglect of animals is a serious mistake. It may also be a mistake for animal ethicists to ignore the long-term future. The basic argument for longtermism applies at least as well to animals as it does to humans, and the longtermist framework may be a useful way of identifying highly effective interventions to improve the wellbeing of animals into the long-term future. Much more research is required to determine which interventions would be most effective, and how they compare with short-term or human-focused ones. Finally, there are many important questions relating to the long-term future of animals that I have not addressed here. For example, there may be important non-hedonic values relating to animals that we might want to promote in the long term, such as the aesthetic value of biodiversity or of having a wide variety of forms of sentient happy animal life. We may also have some duties to make the long-term future better for animals

which are grounded in the past. As Ord suggested, we might want to try to 'make amends' for our past mistreatment of animals, though by considering individual animals rather than species, and with an awareness of the scale of WAS. Or, given the hundreds of millions of years of animal suffering which were necessary for the evolution of our species, we might think that we have reason to make the future better for animals, so that, in Parfit's words, "the existence of the Universe will have been on the whole good".[104] The history of sentient life may be only just beginning, and we ought to use our influence to ensure it goes as well as possible.

---

[104] Parfit (2011) and Kahane (2023).

# How Many Animals Should There Be? Population Ethics, Species, and the Repugnant Conclusion

**Abstract:** Totalism is the theory in population ethics which states that one population is better than another if and only if it has higher total welfare. One counterintuitive consequence of this theory is the Repugnant Conclusion (RC). In this paper I will argue that Totalism also naturally leads to the Animal Repugnant Conclusion (ARC). This may pose a stronger axiological challenge to Totalism than the RC. Worse, the strategies that have been used to avoid the troubling normative implications of the RC don't work in the animal case, so we may have reason to bring about an animal Z population. I introduce the notion of 'Efficiency of Welfare Production' (EWP). This is the idea that animals of different species vary in the degree of efficiency with which they convert resources into welfare. If we want to maximize total welfare, without any speciesist bias, we should identify which species has the highest EWP and try to maximize the population of that species. This leads to very counterintuitive implications, whether we endorse simple hedonism or a more sophisticated theory of welfare. If these implications are unacceptable, then Totalism may not be an adequate theory of population ethics when nonhuman animals are included in the populations to be evaluated.

## 1 Introduction

Since the appearance of Parfit's *Reasons and Persons,* a huge literature on population ethics has emerged. A key aim of work in this area has been to avoid the Repugnant Conclusion (RC). The RC is a natural implication of *Totalism*, the view that one population is better than another if and only if it has a higher total welfare (Greaves 2017, p.3). Though it can be reached without assuming Totalism, the RC is generally seen as a challenge to that theory. Several alternatives to Totalism have been developed, often with the explicit aim of avoiding the RC. These include *Averagism*, *Critical Level* views, and *Variable Value* views. Though

these alternatives may avoid, or at least moderate, the RC, each has problems of its own which may be as difficult to accept as the RC. It has been argued that no theory in population ethics will avoid all counterintuitive results (Arrhenius, 2000; Greaves, 2017). Perhaps because of the apparent impossibility of formulating a population axiology that satisfies all our intuitions, some population ethicists have argued that we should endorse the RC (Tännsjö, 2002; Huemer, 2008). Others have agreed that avoiding the RC should no longer be seen as a criterion for an acceptable population axiology (Zuber et al., 2021). General strategies adopted by those who advocate endorsing the RC include debunking the intuitions that lead us to view accepting the RC as repugnant, pointing out that *all* population axiologies have some deeply counterintuitive implications, and blocking the implication that accepting the RC entails a commitment to trying to bring about such a state of affairs.

Despite the voluminous literature on the problem, a notable omission is the lack of attention paid to nonhuman animals. This may simply be an oversight, as most ethicists are much more concerned with humans than with animals. Alternatively, population ethicists may have assumed that including animals will not make a significant difference to their theories. In this paper I will explore how the inclusion of animals bears on population ethics, paying particular attention to the RC. I will argue that, on some plausible assumptions about welfare, Totalism not only leads to the RC but to the Animal Repugnant Conclusion (ARC), that is, the conclusion that, for any large population of human beings living excellent lives, there is some much larger imaginable population of very simple animals living barely worth living lives that is better. This idea is relatively familiar – Parfit (2016, 2017) recognized that the RC could be extended to include animals, and Sebo (2023a) and Williamson (2021) also make reference to the ARC. Recognizing that accepting the RC entails acceptance of the ARC too may put some intuitive pressure on those who argue that we ought to accept the RC in the human case. If accepting the ARC is more counterintuitive than accepting the RC then

this puts some pressure on Totalism, and any other view that entails the RC. We can call this the *Axiological Problem* of the ARC. I will argue that the ARC also raises a *Normative Problem*, that we may have reasons to try to bring about something like the population in the ARC, and this may be more challenging than the Axiological Problem. I will argue that there are differences between humans and animals which make avoidance of the Normative Problem much more difficult in the animal case than in the human case. If I am right, then this poses a separate, perhaps more serious, challenge to Totalism.

In section 2 I will present the RC and describe how it is entailed by Totalism and why it has been seen as an axiological challenge. I will then outline the arguments in favour of accepting the RC. I will describe the Normative Problem and outline the ways in which it can be avoided in the human case. In section 3 I will show how the RC can be extended across species lines to generate the ARC. I will then describe the Normative Problem posed by the ARC and argue that the strategies used to avoid the Normative Problem in the human case don't work in the animal case, and so the ARC poses a distinctive problem for Totalism. In section 4 I will elaborate on the animal Z population. I will describe what such a world would look like and argue that it is much more difficult to accept than an equivalently realistic human Z world. I will introduce the notion of *Efficiency of Welfare Production* and argue that Totalists ought to favour a population composed as much as possible of individuals of whatever species is most efficient at turning resources into welfare. This is likely to be a very small, relatively simple animal. In section 5 I will consider the objection that Totalism only entails an animal Z world if we accept an overly simplistic conception of welfare. I will argue that if we accept a more complex conception of welfare, we end up with a world almost exclusively populated by human beings, as they would be the most efficient species. While this world would be appealing in some ways, it would be lacking in biodiversity and wild animals generally. Both worlds, however, would be severely lacking in species-diversity and

in wilderness – these would be sacrificed to the maximization of welfare. On either conception of welfare, then, it seems that it is difficult for Totalists to endorse a mixed species population.

Finding a population ethical theory that can include nonhuman animals in a credible way is important for population ethicists, as no theory of population ethics is complete if it cannot include animals in the populations to be assessed. It is also important for WAS interventionists. Large-scale interventions in nature will inevitably alter the size of animal populations, their total and average welfare levels, and the overall species composition of the biosphere. In order to assess proposed interventions, interventionists need a theory of population ethics. If the arguments in this paper are successful, then Totalism may not be an adequate theory.


## 2 Totalism, the Repugnant Conclusion, and the Normative Problem

### 2.1 The Repugnant Conclusion

The RC has been seen as one of the key challenges in population ethics, and especially as a challenge to Totalism (Greaves, p.3). Parfit's original formulation of the RC was that, for any possible population of at least 10 billion people living excellent lives, there is some much larger imaginable population that is better, although everyone in it has a life that is only barely worth living (1984, p.388). There are at least two ways to get to the RC. First, on a Totalist view it is a simple matter of 'adding up' all the positive welfare in both populations. The latter population will have a higher total of positive welfare despite the low levels of welfare enjoyed by each person simply because there are so many people. In Parfit's own analogy, a large enough heap of milk bottles, each containing only a single drop, will have more milk than another smaller group of bottles each of which is full (2016, p.110; 2017, p.153). On the Totalist view what matters is the total balance of positive over negative

welfare; each unit of welfare is worth as much as any other, and it doesn't matter how that welfare is distributed, or how little is enjoyed by each person.

A second route to the RC, also identified by Parfit, is by 'mere addition' (1984, p.419-441). Imagine a population A of people living excellent lives. Population A+ consists of the same group of people at the same level of welfare, but it also includes a number of additional people living good lives, though not as good as the lives of the original population of A. The mere addition of people living good lives cannot, it seems, make A+ worse than A, and it is plausible that it makes it better. In population B we equalize the welfare of all the individual members of the population, so that the original people from A are at a slightly lower level, while the new members are at a higher level than they were in A+, also adding a slight welfare boost so that the total and average welfare in B is higher than that in A+, despite the fact that the original members of A have a slightly lower level of welfare. It is hard to deny that B is an improvement on A+. It does better on total and average welfare, and it is more equal than the inegalitarian A+. If B is better than A+, which is as good as or better than A, then, given the transitivity of 'better than', B must also be better than A. This procedure can be repeated indefinitely until we arrive at some vast population of people with lives that are just barely worth living, Z, which is better than A.

The RC is seen as a challenge as it is very difficult to accept that the existence of a Z world in which vast numbers of people live lives that are barely worth living could possibly be better than the A world in which 10 billion people live excellent lives. Parfit's original understanding of Z was that it is a drab world in which individuals have few pains which are barely outweighed by the simple pleasures of 'muzak and potatoes', and from which all the greater accomplishments and ecstasies of excellent human lives are absent (1986, p.148). On this understanding of Z, it is indeed very counterintuitive that it could be better than A, no matter how many people Z contains. Meeting the challenge has proven extraordinarily

difficult, however. The premises in the argument for the RC are plausible, and denying them, or aggregating welfare in different ways, often leads to even more counterintuitive conclusions. Even if one is not a utilitarian, it is hard to deny that positive welfare is a valuable thing, and that, all else being equal, a population with higher total welfare is better than one with lower welfare. The steps in the mere addition argument are especially difficult to deny. Valuing average rather than total welfare leads to the sadistic conclusion, and other ways of aggregating welfare lead to similarly unpalatable conclusions. Though the RC is counterintuitive, it seems that no theory in population ethics will give us all the results that we might intuitively wish for, and the RC may even be the least bad option.

*2.2 The Normative Problem*

Some philosophers have endorsed the RC (Tännsjö, 2000; Huemer, 2008) and others have agreed that avoiding it should not be a necessary condition for an acceptable population axiology (Zuber et al, 2021). Both Tännsjö and Huemer agree that the RC is supported by strong arguments and that the opposition to endorsing it is based on appeals to intuitions which may be unreliable. For example, both agree that we may undervalue the Z population, as correctly evaluating it depends on our ability to intuitively grasp very large numbers, something that humans do poorly. Secondly, they claim that part of the reason we tend to prefer the A population is out of egoistic bias – we may be illegitimately imagining ourselves living in either the A or the Z world and evaluating the A world as superior on the grounds that this is the world in which we would prefer to live.

One strategy for those who advocate endorsing or tolerating the RC is to debunk our axiological intuitions that the A world simply must be better than the Z world, no matter how many people the latter may contain. If one endorses the RC and accepts that Z is more valuable than A then a new problem arises, however. The problem is that if we acknowledge

that Z is more valuable than A, and we accept that we have at least a prima facie moral reason to bring about more valuable states of affairs when it is in our power to do so, then it seems that we have a prima facie reason to try to bring about a Z world, or at least something approximating it should that be practically possible. Even if we are willing to endorse the RC in our axiological thinking, many of us would still find it difficult to believe that we have any moral reason to try to bring about such a state of affairs. If we initially accept the RC in our purely axiological thinking, then go on to reject the implication that it gives us any reason to bring about Z-like populations, we might see this as a reductio of the axiological claim. We can call this *the Normative Problem* of the RC to distinguish it from the axiological problem.

Why, however, should we think that the Normative Problem is a separate problem beyond the axiological one? After all, if one genuinely believes that the Z world is more valuable than the A world, then it shouldn't be difficult to accept that we have reasons to try to bring about such a world should that be possible. There are several reasons why the Normative Problem may be important. First, a dialectical reason. Since all theories in population ethics have counterintuitive axiological implications, in deciding between competing theories we may look to features other than their axiological implications to make our decision. If it turns out that Totalism has both counterintuitive axiological implications and counterintuitive normative ones, while some other theories have only the counterintuitive axiological implications, this may be a reason to disfavour Totalism. Note that I am not here claiming that this is in fact the case. It may turn out that all theories in population ethics have both normative and axiological implications that are difficult to accept, and in that case, we would be back to a dialectical deadlock. I only make the case here that, once we include animals, Totalism does in fact have implausible normative implications. Second, if we accept the methodology of reflective equilibrium, then the fact that a given theory has counterintuitive results at more than one level (here, axiological and normative) counts

against that theory, if there are other theories that have counterintuitive implications at only one level. This is because all our intuitions are taken to be evidential, not only the more foundational (axiological) ones. Finally, at least for some people, normative intuitions about what we have reason to do are stronger than highly abstract intuitions about what we ought to value in hypothetical (and perhaps practically impossible) cases. If we give more weight to normative intuitions about what we ought to do in possible cases than we do to our axiological intuitions about impossible cases, then a theory that has implausible normative implications is, all else being equal, weaker than one that does not.

*2.3 Avoiding the Normative Problem*

Though they don't use the term, both Tännsjö and Huemer seem to recognize the Normative Problem, and gesture at strategies to make it less troubling. Huemer explicitly asks whether, if we accept the RC as he urges us to, we ought to "aim at a drab future like world Z, where each of our descendants occupies a single, cramped room, and there is just enough gruel to keep them from hunger?" (p.928). He argues that, given any plausible view about the actual effects of population growth on total and average wellbeing, Totalism will not recommend trying to maximize the number of people with lives barely worth living. Rather the best way to maximize total welfare will be to reach the optimum population, which is likely to be a point at which there is a large population at a still relatively high, though not maximum, average level of wellbeing (p.929). For example, if we doubled the world's population, this might result in the average wellbeing level being much lower than 50% of current levels, thereby reducing total wellbeing too. Totalism therefore would not imply that we ought to maximise the population as that would actually reduce total welfare relative to the optimum population size.

Furthermore, Huemer points out that it isn't clear that increasing the population will necessarily reduce average wellbeing at all. While larger populations mean less space, and a lower share of available resources for each person, they also have positive effects on average welfare. More people could mean more innovation, more art, a greater pool of talent to draw from and so on. If these positive effects of large populations on average welfare outweigh the negative effects, then the RC is effectively rendered practically toothless. The apparent trade-off between total and average welfare would turn out to be an illusion, and though the axiological problem posed by the RC would remain, in practice it would be side-stepped by the positive welfare effects of a large population. Having made this response to the Normative Problem Huemer concludes that Totalism "does not enjoin us, in reality, to pursue the world of cramped apartments and daily gruel" and suggests that critics of Totalism "will therefore look upon the principle with less revulsion than has hitherto been customary" (p.930). This final remark suggests that Huemer believes that at least part of the reason that people find endorsing the RC repugnant is that it seems to have such unpalatable normative implications. We can label Huemer's response to the Normative Problem the *Unpredictable Effects Argument*, as it questions the assumption that we know what the effects of population increase on welfare will be.

Tännsjö wishes to correct the mistaken belief that accepting the RC means that one should also endorse all actual increases in the world's population, as some such increases may actually reduce total welfare. He also reminds us that the RC describes a merely hypothetical situation, and that the population in Z is, in Parfit's own words, an *imaginable* one (Tännsjö, p.342). This suggests a further way of blocking the Normative Problem which we can label the *Impossible Populations Argument.* The Z population is merely imaginable because of its magnitude. Of course, there is no fact of the matter about how many people are in the Z population, but we can make reasonable estimates. If the A population has ten billion

153

people, as Parfit stipulates, and the people in the Z population have lifetime wellbeing levels 100 times lower than those in the A population, then for the total wellbeing in Z to be greater than that in A the Z population would have to be more than 100 times larger. That is more than one trillion people, and it is not at all clear that the planet can sustain that many human beings.[105] The implication is that, since the massive Z population is entirely hypothetical, we needn't worry about the normative implications of endorsing the RC.

Another possible response to the Normative Problem is what I will call *the Control Problem*. If we endorsed the RC and thought that we had reason to try to bring about a very large population of human beings with lives only barely-worth-living, we would have to exert very significant control over human reproduction indefinitely in order to create and maintain it. We could hardly expect a Z-like population to come about through chance, since most people don't reproduce as much as possible, or up to the point at which the lives of their children approach the neutral level. In fact, as economies develop, it seems that people prefer to have smaller families and to invest a great deal of resources into their children rather than spreading their resources more thinly across a greater number of children.[106] To bring about something like a Z population then governments would have to exert a huge amount of control over the reproductive lives of their citizens. This would of course be almost impossible to implement successfully. Furthermore, if states adopted coercive measures to make people have more children this would likely have severe negative effects on the wellbeing of the population, possibly sufficient to counter the positive welfare effects of creating extra people. Finally, even if it were practically possible to successfully implement such extensive control over people's reproductive decisions, many people recognize

---

[105] The UNEP (United Nations Environmental Programme) has produced a survey of 65 estimates of the planet's maximum carrying capacity. Only one of the estimates is as high as 1 trillion people, and the majority of estimates are between 8 and 16 billion people (UNEP 2012).

[106] This negative correlation between fertility and economic development is "one of the most solidly established and generally accepted empirical regularities in the social sciences" (Myrsklä et al 2009).

deontological constraints against such infringements of people's autonomy. If this is correct, then it is neither practically possible, nor morally permissible, to take the measures required to create a Z-like population. Even if we accept that Z is more valuable than A, if it isn't practically possible or morally permissible to move us towards Z then we can't have a moral reason to do so. The RC then would be an awkward, counterintuitive fact about what we ought to value, but it wouldn't have implications for how we ought to act. Endorsing the RC need not mean endorsing 'repugnant actions'.

Finally, even if we were able to bring about a Z world, and if it did turn out that the increased population would have the expected negative effects on average welfare, there is no reason to expect that Parfit's Drab Z would be the result. The barely-worth-living lives in Z might be more like the Roller-Coaster Z, in which the ecstasies in each life just barely outweigh the agonies.[107] An 'exciting Z' might well have most or all of the things present in truly excellent lives, with the repugnance coming from the fact that each life is on balance only barely worth living. However, a world like this one might seem less objectionable than the drab world envisioned by Parfit. Recall that Parfit regretted the loss of the best things in life, like the music of Mozart. But why should we expect population increase to extirpate Mozart's music from the world, or to prevent people from appreciating it? If anything, we should expect the opposite. As the population expands some resources and experiences will certainly become scarcer. In a world of a trillion people, it seems unlikely that more than a minute fraction of them will get to visit Venice, if indeed Venice still exists in its current form and hasn't been turned into a high-density living space. However, goods like Mozart's music, great novels, films, etc. are all easily and cheaply replicated digitally, and there seems no reason why there wouldn't be even more people enjoying these great goods in the Z world

---

[107] See Parfit (2016, p.118) where he introduces Short-Lived Z and Roller-Coaster Z as alternatives to Drab Z. Though he thinks that Short-Lived Z would make the RC significantly less repugnant, he doesn't comment on the value of Roller-Coaster Z.

than in the A world. If a highly populated Z-like world still allowed for the enjoyment of these great goods, rather than mere muzak and potatoes, it seems much less objectionable. If the above considerations are correct then the RC may still pose a theoretical problem in axiology, but it doesn't have a practical bite. Endorsing the RC in theory doesn't entail that we have moral reasons to bring about a Z-like world with a huge human population living lives that are barely worth living, nor need we worry about the loss of all the best things in life.

## 3 The Animal Repugnant Conclusion and the Normative Problem

### 3.1 The ARC

Though his original formulation of the RC concerned only human beings, Parfit recognized that the same logic that led us from excellent lives to muzak and potatoes could be extended across the species barrier, resulting in Z worlds composed of nonhuman animals. His own example was a vast population of lizards basking in the sun, enjoying lives that are barely worth living (2016, p.118). A few others have written on the possibility of extending the RC to include nonhuman animals. Tännsjö also recognizes that the RC really ought to be stated in terms of sentient beings rather than people and claims that it might be speciesist to reject a world with higher total welfare on the basis that the population is composed of nonhumans (2002, p.339, 352). Williamson (2021) uses the ARC as an argument against Critical Level axiologies. Critical Level views are those according to which lives worth living can only contribute to the value of a population if they surpass the critical level. Williamson argues that such views cannot accommodate animals, as they run into insurmountable problems whether they make the threshold universal or they relativise it to species. Sebo (2023a) argues that expanding our moral circle to include animals raises questions about priority setting and what sort of population we have reason to aim at. Assuming that small animals

typically have lower average welfare levels than larger ones, and that large populations of small animals might have higher total welfare than a small population of large animals, Sebo argues that, at least in the long run, we may have reasons to try to bring about a population that has more, but smaller, animals rather than one composed mostly of larger animals. He calls this 'the rebugnant conclusion' as on his view we should also include insects and perhaps even microbes in our moral circle (p.11).

If we reject speciesism, accept interspecies cardinal comparisons of welfare, and endorse a Totalist welfarist axiology it is easy to see how we can get to the ARC. If we start with a population A consisting of 1 million happy Homo sapiens, we can see that population B consisting of 2 million Neanderthals who are each more than half as happy as the humans in A is better, as total welfare is higher. From B we can move to C, a population of 4 million Homo erectus, each of whom is more than half as happy as each of the Neanderthals in B. This operation can be repeated indefinitely, until we reach a vast population of simple animals with a very low capacity for happiness, each of which has a life barely worth living. Perhaps this population would consist of trillions of barely-sentient oysters, each of whom has a life that is just barely worth living. Let's call this population animal Z. If the happiness in each oyster life in Z is a million times less than the happiness in each human life in A, then, as long as there are more than a trillion (1 million million) oysters in Z, the total happiness in Z will exceed that in A.

As Williamson (2021, p.410) points out, for those who found the original RC repugnant, the ARC is likely to be an unwelcome result. If it is repugnant to conclude that a population of humans living excellent lives is less good than a much larger population of humans with lives that are barely worth living, it must be at least as repugnant to conclude that a sufficiently large population of simple animals would also be better. It may be that most people have the intuition that the ARC would be even worse than the RC, yet Totalism

implies that it is as good or better, and this may reduce the plausibility of Totalism. I am not convinced, however, that the ARC must be more troubling to Totalists than the standard RC. Like the standard RC, the ARC is a theoretical result in axiology. It may be surprising and intuitively unappealing, but, like the RC, the arguments that lead to the ARC are strong, and the primary objections to accepting it are based on intuitions that are vulnerable to debunking.

When we are considering populations that are merely logically possible, it is not clear that the ARC need be any worse than the RC. In thought experiments we can make whatever stipulations we like about the welfare levels of the inhabitants of Z, whether human or animal. For example, we can stipulate that the average level of welfare is arbitrarily low, that each life lasts only a few moments as in Short-Lived Z, that there are no higher pleasures, but only 'muzak and potatoes' as in Drab Z, etc. In thought experiments the differences between a human Z population and an animal one can be stripped away, until the only salient difference them is species. If we reject the claim that the value of a unit of welfare depends on the species of the individual in which it is instantiated, however, then there is no reason to prefer a human Z population over an animal one. Furthermore, accepting the ARC does not automatically entail that we have moral reason to bring about something like an animal Z population. One might think that the responses to the Normative Problem of the RC can be applied equally to the ARC, thus blocking the move from a counterintuitive result in axiology to an even more counterintuitive account of our reasons for action. In the next section I will argue that this supposition is false. There are relevant differences between humans and animals which mean that the strategies for responding to the Normative Problem in the human case do not apply in the animal case. This means that we do in fact have moral reason to try to bring about such a population. Then, in section 4, I will flesh out what a practically

possible animal Z population might look like, and highlight the undesirable features of such a world.

*3.2 The ARC and the Normative Problem*

In section 3.1 we saw how Totalism implies the ARC. Like the RC, it doesn't necessarily entail any commitment to 'repugnant actions'. We saw in 2.3 some ways in which those who endorse the RC in their axiology can avoid the Normative Problem. If these same ways of blocking the move from the axiological to the normative work in the animal case, then endorsing the ARC also needn't imply commitment to repugnant actions.

I will argue in this section that these strategies for avoiding the Normative Problem posed by the RC don't work for the ARC. Consider first Huemer's Unpredictable Effects Argument. His claim was that endorsing the RC axiologically doesn't necessarily entail that we should maximize the human population, rather we should seek the optimum population, which may be one with a relatively high average welfare. When we include animals in our considerations, however, things are different. First, as I will argue in greater detail in section 4, since small animals have much lower resource requirements than humans, we could support vastly higher populations of animals, and this could significantly increase total welfare. Second, it should be clear that most animals have lower welfare levels than most humans. It is plausible that most human beings have very good lives compared to small, simple animals like mice. Humans have a wide, diverse, and complex array of pleasures, deep emotional lives, and sophisticated intellectual interests, while mice presumably have comparatively few sources of pleasure, and relatively shallow emotional and intellectual lives. Furthermore, the average human lives for around 80 years, while the average mouse lives for about two years. The lifetime wellbeing of a single lucky mouse may be hundreds or thousands of times lower than that of a lucky human being. Bringing about a Z population of

mice then would certainly lower average wellbeing – the trade-off between average and total may or may not be illusory in the human case, but in the pan-species case the trade-off is all too real.

Secondly, consider Tännsjö's observation about the impossibly large populations involved in the human Z world. It may be correct that a human Z population is simply too large to sustain on Earth and must remain therefore purely imaginary. This is insufficient to block the Normative Problem, however. Even in the human case, the fact that Z is impossible to actualize does not show that we have no moral reasons to increase the population as much as we can in order to increase total welfare. While this would be unlikely to result in a world as unpalatable as Parfit's Z, if the expected negative effects on average welfare obtained then it might still be an intuitively undesirable world. More fundamentally, this 'impossibility objection' simply doesn't hold in the animal case. Consider the number of animals humans farm globally each year. According to estimates by Fish count (2019a and 2019b), in 2017 between 250 and 600 billion crustaceans were farmed, as well as 51-167 billion fish. We also slaughter an estimated 50 billion chickens each year, and several billion sheep, pigs, and cows (Thornton, 2019). These numbers are meant to illustrate the practical possibility of creating and maintaining very large populations of nonhuman animals. If we focused on creating smaller animals (e.g. mice instead of chickens and cows, oysters instead of salmon and cod) it seems likely that we could support even larger populations, easily rivalling those of the hypothetical human Z population. The impossibility objection may have some force in the human case, but not in the animal case.

Finally, consider the Control Problem. While it would be very difficult or impossible to control human reproductive behaviour sufficiently to bring about something resembling a Z population, it seems relatively easy to control the reproductive behaviour of some animals. It would be relatively simple for example to induce captive mice to reproduce in very large

numbers. In fact, we already do breed large numbers of mice to use in research labs.[108]

Questions of the permissibility of such control seem less salient in the animal case too. While it would be monstrous to interfere significantly with the reproductive decisions of humans, this seems far less problematic in the case of simple animals like mice and oysters. Humans make autonomous decisions about their reproductive lives based on many considerations about what they value and how they want to live their lives. Mice don't seem to make any reproductive decisions at all – rather their reproductive behaviour seems instinctive, and so there would be no issue of undermining the autonomy of mice by controlling their reproductive behaviours. This is even more obviously true in the case of creatures like oysters who don't seem to make any decisions about their behaviour – breeding them in large quantities seems no more objectionable than growing large amounts of crops, so long as no pain is inflicted on them in the process.

Sebo (2023a, p.9-12) argues that, if we can maximize total utility by creating many small animals with low average levels of welfare, then we may have reason to do so, at least in theory. He recognises that this is a counterintuitive conclusion, which many might want to avoid. In my terms, Sebo recognizes the Normative Problem, though he does not think it is necessarily a conclusion that we should try to avoid (p.12). Nevertheless, he offers two considerations which might allow us to block the practical normative implications of extending the RC to animals, even if we still accept the axiological claim that the animal Z world might be better than our own world. We can consider these as solutions to the Normative Problem that are specific to the ARC. The first solution concerns the tractability of creating a large population of very small animals and ensuring that they have net good lives. Sebo argues that this may be much more difficult than creating smaller populations of happier large animals. Though he thinks it would be tractable to create more insects and other small

---

[108] Carbone (2021) estimates that over 100 million are used each year in the United States alone.

animals, it might be much more difficult to ensure that they have net positive experiences, especially if they continue to reproduce through r-selection.

This suggests that Sebo has in mind the creation of large numbers of *wild* animals; r-selection is only a welfare problem for wild animals, as they quickly reproduce beyond the carrying capacity of their environments. If small r-selecting animals were bred in captivity, then it would be relatively easy to provide them with sufficient resources, and to control their breeding so that they don't produce more offspring than can survive. Sebo might also think that, even if we set aside the welfare problems caused by r-selection, there may be a difficulty in ensuring that insects are having good experiences, since they are so different from us. This is an epistemological problem – as we move further away from humans on the tree of life, our confidence in our ability to measure and promote welfare in other animals might reasonably waver. Both of these problems can be solved if we raise small mammals in captivity. This would solve the welfare problems caused by r-selection, as we could effectively control the reproduction of captive animals to ensure their population growth doesn't outstrip available resources. It would solve the epistemological problem (or mitigate it very significantly) as we can be much more confident in our assessment of the welfare of animals that are evolutionarily closer to us than we can be with insects. I will investigate this possibility in greater detail in section 4.

Secondly, Sebo argues that when we consider the long-term effects of our actions it may no longer be optimal to create large populations of small animals in the short term. Rather, we might best promote welfare in the long term by continuing to focus our efforts on improving the lives of human beings as, in the long run, humanity will be in a better position to maximize welfare than we are today. This suggests that, even if prioritizing the welfare of smaller animals is the best way to maximize welfare in the present, we are permitted to prioritize humans for now as this will have the best consequences in the long term. As Sebo

acknowledges, however, this isn't really a solution to the Normative Problem. Considering the optimal strategy for maximizing welfare in the long term "might save us from a surprising neartermist implication, [that we have reason today to create large populations of happy, small animals] but it would do so only by replacing it with a surprising longtermist implication" (p.13). Furthermore, it is not clear that even in the short term the best way to maximize total wellbeing will be to focus on human beings. Some ways of helping humans today will increase our future ability to help small animals, but others will not. For example, saving the lives of current day humans, or lifting them out of poverty might instead lower total wellbeing, if those humans will consume animal products from factory farms (Plant, 2022). Furthermore, as Beckstead (2012) has argued, not all human beings will be able to have positive effects on the long-term trajectory of humanity, and, in the case of such people, the Totalist might prefer to spend resources on creating large numbers of happy animals in the present.

The upshot of these arguments is that while the RC may be merely a theoretical problem in axiology, the ARC also poses a Normative Problem. Totalists could safely ignore the RC when deciding how to act since it is impossible to actualize something like the human Z population. Blocking the natural move from the axiological to the normative makes endorsing the RC, and thus holding on to Totalism, less unappealing. But since it is possible to bring about something like an animal Z population, however, and since the ARC also arises naturally from Totalism, this may count against the plausibility of Totalism.


## 4 Efficiency of Welfare Production and Mouse World

In section 3 we saw how Totalism leads not only to the RC, but to the potentially more troubling ARC. We also saw that the strategies to block the Normative Problem posed by the RC don't work in the animal case. This suggests that we would have reason to try to bring

about something like an animal Z population if that would maximize total welfare. So far, this is still hypothetical – I have given no reason to think that a large population of small animals would in fact maximize total welfare, nor have I said anything about what such a world would actually be like. In this section I will argue that (i) small animals are, in fact, more efficient at converting resources into welfare than humans are and (ii) the ideal population in terms of total welfare would be one composed in so far as possible of individuals of whatever species turns out to be most efficient. This world, though high in total welfare, has very low average welfare, lacks the 'best things in life' that are distinctive of truly excellent human lives, and is lacking in both species-diversity and wilderness. If Totalism suggests that we have reasons to try to bring about such a world, then this counts against the plausibility of Totalism as an adequate population axiology when nonhuman animals are included in the populations being assessed.

*4.1 Population engineering 1: Mouse World*

Sebo (2023a) has argued that we might be able to best maximize welfare by creating very large populations of small animals. He seems to assume that there would be animals of a variety of species in this world, and that some of them would be wild. I will argue that both assumptions are likely to be false. First, if our aim is to maximize total welfare, it isn't enough to recognize that large populations of small animals in general produce more welfare than small populations of large animals. Rather, since we have finite resources, we should aim to produce as many individuals as possible of *whatever species* turns out to be best at producing welfare. This means that the best population according to Totalism will not only be composed of small animals, but it will also, to the extent possible, be composed of individuals of only a single, or a small number of, species. Secondly, if we think that most wild animals have lives that are not worth living, then wild nature becomes a waste of

resources, as more welfare could be produced by a population of happy captive animals. Even if life for wild animals is on average net positive, it is unlikely that this is the most efficient way possible of producing welfare. This is simply because wild nature isn't optimised for *anything*, and certainly not the production of welfare. Compare the efficiency in producing calories for human consumption of farmland with that of wild land used by hunter gatherers to see the difference made by deliberate optimization. If these points are correct, then the ideal world in terms of total welfare will be one which is lacking in both biodiversity and wildness.

Let us assume the position of an ambitious would-be 'population engineer' who wishes to maximize total welfare in a non-speciesist way. Given that she has limited resources, she needs to think about what kind of population will most efficiently maximize welfare. Most importantly, she will need to determine whether there are any salient differences between species that will make some better generators of welfare than others. I suggest that there are such differences, and which species are most efficient for welfare-generation will depend on how we understand welfare. If we think that all sentient beings have the capacity for wellbeing, and that there are no sharp discontinuities between the kinds of wellbeing available to humans and those available to simpler animals, then we might think that the best possible population is one composed of very simple animals such as mice or oysters.[109] The reason for this is straightforward. Small simple animals generally require very few resources and little space, while large and complex animals like humans require a large amount of both. Given the obvious fact that we inhabit a planet with finite resources, if we are to maximize the welfare value of the population then we have good reason to fill the world with small simple animals rather than large complex ones, so long as the total amount

---

[109] This conception of welfare is endorsed by Bentham (1789), Sidgwick (1907), McTaggart (1927), and Katarzyna de Lazari-Radek and Peter Singer (2014).

of happiness in the 'small animal world' would be greater than that in the 'large animal world'.

In this section I will assume this conception of welfare is correct, and there is no sharp discontinuity between the kinds of wellbeing available to humans and those available to simpler animals. Given her goal of creating a population with the highest welfare value, her rejection of speciesism, and the limited resources available to her, how should our population engineer proceed? An obvious method is to identify which species of animal generates the greatest balance of positive wellbeing per unit of resources, and then attempt to move us towards a large population that is composed of individuals of that species to the extent possible. Human lives are generally high in welfare, but the cost of maintaining human lives is also high. It may turn out that there are simpler animals who can 'convert' resources into welfare more efficiently than humans do. Let's call this conversion rate 'Efficiency of Welfare Production' (EWP).

Since we have limited resources, we must consider EWP when thinking about what sort of population to bring into existence. EWP is likely to vary not only among different ways of producing welfare, and between different individuals of the same species, but also between species. This is because different animals may have comparable capacities for welfare, while some animals require much fewer resources than others. In particular, given plausible assumptions about the levels of welfare than nonhuman animals are capable of, human beings are inefficient welfare producers. Humans are large, complex animals. Compared to many smaller and simpler animals, they require a huge amount of resources to survive and thrive. Smaller animals may attain similar levels of welfare as humans do, while requiring far less food, space, and other resources. In the rest of this section, I will compare the EWP of human beings with that of mice, and I will argue that, given plausible

assumptions about the wellbeing experienced by mice, we have reason to try to bring about a large population of happy mice instead of a much smaller population of happier humans.

I have chosen mice for several reasons. First, their physical size means that they require much less food, space, and other resources than humans do. Second, they are mammals. This isn't to assume that there is any greater value in mammal happiness than happiness in non-mammals. Rather it is for epistemological reasons. While there are certainly many simpler animals than mice, once we leave the mammalian order, doubts about their sentience, and the kinds and degrees of wellbeing available to them, become much greater. No one reasonably doubts mouse sentience, and it seems that they enjoy many kinds of wellbeing which humans share and understand – pleasures like eating, playing, socialising, sex, exploring their environment, and so forth.

While it's certainly possible that we could produce more happiness by creating a very large population of animals simpler than mice (oysters[110] perhaps) this is less certain given our current understanding of sentience and wellbeing, and the physical structures required for their realisation. Of course, if one is more confident about the welfare capacities of even simpler animals, and of our ability to promote their welfare, then an even stronger argument can be made using those simpler animals in place of mice. Sebo (2023a) argues that, at least in theory, we should focus our efforts on even simpler animals such as insects, nematodes, and even microbes. This is because even if our credence that they are sentient is low, there are so many of them that the expected value of trying to increase their welfare is very high. Since this view relies on fanaticism about decision-making under uncertainty (see Bostrom 2009),

---

[110] Oysters, if we knew them to be sentient and capable of experiencing pleasure, would be ideally efficient pleasure producers. They take up almost no space and cost almost nothing to feed and maintain. There is some evidence of opioid receptors in bivalves such as mussels and oysters. We could easily raise trillions of bivalves, perhaps in a solution containing mild opiates. McTaggart's happy oyster could become a reality, albeit in the form of trillions of constantly replaced mortal oysters, instead of one immortal one (see also Crisp, 1997, pp.24-25). Perhaps we could also 'hedge our bets' by producing an animal Z world containing mice and oysters, especially since they live in different environments.

which leads to many difficult problems, I will set this position aside and focus on animals we can be quite sure are sentient.

In order to compare the EWPs of humans and mice we need to know how much welfare each is capable of realising, and what quantity of resources it takes for them to do so. This requires having a theory of welfare, and being able to measure and compare the extent to which humans and mice are capable of enjoying it. Let's assume simple Benthamite hedonism as our theory of welfare. We then need to ask how a mouse's capacity for pleasurable experiences compares to that of a human. Clearly it is not possible to measure intensity of pleasurable experiences directly. However, we can get a clearer picture by thinking about the degree of pleasure mice and humans experience from engaging in the same kind of activity. Consider the pleasure a human gets from eating a standard meal. Let's assume that the meal generates 1 unit of pleasure. Now imagine how much pleasure a mouse gets from eating a meal. Certainly, it's plausible that mice take pleasure in eating, for much the same evolutionary reasons that humans do. But how much pleasure does a mouse take in eating? Approximately the same amount? Ten times less than a human does? One hundred times less?  It seems implausible that the pleasure the mouse gains from her meal could be any less intense than that.

So, all else being equal, if we had the choice of giving a meal to a human or to a mouse, we ought to give it to the human. Not on speciesist grounds, but simply because the amount of pleasure generated for the human is likely more than the mouse would experience, and so we can best maximize welfare by feeding the human rather than the mouse. However, all is not equal. On average a human being eats 1.8kg of food per day. A mouse eats 3 grams. Even if the human being's pleasure in her meal is 100 times greater than the pleasure the mouse takes in hers, a human eats 600 times more than a mouse. If we want to maximise pleasure, without any speciesist bias, then we have reason to invest our limited resources in

producing food for mice instead of food for humans, as this will be at least 6 times more efficient, even when ignoring the fact that the financial cost of producing food palatable for humans is likely much higher than producing the same amount of food palatable to a mouse.[111]

This suggests that, at least when it comes to the pleasure of eating, mice have a higher EWP than humans do, unless we make implausibly low estimates for the amount of pleasure mice get from eating. Of course, it isn't just the pleasure of eating that we are trying to maximize, but overall welfare. This too is likely maximized by directing our resources into making happy mice. Like humans, mice take pleasure in sex, play, physical activity, exploring their environment, and so on. Again, it may be that the pleasure humans take in these activities is greater than that of mice. As we saw above, however, we can feed at least 600 mice for the same investment as it takes to feed a single human. For human pleasure to outweigh mouse pleasure then we would have to believe that the pleasure experienced by a human being engaged in one of the above activities is at least 600 times greater than the pleasure of a mouse engaged in the same kind of activity.

The above examples help clarify the notion of EWP and suggest that mice may be better converters than humans. It would be useful to have harder data on the welfare realising capacities of mice and of the costs of keeping them alive and happy. New research by the organisation *Rethink Priorities* (2022) may help shed some light on the former. The 'Moral Weight Project' is an effort to get a better understanding of the capacities for welfare of several commonly farmed animals. The study assumes hedonism; then based on several hypotheses about the function of valenced experience they identified 96 empirically measurable traits that can plausibly be used as proxies for welfare capacity. By surveying the

---

[111] A quick search for mouse food shows that one can purchase enough food to feed a mouse for well over a year for about £7.

literature on the behavioural and cognitive capacities of the selected farm animals they were able to come up with figures for the welfare ranges of each, where welfare range is the difference between the most intense positively-valenced state the individual can realise and the most intense negatively-valenced one. These ranges were then compared to the human range. The results of course cannot be expected to be fully precise, but they are suggestive. The welfare ranges of mice were not estimated as they are not farmed animals. However, the welfare range of pigs was estimated to be slightly over half that of human beings, and the range for chickens was one third that of humans.

It seems unlikely that mice would have significantly lower welfare ranges than chickens. As a reasonable figure, based on the findings of *Rethink Priorities* for pigs and chickens, I will assume that mice have a welfare range that is one quarter that of humans. That is, mice are capable of a range of positively and negatively valenced states up to 25% of the intensity of human states at any given moment. Mice are social and intelligent animals, and they seem to display many of the traits that the study uses as proxies for welfare capacity, such as play behaviour. Second, one tentative conclusion of the study is that, for all the vertebrates studied, none has a welfare range that is more than twice that of any other. Since the study also included carp and salmon, we would have to believe that mice have a lower welfare range than these animals in order to have a range lower than 0.25. Of course, this may be the case, but it is very counterintuitive. Finally, mice are more closely related to humans than chickens *and* pigs. Our last common ancestor with mice lived around 70 million years ago, compared with 90 million for pigs, and over 300 million for chickens. None of these considerations are decisive, but they do suggest a relatively high welfare range for

mice. Until further research is done, I will continue to assume a welfare range for mice of 0.25 that of humans.[112]

Based on the trivial cost of feeding and housing mice, and the reasonable assumption that they can generate 25% as much positive welfare as a human being at a given time, it seems that a population engineer who was impartially concerned with creating as much positive welfare as possible might try to move us towards a population composed more of mice than of people and other species. This is because they have a higher EWP than humans do, and in a world with finite resources we have reason to take the most efficient route to our goals. If our goal is maximizing welfare, we can do this more efficiently by making happy mice rather than making happy people. If the arguments in this section are right, then we have reason to aim for a very large population of happy, domesticated mice, with the minimum human population necessary to maintain it. Ultimately, we may not need humans at all, if we can be replaced with artificially intelligent machines who can be given the task of looking after the mice for us. Mice, in virtue of their very low resource requirements and consequent high EWPs, are something like a utility monster. Instead of a single being who enjoys much more utility than all others combined, though, they dominate our utility calculations through sheer numbers – we could label them a 'utility swarm'. If we develop sentient AIs in the future, it may be possible to create even simpler beings with even lower resource requirements who would make up a Z population with even higher total welfare, that is even more difficult to accept.[113]

---

[112] Bob Fischer at Rethink Priorities is currently researching the welfare ranges of mice. In personal correspondence he said he would put the welfare range of mice at around the same level as that of chickens, that is about 0.33 that of a human. However, he stressed that he has low confidence in this estimate.
[113] Sebo describes humans as a kind of utility monster, from the point of view of smaller animals.

*4.2 What's bad about Mouse World?*

The idea that animal Z might be more valuable than an A world of extremely happy humans seems repugnant; the idea that we might have moral reason to try to move closer to such a world seems even worse. Yet if we accept the assumptions I have made above, it looks like we do have such reasons. It is worth spelling out exactly what seems so repugnant about that conclusion, and what would be missing from animal Z.

The first thing that seems bad about animal Z is the very low welfare in each life. The welfare range of a mouse may be about 25% that of a human being, meaning the maximum positive welfare they can experience at a given time is about one quarter as intense as that available to a human. Even worse, since each mouse only lives for a couple of years, while humans might live to 80, the total lifetime welfare of each mouse must be divided again by 40. Taking these figures literally then, a mouse in animal Z would have a lifetime welfare level about 160 times lower than a human in any possible human Z world. Although in thought experiments we can imagine humans having such low lifetime wellbeing levels, in practice it seems hardly possible to have a viable population of human beings at such a low level of lifetime wellbeing. Average welfare in animal Z is much lower than that in any possible human world approaching Z.

The second bad thing about animal Z is the lack of all the higher pleasures and best things in life. As we saw, these things might also be absent from a human Z world, at least in theory. In practice, though, it is hard to see why any possible human Z-like world would lack all these good things, especially since many of the higher pleasures are in fact cheap to provide (think of digital copies of all of Shakespeare's work, etc.). Animal Z by contrast would certainly lack all of these good things. Worse, it is not only the higher pleasures that are lacking in animal Z – humans and other animals enjoy a range of pleasures that would simply be absent in a world of mice. It should be stressed here that these are not purely

theoretical points; rather they are made more pressing by the fact that we could bring about such a state of affairs. As I have tried to emphasize, the difference between a human Z and an animal one may be relatively trivial when we consider those worlds as thought experiments. Without the constraints of reality, we can imagine humans with lifetime welfare levels as low as those of mice, humans who, for some reason, are unable to enjoy any of the nobler pleasures characteristic of human life and so on. A realistic human Z world is unlikely to be quite so bad – it is likely that there is a practical 'welfare floor' below which it would not be possible to sustain a large population of humans, and there is no reason to think that all of the best things in life would be lacking from any realistic human Z world - whereas a realistic animal Z could be very bad indeed in terms of average welfare and loss of the best things in life.

The third thing lacking from animal Z is species-diversity. The reason for this is simple. Our population engineer's strategy for maximizing welfare given her limited resources was to discover the species with the highest EWP and to try to increase their population as much as possible. If successful, this would involve taking resources away from less efficient species and allocating them to the more efficient ones. Related to this is a lack of wild animals generally. Wild animals are a waste of resources in two ways. First, most of them seem to have net negative lives,[114] and second, even for those that have positive lives, elephants and chimps for example, they are likely to be much less efficient at converting resources into wellbeing than captive mice. This consideration is absent from Sebo's account of the ARC. While a committed total utilitarian might be happy to sacrifice biodiversity and wildness to attain more total welfare, most people would not be so blasé about this loss. If Totalism implies that the best world is one which lacks these values, then this may make it more unappealing as an account of the value of populations when nonhumans are included.

---

[114] See Ng (1995), Tomasik (2009b) and Horta (2010).

In practice, a perfectly implemented animal Z world seems unlikely, even if it is possible. However, the Normative Problem of the ARC doesn't go away entirely just because animal Z is unlikely to be fully implemented. The cost-effectiveness of creating happy mice seems so high compared to effective human interventions that it ought to be taken seriously as a way of making the world better, at least by utilitarians, and it is surely implausible that they have all-things-considered reasons to create colonies of happy mice.[115] Secondly it suggests that we should not regret losses in wilderness and losses in species, in so far as these bring us closer to the ideal state of affairs. The idea that we have moral reason to try to create a world which has (1) such low average welfare (2) lacks the best things in life (3) is severely lacking in biodiversity and (4) lacks wild animals living free from human control is enough to make us doubt that a Totalist welfarist axiology is sufficient for thinking about the value of populations that differ in species composition.

Furthermore, many of those who are concerned with animal welfare should also find this world unappealing. Though Mouse World takes animal welfare seriously, and, indeed, maximizes it, this doesn't seem like the kind of world those who care about animals might wish for, since it is so lacking in wildness and biodiversity. Some who emphasize the reduction of animal suffering, such as Tomasik, might welcome something like Mouse World. Others might find the loss of non-welfare value too high. The fact that the world with the highest total welfare might be so lacking in biodiversity seems to have been missed, by Sebo and others. Johannsen for example suggests that a world with more individual animals and more species of animals is better than a world containing fewer, and this is at least partly

---

[115] For an argument that egalitarianism might require a massive shift of resources from humans to mice, since mice are so badly off compared to humans, see Vallentyne (2005). Matheny and Chan (2005) argue against the 'logic of the larder', that is, the idea that animal agriculture benefits animals as it causes them to exist with lives worth living. Part of their argument is that, if we take the logic of the larder seriously, and want to use our money to create happy animals, there are more efficient ways of doing so than farming animals. Since mice have very low resource requirements, we ought to renounce animal agriculture and "adopt a vegan (vegetarian) diet and invest savings in colonies of mice" (p.589).

because such a diverse world would have higher total utility (2021, p.73). This seems to be false, however – the world with the greatest utility might turn out to be one with only one or a very small number of species of 'efficient' animals. Animal advocates who prefer a world with a diverse range of animal life then ought to defend the value of that world in terms of values other than total welfare.

**5 The Best Things in Life and Human World**

*5.1 Population engineering 2: Human World*

In section 4 I assumed a simple hedonistic conception of welfare. On this conception there are no sharp discontinuities between the constituents of welfare available to humans and those available to simpler animals. This opened the door to the ARC – since small simple animals are capable of attaining levels of welfare comparable to those of humans, while requiring far fewer resources, a population engineer attempting to maximize the welfare value of the population might do so by maximizing the population of those simple animals with the highest EWP and minimizing the populations of more complex animals such as humans.

In this section I will assume that there is something special about the constituents of welfare available to typical humans which means that human beings are the species with the highest EWP. This could be because the higher pleasures available to human beings are lexically superior to those lower pleasures available to simpler animals as Mill claimed. This kind of strong lexicality view implies that a single human life containing the higher pleasures would be better than any number of lives without those pleasures. Some have tried to use more sophisticated forms of lexicality to avoid the RC without entailing this equally implausible conclusion. In particular, Nebel (2022) has developed a view of lexical thresholds which allow him to avoid this conclusion, and Parfit (2016) has suggested a form

of perfectionism according to which the 'best things in life' are imprecisely lexically superior to lower quality components of welfare, such that any loss of these best things cannot be compensated for by increases in the amount of lower quality constituents of welfare.

Since I am concerned with the Normative Problem of the ARC rather than the axiological problem, I need not assume any form of lexicality. If my concern was with axiology, then I would have to provide an account of lexicality that gives the right answers in hypothetical cases involving the existence of impossibly vast numbers of nonhuman animals and much smaller human populations. Claiming the superiority of the smaller human population in imaginary cases requires something like lexicality, as we can simply keep increasing the numbers of animals in the imagined animal population. Since I am concerned with the Normative Problem posed by the ARC, I can simply assume that those human pleasures produce vastly, but finitely, more welfare than lower animal ones, as there is of course a limit to how resource efficient simple animals can be compared with humans.

I will simply assume then that there is a certain class of constituents of welfare available only to humans which are so much more valuable than those available to simpler animals that, in practice, the greater resource efficiency of simpler animals will not be sufficient to make their EWP greater than that of humans. That is, there is a class of distinctively human activities that contributes so much to welfare compared to those simpler pleasures also available to nonhuman animals that it would be practically impossible to create enough simple animals like mice living good lives that would produce more welfare value than a smaller population of humans engaged in those activities. I assume that this class of distinctively human activities contains things similar to those identified by Mill as the higher pleasures, and by Parfit as the best things in life, such as doing philosophy, listening to excellent music, and so on.

Again, our population engineer wishes to maximize the welfare value of the global population given finite resources, and once again there is one species which has a higher EWP than any other, humans. How might she proceed? A rational programme would be to increase the human population as much as possible, in so far as this is compatible with humans continuing to engage in these highly welfare efficient activities. This would presumably involve maximizing the living space available to human beings – in so far as it is possible, she would want to 'pave over' wilderness to make room for more humans. This would mean reducing wild animal populations significantly. It might also involve significant reductions to the population of domesticated animals. Since most domesticated animals are used for food, and since plant-based diets are far more resource efficient, she would have reason to eliminate wasteful practices of animal agriculture in order to maximize the human population. The end goal of her efforts would be a vast population composed almost exclusively of human beings, each of whom spends their lives enjoying the most resource-efficient of the human welfare-producing activities.

## 5.2 What's bad about Human World?

Appealing to a more sophisticated conception of welfare according to which humans are vastly better at generating welfare than simple animals has allowed us to avoid the ARC in practice. In some ways the resulting human world is quite appealing – there is a large human population enjoying a very high total and average level of wellbeing, and not only are the 'best things in life' not missing, they are central to the lives of each person. There is much more diversity of pleasures in the human world than there is in the mouse world. In Mouse World the same simple murine pleasures were repeated endlessly. The pleasant activities and contributors to positive welfare in humans are much more varied. However, it is still likely to be the case that there is less diversity of welfare-generating activities in this human world

than there is in our own; since EWP varies among different kinds of activities as well as among species, we can expect our population engineer to try to ensure that the population enjoys the most efficient welfare-producing activities available, and this may be a relatively small subset of those enjoyed by humans today. Human World would also lack all those welfare-producing activities enjoyed by other animals which humans don't enjoy, such as the pleasures dogs take in scents, that birds take in flight and so on.

There are some potentially important things missing from this world too, however. Like Mouse World, this world is severely lacking in diversity of animal life and of wilderness generally. The reason for this lack is the same in both worlds – in each world there is one species with a higher EWP than all others, and, given the fact of finite resources, in order to maximize the welfare of the population our population engineer ought to try to shift the species composition of the population so that it contains as many of these efficient animals as possible. Other, less efficient animals ultimately become a waste of resources. Wilderness too is a waste – if most wild animals have bad lives, then instead of having wild spaces in which free living animals experience net negative welfare it is obviously better, in so far as it is possible, to replace these wild spaces with human-controlled ones. Even if animals in the wild lived marginally net positive lives, it is unlikely that uncontrolled wilderness would be the optimal way to produce the highest total welfare.

This human world may be appealing to committed welfarist Totalists who also believe in the superiority of human welfare-generating activities. It may also be appealing, or at least unproblematic, to those WAS interventionists who argue for the elimination of wild spaces to reduce WAS (Tomasik, 2016). It is less appealing for those who are not committed welfarist Totalists, and for those interventionists who would prefer to solve the problem of WAS by making animals happy rather than making them extinct. In practice, our population engineer is unlikely to be able to fully implement this human world either. Still, some unpalatable

normative conclusions remain even if full realisation proves impossible. First, if the arguments in this section are correct then we have less reason to worry about human caused extinctions, and the constant encroachment of human beings into formerly wild areas. In fact, we have reason to continue such activities so long as continued human expansion doesn't threaten our long-term ability to maximize our own population. If humanity ever spreads to other planets, we should certainly not bring animals with us. We should see animals not as wondrous fellow creatures sharing the world with us, but as regrettable axiological cul-de-sacs, resource sinks that prevent us from maximizing total welfare. We might wish to care for those animals who exist, while arranging for their gradual extinction. The long-term future ought to be a human one.

## 6 Conclusion

In this paper I have investigated some of the consequences of including animals in our thinking about the value of populations. I have argued that the RC can easily be extended to animals, resulting in the ARC. While the RC may be a merely theoretical problem, endorsing the ARC has practical implications that few would be willing to accept, and this puts pressure on those who accept the RC. Either they must (1) reject the RC; (2) reject the ARC; or (3) find a way to block the natural connection between the axiological and the normative implications of the ARC. For those who are undecided, Totalism's unappealing normative implications once animals are included in the analysis may count against it as an adequate population axiology.

Committed Totalists might accept the normative implications I have outlined, despite their unpalatability. Which kind of population best maximizes total welfare then turns out to depend essentially on what conception of welfare we adopt. If we adopt simple hedonism then the fact that some animals can achieve levels of welfare comparable to those attained by

humans, but in a much more resource efficient way, suggests that the ideal population would be one composed of very simple animals such as mice. If, on the other hand, we think that human welfare is far superior to that of other animals, then the ideal population may be one composed entirely of human beings engaged in typically human activities of the sort described by Mill and Parfit as higher pleasures/the best things in life. I suggested that neither of these populations is intuitively very appealing, primarily because both worlds are severely lacking in species-diversity, wild animals, and diversity of sources of welfare. If welfarist Totalism really implies that one of these two worlds would be ideal, then this is a counterintuitive implication of that view. The populations I described may also be unappealing to those animal advocates who value the continued existence of a wide variety of wild animals, though it may be amenable to those, like Tomasik, who propose paving over nature as a solution to wild animal suffering. For those of us who, like me, find the two worlds I have described above unappealing, it may be the case that we need to adopt a richer population axiology than welfarist Totalism if we are to accommodate animals in the populations analysed.

# Conclusion

In this thesis I have explored some of the moral implications of wild animal suffering and the moral complexities that arise once we begin to systematically intervene in nature to reduce that suffering. I began in the introductory chapter by describing the nature and scale of WAS. I listed and categorized its major causes, argued that naturogenic causes of suffering are both more important practically, and more interesting theoretically, than anthropogenic causes, and distinguished and labelled importantly different claims about the scale and severity of WAS. I briefly outlined the case for intervention in nature to reduce animal suffering. I argued that, for utilitarians, the case for intervention is very simple, and depends only on the facts about WAS, and our ability to mitigate it safely and effectively. The case for intervention is also very strong for non-utilitarians who accept a principle of beneficence, though they may not accept all the same specific interventions that utilitarians do. This is because some interventions may violate animals' negative rights and others may not be compatible with other values. I listed the various proposals for intervention that have been made in the literature, then briefly discussed some of the arguments against intervention. Finally, I summarized the main aims of the papers that constitute this integrated thesis, and explored the connections between them, as well as some of the overarching themes of the thesis.

*Summary of papers*

In the first paper I explored Milburn's (2022) version of the relationality objection to intervention. For the sake of argument, I assumed that positive duties only arise in the context of morally salient relationships. I then argued that, if we accept Milburn's claim that we are morally entangled with domesticated animals in virtue of the actions of our ancestors, then we are also morally entangled with wild animals in significant and pervasive ways. This is because our ancestors, both in historical times and in prehistory, acted in ways that created

vulnerability in wild animals on a very large scale. I suggested several ways in which the actions of our ancestors have morally entangled us with wild animals. By hunting large herbivores to extinction, it is possible that our ancestors inadvertently created ecosystems with more suffering than the ones they replaced, primarily because it allowed for the greater proliferation of smaller animals, who are overwhelmingly r-strategists, and likely have lives that are much less good than those of larger, K-selecting animals. Humans have also introduced animal and plant species into new environments all over the Earth, from prehistory to modern times. This created new vulnerabilities in the native animals outcompeted, displaced, or predated by the introduced species. Finally, humanity has taken over and transformed most of the habitable land on the planet, thus forcing wild animals to adapt to these transformed environments. I concluded that, even on a relational account, it may be the case that we are sufficiently entangled in the lives of wild animals that we have significant positive duties towards them.

Furthermore, this pervasive entanglement with wild animals may also be relevant to those who reject relational accounts and argue for intervention on the grounds of beneficence. Johannsen in particular argues for a collective obligation to intervene in nature on the basis of beneficence but accepts that we may also have duties grounded in rectificatory justice. He thinks that rectificatory justice will play a relatively minor role, however, as most harms endured by wild animals are not anthropogenic. If the arguments I have made in the first paper are correct, then we are much more entangled with wild animals than previously thought. It is not only modern impacts on wild animals, such as climate change, that entangle us in their lives. Rather, we are, and have been, significantly entangled with many wild animals for centuries. This dramatically increases the extent of our rectificatory duties towards them. If duties of justice are stronger than duties of beneficence, then our moral obligation to intervene in nature is even stronger. Of course, all of this depends on a

controversial view of moral responsibility, namely, the idea that humanity as a whole forms a morally significant collective, and that people today can have duties to rectify harms caused by their ancestors. The case for reparations for historical wrongdoing between human groups is already extremely controversial – the idea that we may have similar duties to animals because of their historical mistreatment by our ancestors seems even more difficult to accept.

In the second paper I argued against Kyle Johannsen's claim that the non-identity problem doesn't affect our reasons to intervene in nature. I argued that the non-identity problem changes both the nature and the strength of our reasons to intervene. By intervening in nature on a large scale we change which animals come into existence. In the case of harmful animals, we foreseeably enable them to inflict harms on other animals. In the case of harmed animals, we put them in harm's way, by causing them to endure harms instead of some other animals. By intervening significantly in nature, we can transform naturogenic harms, for which we are not responsible, into anthropogenic ones, for which we are. Even if intervention reduces the total number of harms endured, this doesn't eliminate our duties to rectify the harms that we have inflicted on the animals we do cause to exist. Furthermore, our duties to rectify harms we have done are duties of justice, and these are usually stronger than our beneficence-based reasons. By intervening in nature then we take on secondary duties towards and regarding the animals we have caused to exist, and these duties may be extremely demanding.

If these secondary duties are costly enough, then we might think that intervention is supererogatory. If WAS is sufficiently bad, however, then we may still have a duty to intervene in nature, despite the heightened costs of doing so. That is, the worse the suffering in nature is, the stronger our reasons of beneficence are to intervene. Furthermore, some of the other papers in this thesis may provide additional support for intervention. The paper on longtermism, by taking the very long-term future into account, highlights the scale of WAS.

If we do nothing to mitigate it then we can expect WAS to continue on our planet as long as it is habitable, that is, for another billion years or so. If the suffering of all these future animals also matters morally, this increases the strength of our moral reasons to begin a process of intervention to prevent or reduce it. If my arguments in the first paper are correct, then human beings may already be significantly entangled with wild animals. If we are already partly responsible for many of the apparently naturogenic harms endured by wild animals, then we already have reasons to intervene that are based on rectificatory justice. While we might be permitted to refrain from intervening out of beneficence if the costs of doing so are too high, we are not permitted to shirk our duties to rectify harms that we are responsible for. If this is correct, then intervention may be mandatory, despite the high secondary costs.

Once we begin intervening in nature in large-scale, systematic ways, however, it may no longer be permissible for us to stop. The more we intervene in the workings of the natural world, the more responsibility we take on for what happens there. After a sufficient degree of intervention, it is no longer accurate to describe the natural world as something independent of human choices. Once wild animals have been brought into the human world, and under our control, we take on duties of justice towards them. Even if we have done so for good reasons, such as beneficent impulses to prevent harm, once we have brought beings from the natural world into the social world, we can no longer claim to have merely negative duties towards them. Once we bring wild animals into the human world, we will, in a sense, have domesticated them, and will owe them positive duties, too. It may be, for example, that once wild animals are sufficiently under our control, they will become something like citizens, and questions of justice in the distribution of resources may apply to them too (Donaldson and Kymlicka, 2011).

The third paper was an investigation into the implications of WAS for spreading life beyond the Earth. I argued that recent proposals to spread life throughout the galaxy by

directed panspermia are impermissible, on both consequentialist and non-consequentialist grounds. The consequentialist argument rested on the predictability of WAS in an unguided biosphere. I argued that WAS is not the contingent outcome of evolution on our planet, but the predictable result of necessary features of evolution by natural selection. This means that we can expect net suffering, or at least net bad lives, on any planet on which sentient life evolves without intelligent oversight. For a classical utilitarian, launching a directed panspermia mission is impermissible if the expected outcome is net suffering. For consequentialists who also value certain distributions of welfare, such as egalitarians or prioritarians, or for those who place more moral weight on suffering than on wellbeing, directed panspermia missions are likely to be impermissible if the weaker Net Bad Lives claim is true.

I also gave a non-consequentialist argument against directed panspermia, based on the duties that creators have to their sentient creations. Drawing on the literature on the ethics of procreation, I proposed two thresholds for permissible creation. The weaker threshold is that creators must ensure that their sentient creatures will have at least a life that is worth living. The stronger threshold is that creators owe their creations a reasonable chance of flourishing. If the Net Bad Lives claim is true, then creation of creatures by directed panspermia cannot meet either threshold for permissible creation. If Net Bad Lives is false, but Net Failure to Flourish is true, then perhaps only the stronger threshold cannot be met.

I also considered four objections to my arguments. The first objection is that we may have an overriding duty to safeguard the existence of life in the universe, and directed panspermia missions may be the only currently technologically feasible means of doing so. I pointed out that there are other ways of discharging our putative duty to preserve life that do not run the risk of causing huge amounts of WAS. If we can wait until these methods are technologically feasible then this would be preferable. If it is imperative that we act to

185

preserve life immediately, then we could launch a modified directed panspermia mission that would only begin the seeding process if the craft loses contact with Earth. The second objection concerned the possibility of creating 'designer microbes' that will not evolve the capacity to suffer in the future. I argued that this may be impossible. If such microbes could be designed, however, then a mission to seed the galaxy with them would count as a guided life-propagation mission rather than an unguided one, and may be permissible.

The third objection pointed out that my argument might commit us to regretting that life ever evolved on our own planet. I defended the rationality of being glad that life evolved on Earth, while also maintaining that launching a directed panspermia mission is impermissible. My defence, however, relies on an uneasy mixture of optimism about our future, and pessimism about the prospects of happy or intelligent life evolving in an unguided biosphere. Finally, I considered the objection that my consequentialist argument against launching a directed panspermia mission commits me to saying that we should sterilize any planets we discover that already harbour indigenous microbial life. I argued that the relative costs of launching a directed panspermia mission and sterilizing planets with indigenous life, as well as the doing/allowing distinction, might make a meaningful moral difference between the two cases. I accepted, however, that at least in some unlikely situations my consequentialist argument commits me to the counterintuitive conclusion that we should sterilize planets with native microbial life.

In the fourth paper I argued that the basic argument for longtermism applies just as much to animals as it does to humans. In so far as they ignore animals, then, longtermists have made an important mistake. Drawing on the literature on longtermism, I outlined the basic argument common to most longtermists. I then showed in detail how the main premises apply to animals. In particular, I argued that *Big Future* and *Causal Efficacy* are true of nonhuman animals. In order to deny *Big Future,* one would have to show that it is

overwhelmingly likely either that animal populations will drop to zero or near zero and remain there, or that human, but not animal, populations will be multiplied dramatically by spreading beyond Earth, by being instantiated in digital minds, or both. Since this cannot be done, the expected population of nonhuman animals must be admitted to be large enough to secure the truth of *Big Future.*

Using the longtermist concepts of persistent states and value lock-in, I argued that *Causal Efficacy* is also true of animals. That is, there are acts we can perform in the near-term future that can be expected to make the long-term future go better or worse for animals. One way in which we could make things worse is by spreading WAS beyond the Earth, for example, by directed panspermia, or by accidental contamination of another planet. This would increase both the scale of WAS and its persistence, since we can expect WAS on multiple planets to last longer than animal suffering confined to Earth. We can also work now to ensure that the long-term future goes better for wild animals, by researching safe and effective ways of intervening in nature to improve the lives of wild animals on a large scale.

It is also possible that we make the long-term future go worse for animals by locking in a set of values that are not conducive to the wellbeing of animals. One way in which we could do this is by locking in the speciesist values of current-day society. If this were to happen, then we might continue exploiting animals on an industrial scale for eons. Regarding value lock-in, I suggested that we may be in an especially critical time for animals. If it is the case that humanity begins to colonize the galaxy later this century, then we have a brief window of opportunity to abolish the exploitation of animals before colonization begins. If we succeed in doing so, we would effectively be locking-out the possibility that colony worlds across the galaxy would also exploit animals, thus making the future go significantly better for animals.

Finally, I briefly suggested some promising longtermist interventions that longtermists, or animal advocates, might consider carrying out. These included: the creation of strict laws governing the spreading of life beyond Earth, as well as the enforcement of strict decontamination procedures for spacecraft originating from Earth, in order to avoid spreading WAS; investing in research to develop means of effectively reducing WAS, and supporting efforts to reduce the risk of human extinction, in order to increase the chances that WAS is reduced dramatically in the future; trying to cultivate values that are more conducive to animal wellbeing, by promoting anti-speciesism/moral circle expansion, and educating people on the problem of wild animal suffering. To determine whether these interventions are likely to be effective will require significant further research.

The fifth paper explored how animals can be included within our population ethical thinking. This is important for interventionists, as our interventions in nature will inevitably change wild animal populations in morally important ways, such as their size, total and average welfare, and species composition. If humanity takes control of the biosphere, and determines what the global population will be like, then we should try to understand how animals can be incorporated into our theories of population ethics. I argued that Totalism has counterintuitive normative implications once we include nonhuman animals in the populations to be assessed. I argued that Totalism implies either that we should try to bring about a population composed as much as possible of mice, as they are so efficient at converting resources into welfare, or that we should bring about a population composed exclusively of humans, if we think that there is something special about human welfare that makes humans the species with the highest EWP. In either case, Totalism implies that we should bring about a world with very low species-diversity. This is because, if we are trying to maximize total welfare, we should try to populate the world with whichever species has the highest EWP. Having animals around with lower EWPs would simply be a waste of

resources, leading to a less than optimal population in terms of total welfare. Similarly, Totalism implies that a world without wilderness is better than one with it. This is because wild spaces, that is, those spaces not intensively managed and controlled by intelligent beings, aren't optimized for anything. Much in the same way that wild lands aren't optimal for the production of calories for human consumption, but intensively farmed fields of wheat or soybeans are, wild areas aren't optimized for maximizing total welfare. If we are good Totalists, then, we ought to prefer a large population of whatever species has the highest EWP, living in a world that has as little wilderness, and as much intensively managed welfare-producing land as possible. If this is too unappealing, then we may need to adopt a richer population ethical theory than Totalism in order to adequately accommodate nonhumans within our population ethical thinking.

*Overall conclusions, contribution to literature, and directions for further research*

Throughout this thesis, I have assumed that the suffering of wild animals is bad enough that it gives human beings strong moral reasons to intervene in nature to reduce that suffering, and that, fundamentally, these are reasons of beneficence. That is, I've assumed that even if human beings are not responsible, causally or morally, for the suffering endured by wild animals, we nonetheless have very strong reasons to reduce that suffering, simply because it is good to help others when we are in a position to do so.

However, in several of the papers I have argued that humans may in fact have more responsibility for the suffering of wild animals than is commonly thought. In the first paper I suggested ways in which human beings may have entangled themselves in the lives of wild animals, from prehistory up to the present day. I made this argument in response to Milburn's version of the relational objection, which depends on the idea that modern humans can be responsible for the actions of their ancestors. My aim was to show how even on a relational

189

account, we can have strong duties to intervene in nature. If we accept the idea that we can have moral duties today in virtue of the actions of our ancestors, then this has significant implications. I argued that human beings have altered the biosphere in such significant, pervasive, persistent, and often harmful ways that we may well have extremely wide-ranging duties to wild animals today, quite aside from duties grounded in beneficence. If this is true, then much WAS on Earth really isn't naturogenic at all; rather it is the result of human choices, and so our duties to wild animals are far more pervasive and stringent than commonly assumed. The paper is, to my knowledge, the first to respond to Milburn's version of the relationality objection, and the first to suggest that humanity's historic, and prehistoric, impacts on the biosphere may be sufficient to generate rectificatory responsibilities for modern humans. Defending the idea that humanity as a whole forms a morally relevant collective, and determining precisely what responsibilities this generates for us today, would require far more work.

The second paper also suggest ways in which human action can make us responsible for what are typically considered to be naturogenic harms. Following Johannsen, I agree that rescuing harmful animals can make us responsible for the harms that they will go on to inflict. Developing this idea, I argue that we can also become responsible for animals by creating them, and even by causing them to exist by performing identity-affecting actions, so long as all three responsibility-generating conditions are met (causal responsibility, foreseeability, and non moral agency). Furthermore, I point out that even our beneficent interventions in nature to reduce suffering generate secondary moral duties towards the animals we thereby cause to exist. Since these costs may be very high, it may be the case that intervention is supererogatory. The extent of WAS, and the possibility that we already have extensive duties to rectify harms done to animals (paper 1), however, likely mean that intervention is still obligatory. The more we intervene, the less natural the natural world is,

and the more it becomes part of the human world. Ultimately this may mean that humanity has a duty to take control over the biosphere, bringing it within the sphere of justice. In Nussbaum's evocative phrase, humanity must work to ensure "the gradual supplanting of the natural by the just" (2006, p.400).

This paper is to my knowledge the first to explore how the non-identity problem affects our reasons to intervene in nature to reduce WAS. It may also be original in investigating how the non-identity problem applies to lives that are not worth living. Typically, the non-identity problem concerns either the creation of a life that is flawed, but worth living, or the choice between two lives, both of which are worth living but one of which is significantly worse than the other. The choices I explore in this paper involve lives that are not worth living. The choice in *Mutagenic Gas* is a same-number case choice between allowing a very badly off person to come into existence, or acting to cause a less badly off person to come to exist instead. The choice in *Beneficent Intervention* is a different-number choice, between allowing a large number of badly off animals to come into existence, or acting to cause a smaller number of equally badly off, but different animals to come into existence instead. My claim is that we may have reason to compensate the badly off individuals we have caused to exist, even if causing them to exist was, overall, the right thing to do, since it prevented worse-off, or a greater number of equally badly-off, individuals from coming into existence. This aspect of the non-identity problem sems to be worthy of further research.

Several of the papers, especially the third and fourth, emphasized the importance of the long-term future for animals. The third paper is the first to deeply explore what WAS entails for spreading life beyond the Earth. Making the connection between space ethics, procreative ethics, and animal ethics seems to be an original contribution, and I hope that it will prove useful for further work. Also in the third paper, the arguments that net WAS or Net

Bad Lives is not merely a contingent outcome of evolution on our own planet, but a predictable outcome of the evolutionary process generally, are worthy of further research. The fourth paper is one of the first to systematically argue that the basic argument for longtermism applies to animals as well as to humans, and to explore what this means for both longtermists and animal advocates. Working out which long-term interventions are likely to be most effective will require a great deal more work, however.

Finally, the fifth paper ought to be of interest both to advocates of beneficent intervention in nature and to population ethicists generally. Interventionists need a viable theory of population ethics that applies to mixed species populations. Without one, it is difficult to assess the value of interventions in nature that change not only total and average wellbeing, but the numbers of animals who exist and the species composition of the biosphere. Population ethicists should aim to discover a theory that can encompass not only human beings, but animals too. Without one, their theories must remain incomplete. I have argued that Totalism faces difficult challenges when animals are included in the populations to be assessed. The Normative Problem arises once we realise that an animal Z world, unlike a human one, is a realistic possibility, and not merely a theoretical challenge. The fact that different species of animal have different EWPs opens the door to the troubling conclusion that the best attainable population, in terms of total welfare, is one composed of a huge number of individuals of the most efficient species. It may also be the case that other population ethical theories fare equally badly with animals, but to determine whether this is so requires more research. Much more work needs to be done in this area, both by wild animal ethicists and by population ethicists.

Human beings are one species among many, but they are undoubtedly a special one. Our rational, technological, and moral powers generate important responsibilities which we must rise to meet. The other animals with whom we share our world call on us in several

ways. The domesticated animals currently living under human domination and enduring human exploitation call on our mercy, while those living in the wild, suffering purely naturogenic harms, call on our beneficence. The existence of nonhuman animals exerts a call on us to improve our philosophical theorizing as well. Taking nonhumans seriously means reassessing our philosophical assumptions and theoretical frameworks, often in radical ways. I have argued in this thesis that taking animals seriously means reassessing our theories in population ethics, longtermism, environmental ethics, and space ethics. Undoubtedly taking animals seriously will disrupt other areas of philosophy as well. Identifying how animals affect our theories, and working out how to include them in our picture of the world pose significant philosophical challenges. Working out how to make the world a good place, not only for humans, but for all animals, raises ethical and technological challenges which may take generations to solve. We are not responsible for the existence of suffering, but we are responsible for abolishing it. We should begin.

# Bibliography

Agar, Nicholas (1997) 'Biocentrism and the concept of life', *Ethics,* 108(1): 147-168.

Animal Ethics (2023a) 'Animal exploitation: Introduction'. Available at: https://www.animal-ethics.org/animal-exploitation-introduction/ [accessed 21/07/23]

Animal Ethics (2023b) 'Helping animals in the wild'. Available at: https://www.animal-ethics.org/helping-animals-in-the-wild/ [accessed 21/07/23]

Animal Ethics (2023c) 'The situation of animals in the wild'. Available at: https://www.animal-ethics.org/situation-of-animals-wild/ [accessed 21/07/23]

Animal Ethics (2020) *Introduction to wild animal suffering: A guide to the issues*. Oakland: Animal Ethics. Available at: https://www.animal-ethics.org/wp-content/uploads/Introduction_to_Wild_Animal_Suffering.pdf [accessed 21/07/23.

Anthis, Jacy Reese, and Paez, Eze (2021) 'Moral circle expansion: A promising strategy to impact the far future', *Futures,* 130.

Anthony, Andrew (2021) 'From fringe to mainstream: how millions got a taste for going vegan', *The Guardian.* Available at: https://www.theguardian.com/lifeandstyle/2021/oct/10/from-fringe-to-mainstream-how-millions-got-a-taste-for-going-vegan [accessed 21/07/23]

Armstrong, Stuart and Sandberg, Anders (2013) 'Eternity in six hours: Intergalactic spreading of intelligent life and sharpening the Fermi paradox' *Acta Astronautica,* 89: 1-13.

Arrhenius, Gustaf (2000) 'An impossibility theorem for welfarist axiologies', *Economics and Philosophy,* 16(2): 247-266.

Austad, Steven and Hoffmann, Jessica (2018) 'Is antagonistic pleiotropy ubiquitous in aging biology?', *Evolution, Medicine and Public Health,* 1: 287-294.

Axelrod, Robert and Hamilton, William D. (1981) 'The evolution of cooperation', *Science,* 211(4489): 1390-1396.

Bakker, Elisabeth et al. (2015) 'Combining paleo-data and modern exclosure experiments to assess the impact of megafauna extinctions on woody vegetation', *PNAS,* 113(4): 847-855.

Banks, Sam et al. (2007) 'Sex and sociality in a disconnected world: a review of the impacts of habitat fragmentation on animal social interactions', *Canadian Journal of Zoology,* 85(10).

Barraclough Jones, Katy and Wallace, Stuart (2021) 'Toombes v Mitchell: Wrongful life v wrongful birth'. Available at:

https://www.dacbeachcroft.com/es/gb/articles/2021/march/toombes-v-mitchell-wrongful-life-v-wrongful-birth/ [accessed 27/07/23]

Beckstead, Nick (2013) *On the overwhelming importance of shaping the far future.* PhD Thesis, Rutgers, New Jersey.

Bekoff, Mark and Pierce, Jessica (2009) *Wild justice: The moral lives of animals.* Chicago: University of Chicago Press.

Benatar, David (2006) *Better never to have been: The harm of coming into existence.* Oxford: Oxford University Press.

Bentham, Jeremy (1789) *An introduction to the principles of morals and legislation.*

Birch, Jonathan (2018) 'Degrees of sentience', *Animal Sentience,* 21(11).

Borchert, Rhys and Dewey Aliya R. (2023) 'In praise of animals', *Biology and Philosophy,* 38(4).

Bostrom, Nick (2013) 'Existential risk prevention as global priority', *Global Policy,* 4(1): 15-31.

Bostrom (2009) 'Pascal's mugging', *Analysis,* 69(3): 443-445.

Bostrom, Nick (2008) 'Letter from utopia', *Journal of Evolution and Technology,* 19(1): 67-72.

Bostrom, Nick (2003) 'Astronomical waste: The opportunity cost of delayed technological development', *Utilitas,* 15(3): 308-314.

Bramble, Ben (2021) 'Painlessly killing predators', *Journal of Applied Philosophy,* 38(2): 217-225.

Brennan, Pat (2022) 'Cosmic milestone: NASA confirms 5,000 exoplanets', NASA Exoplanet Exploration. Available at: https://exoplanets.nasa.gov/news/1702/cosmic-milestone-nasa-confirms-5000-exoplanets/ [accessed 20/07/23]

Browning, Heather, and Veit, Walter (2023) 'Positive wild animal welfare', *Biology and Philosophy,* 38(14).

Bruers, Stijn (2019) 'Wild animal suffering, longtermism and population ethics'. Available at: https://stijnbruers.wordpress.com/2019/09/03/wild-animal-suffering-longtermism-and-population-ethics/ [accessed 19/07/23]

Callicott, J. Baird (1989) *In defense of the land ethic: Essays in environmental philosophy.* Albany: State University of New York Press.

Caviola, Lucius, Everett, Jim A. C., and Faber, Nadira S. (2019) 'The moral standing of animals: Towards a psychology of speciesism', *Journal of Personality and Social Psychology,* 116(6): 1011-1029.

Carbone, Larry (2021) 'Estimating mouse and rat use in American laboratories by extrapolation from Animal Welfare Act-regulated species', *Scientific Reports,* 11(493).

Chalmers, David (2022) *Reality+: Virtual worlds and the problems of philosophy*. Penguin Books.

Conway Morris, Simon (2003) *Life's solution: Inevitable humans in a lonely universe.* Cambridge: Cambridge University Press.

Crick, F.H.C. and Orgel, L.E. (1973) 'Directed panspermia', *Icarus,* 19(3): 341-346.

Crisp, Roger (2022) 'Pessimism about the future', *Midwest Studies in Philosophy,* 46: 373-385.

Crisp, Roger (1997) *Mill on Utilitarianism.* London: Routledge.

Crowl, Adam, Hunt, J., and Hein, Andreas M. (2012) 'Embryo space colonisation to overcome the interstellar time distance bottleneck', *Journal of the British Interplanetary Society* 65(7): 283–285.

Cowen, Tyler (2003) 'Policing nature', *Environmental Ethics,* 25(2): 169-182.

de Lazari-Radek, Katarzyna and Singer, Peter (2014) *The point of view of the universe.* Oxford: Oxford University Press.

Delon, Nicolas and Purves, Duncan (2018) 'Wild animal suffering is intractable', *Journal of Agricultural and Environmental Ethics,* 31(2): 239-260.

Dembitzer, Jacob et al. (2022) 'Levantine overkill: 1.5 million years of hunting down the body size distribution', *Quaternary Science Reviews,* 276.

Duncan, Ian (2006) 'The changing concept of animal sentience', *Applied Animal Behaviour Science,* 100(1-2): 11-19.

Dworkin, Ronald (1993) *Life's dominion: An argument about abortion, euthanasia, and individual freedom.* New York: Knopf.

Eaton, C. D. (2008) 'Coevolutionary research', in Jørgensen, Sven Erik (Ed.), *Encyclopedia of ecology.* Amsterdam: Elsevier.

Eckerström-Liedholm, Simon (2019) *Persistence and reversibility: Long-term design considerations for wild animal welfare interventions.* Wild Animal Initiative. Available at: https://static1.squarespace.com/static/5f04bd57a1c21d767782adb8/t/5f160c91bc0bff4abe964d5a/1595280529848/WAI_PersistenceAndReversibility_Dec2019.pdf [accessed 19/07/23]

Edwards, Matthew R. (2021) 'Android Noahs and embryo Arks: ectogenesis in global catastrophe survival and space colonization', *International Journal of Astrobiology,* 20(2): 150-158.

Elliot, Robert (1982) 'Faking nature' *Inquiry,* 25(1): 81-93.

European Environment Agency (2020) 'Invasive alien species: A growing problem for environment and health'. Available at: https://www.eea.europa.eu/highlights/invasive-alien-species-a-growing [accessed 26/07/23]

European Food Safety Authority (2005) 'Aspects of the biology and welfare of animals used for experimental and other scientific purposes', *ESFA Journal*, 292: 1-46.

Faria, Catia (2023a) *Animal ethics in the wild: Wild animal suffering and intervention in nature.* Cambridge: Cambridge University Press.

Faria, Catia (2023b) 'Vulnerability and the ethics of environmental enhancement', *Ethics, Policy, and Environment,* (Published online 14th April 2023).

Faria, Catia (2016) *Animal ethics goes wild: The problem of wild animal suffering and intervention in nature.* PhD thesis, Universitat Pompeu Fabra, Barcelona.

Faria, Catia and Horta, Oscar (2019) 'Welfare biology' in Fischer, Bob (Ed.), *The Routledge handbook of animal ethics*. New York: Routledge.

Feinberg, Joel (1986) 'Wrongful life and the counterfactual element in harming', *Social Philosophy and Policy*, 4(1): 145-178.

Ferrin, Asia (2019) 'Nonhuman animals are morally responsible', *American Philosophical Quarterly,* 56(2): 135-154.

Fish Count (2019a) 'Numbers of farmed decapod crustaceans'. Available at:

http://fishcount.org.uk/fish-count-estimates-2/numbers-of-farmed-decapod-crustaceans

[accessed 21/07/23]


Fish Count (2019b) 'Numbers of farmed fish slaughtered each year'. Available at:

http://fishcount.org.uk/fish-count-estimates-2/numbers-of-farmed-fish-slaughtered-each-year

[accessed 21/07/23]


Fish Count (2019c) 'Reducing suffering in fisheries'. Available at: http://fishcount.org.uk

[accessed 21/07/23]


Frick, Johann (2017) 'On the survival of humanity', *Canadian Journal of Philosophy,* 47(2-3): 344-367.


Gardner, Molly (2022) 'Suffering and meaning in the lives of wild animals', *Midwest Studies in Philosophy,* 46: 355-371.


Gould, Stephen Jay (1989) *Wonderful life.* New York: Norton.


Greaves, Hilary (2017) 'Population axiology', *Philosophy Compass,* 12(11): e12442.


Greaves, Hilary, and MacAskill, William (2021) 'The case for strong longtermism', GPI working paper available at: https://globalprioritiesinstitute.org/hilary-greaves-william-macaskill-the-case-for-strong-longtermism-2/ [accessed 19/07/23]

Groff, Zach, and Ng, Yew-Kwang (2019) 'Does suffering dominate enjoyment in the animal kingdom? An update to welfare biology', *Biology and Philosophy,* 34(40).

Gros, Claudius (2019) 'Why planetary and exoplanetary protection differ: The case of long duration genesis missions to habitable but sterile M-dwarf oxygen planets', *Acta Astronautica,* 157: 263-267.

Gros, Claudius (2016) 'Developing ecospheres on transiently habitable planets: The genesis project', *Astrophysics and Space Science,* 361(324).

Harman, Elizabeth (2004) 'Can we Harm and benefit in creating?', *Ethics,* 18(1): 89-113.

Hastings, Kenneth L. and Bala, Shukal (2005) 'Animal models of immunodeficiency', in Vohr HW. (Ed.), *Encyclopedic reference of immunotoxicology*. Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-27806-0_69

Hecht, Luke (2020) 'The relevance of trophic interactions to wild animal welfare'. Available at: https://www.wildanimalinitiative.org/blog/trophicinteractions [accessed 24/03/23]

Hills, Alison (2010) 'Utilitarianism, contractualism and demandingness', *The Philosophical Quarterly,* 60(239): 225-242.

Horta, Oscar (2022) *Making a stand for animals.* Abingdon: Routledge.

Horta, Oscar (2017) 'Animal suffering in nature: The case for intervention', *Environmental Ethics,* 39(3): 261-279.

Horta, Oscar (2010) 'Debunking the idyllic view of natural processes: Population dynamics and suffering in the wild', *Telos* 17: 73-88.

Horta, Oscar and Teran, Dayron (2023) 'Reducing wild animal suffering effectively: Why impracticability and normative objections fail against the most promising ways of helping wild animals', *Ethics, Policy and Environment,* (Published online 12th of May 2023).

Huemer, Michael (2008) 'In defence of repugnance', *Mind,* 117(468): 899-933.

International Space Studies University (2020) *Lunar agriculture: Farming for the future.* Available at: https://isulibrary.isunet.edu/doc_num.php?explnum_id=1714 [accessed 19/07/23]

IUCN (2014) *Numbers of threatened species by major groups of organisms (1996–2014).* Available at:

http://cmsdocs.s3.amazonaws.com/summarystats/2014_3_Summary_Stats_Page_Documents/2014_3_RL_Stats_Table_1.pdf [accessed 20/07/23]

Jalagania, Beka (2021) 'Wild animals and duties of assistance', *Journal of Agricultural and Environmental Ethics* 34(9).

Johannsen, Kyle (2021) *Wild animal ethics: The moral and political problem of wild animal suffering.* New York: Routledge.

Johannsen, Kyle (2020) 'To assist or not to assist? Assessing the potential moral costs of humanitarian intervention in nature', *Environmental Values,* 29(1): 29-45.

Johannsen, Kyle (2017) 'Animal rights and the problem of r-strategists', *Ethical Theory and Moral Practice,* 20: 333-345.

Kahane, Guy (2023) 'Was evolution worth it?', *Philosophical Studies,* 180: 249-271.

Kahane, Guy (2019) 'History and persons', *Philosophy and Phenomenological Research,* 99(1): 162-187.

Korsgaard, Christine M. (2018) *Fellow creatures: Our obligations to the other animals.* Oxford: Oxford University Press.

Leopold, Aldo (1949) *A sand county almanac.* New York: Oxford University Press.

Lopez, Jose V., Peixoto, Raquel S., and Rosado, Alexandre S. (2019) 'Inevitable future: Space colonization beyond Earth with microbes first', *FEMS Microbiology Ecology*, 95(10).

MacArthur, Robert and Wilson, Edward O. (1967) *The theory of island biogeography.* Princeton: Princeton University Press.

MacAskill, William (2022) *What we owe the future*. New York: Basic Books.

MacClellan, Joel (2023) 'Is biocentrism dead? Two live problems for life-centered ethics', *The Journal of Value Inquiry,* (Published online 13th of July 2023).

Martin, Paul (1967) 'Prehistoric overkill' in Martin P. S. & Wright H. E. (Eds.), *Pleistocene extinctions: The search for a cause*, pp. 75–120. Yale University Press.

Matheny, Gaverick and Chan, Kai M. A. (2005) 'Human diets and animal welfare: The illogic of the larder', *Journal of Agricultural and Environmental Ethics,* 18: 579-594.

Mautner, Michael (2010) 'Seeding the universe with life: Securing our cosmological future', *Journal of Cosmology,* 5: 982-994.

Mautner, Michael (2009) 'Life-centred ethics, and the human future in space', *Bioethics,* 23(8): 433-440.

Mautner, Michael (2004) *Seeding the universe with life: Securing our cosmological future*. Legacy Books.

McMahan, Jeff (unpublished manuscript) 'Compensation for wrongful life?'. Available at: https://www.law.nyu.edu/sites/default/files/Jeff%20McMahan_Compensation%20for%20Wrongful%20Life_general%20%283%29.pdf [accessed 27/07/23]

McMahan, Jeff (2021) 'Climate change, war, and the non-identity problem', *Journal of Moral Philosophy,* 18*,* 211-238.

McMahan, Jeff (2016) 'The moral problem of predation' in A. Chignell, T. Cuneo & M. Halteman (Eds.), *Philosophy Comes to Dinner.* Oxford: Routledge.

McMahan, Jeff (2013) 'Causing people to exist and saving people's lives', *Journal of Ethics,* Volume 17(1/2): 5-35.

McMahan, Jeff (2010) 'The meat eaters', *The New York Times.* Available at: https://archive.nytimes.com/opinionator.blogs.nytimes.com/2010/09/19/the-meat-eaters/ [accessed 20/07/23]

McMahan, Jeff (2009) 'Asymmetries in the morality of causing people to exist', in Roberts, Melinda and Wasserman, David (Eds.), *Harming future persons: Ethics, genetics, and the Nonidentity Problem.* Dordrecht: Springer.

McGuire, Luke, Palmer, Sally B., & Faber, Nadira S. (2023) 'The development of speciesism: Age-related differences in the moral view of animals', *Social Psychological and Personality Science,* 14(2): 228-237.

McTaggart, John M. E. (1927) *The nature of existence* (volume 2)*.* Cambridge: Cambridge University Press.

Mellor, D.J. and Diesch, T.J. (2007) 'Birth and hatching: Key events in the onset of awareness of the lamb and chick', *New Zealand Veterinary Journal,* 55(2): 51-60.

Milburn, Josh (2022) *Just fodder.* Montreal and Kingston: McGill-Queen's University Press.

Milburn, Josh (2020) 'Animal property rights' in Kocsis, M. (Ed.), *Global Encyclopedia of Territorial Rights*. Springer, Cham. https://doi.org/10.1007/978-3-319-68846-6_72-1

Milburn, Josh (2017) 'Nonhuman animals as property holders: An exploration of the Lockean Labour-Mixing account', *Environmental Values* 26(5): 629-648.

Mill, John Stuart (1864/2014) *Utilitarianism.* Cambridge: Cambridge University Press.

Minassian, Liana (2022) 'Why the global rise in vegan and plant-based eating is no fad. *Food Revolution Network.* Available at: https://foodrevolution.org/blog/vegan-statistics-global/#:~:text=Vegan%20Stats%3A%20Global%20Demand%20Increases%20for%20Plant%2DBased%20Foods&text=This%20means%20that%20plant%2Dbased,to%20increase%20fivefold%20by%202030 [accessed 24/07/23]

Moen, Ole Martin (2016) 'The ethics of wild animal suffering', *Etikk I Praksis - Nordic Journal of Applied Ethics* 1(1): 91-104.

Mogensen, Andreas L. (2022) 'The only ethical argument for positive δ? Partiality and pure time preference', *Philosophical Studies,* 179: 2731-2750.

Moore, Lauren et al. (2023) 'Demographic effects of road mortality on mammalian populations: a systematic review', *Biological Reviews* (early view). Available at: https://onlinelibrary.wiley.com/doi/10.1111/brv.12942 [accessed 25/03/23]

Myrsklä, Mikko, Kohler, Hans-Peter and Billari, Francesco C. (2009) 'Advances in development reverse fertility declines', *Nature,* 460(7526): 741-743.

Nebel, Jacob M. (2022)  'Totalism without repugnance', in  J. McMahan,  T. Campbell,  J. Goodrich, &  K. Ramakrishnan (Eds.),  *Ethics and existence: The legacy of Derek Parfit*. Oxford: Oxford University Press.

Newberry, Toby (2021) 'How many lives does the future hold?', GPI working paper available at: https://globalprioritiesinstitute.org/how-many-lives-does-the-future-hold-toby-newberry-future-of-humanity-institute-university-of-oxford/ [accessed 19/07/23]

Ng, Yew-Kwang (1995) 'Towards welfare biology: Evolutionary economics of animal consciousness and suffering', *Biology and Philosophy* 10: 255-285.

Nooijer, Silvester, Holland, Barbara R. and Penny, David (2009) 'The emergence of predators in early life: There was no Garden of Eden', *PLoS One,* 4(6): e5507.

Nussbaum, Martha (2006) *Frontiers of Justice,* Harvard University Press.

O'Brien, Gary David (2023) 'On our moral entanglements with wild animals', *Food Ethics,* 8(15).

O'Brien, Gary David (2022a) 'Beneficence, non-identity, and responsibility: How identity-affecting interventions in nature can generate secondary moral duties', *Philosophia* 50: 887–898.

O'Brien, Gary David (2022b) 'Directed panspermia, wild animal suffering, and the ethics of world-creation', *Journal of Applied Philosophy,* 39(1): 87-102.

Ord, Toby (2020) *The precipice.* New York: Hachette Books.

Palmer, Clare (2021) 'Assisting wild animals vulnerable to climate change: Why ethical strategies diverge', *Journal of Applied Philosophy* 38(2): 179-195.

Palmer, Clare (2010) *Animal ethics in context.* New York: Columbia University Press.

Parfit, Derek (2017) 'Future people, the Non-Identity Problem, and person-affecting principles', *Philosophy and Public Affairs,* 45(2): 118-157.

Parfit, Derek (2016) 'Can we avoid the Repugnant Conclusion?', *Theoria,* 82(2): 110-127.

Parfit, Derek (2011) *On what matters* (volume 2). Oxford: Oxford University Press.

Parfit, Derek (1986) 'Overpopulation and the quality of life', in P. Singer, (Ed.), *Applied ethics*, Oxford: Oxford University Press, 145–164.

Parfit, Derek (1984) *Reasons and persons.* Oxford: Oxford University Press.

Pearce, David (2009) 'Reprogramming predators'. Available at:

https://www.hedweb.com/abolitionist-project/reprogramming-predators.html [accessed 20/07/23]

Pearce, David (1995) *The Hedonistic Imperative.* Available at: https://www.hedweb.com [accessed 20/07/23]

Pepper, Angie (2019) 'Adapting to climate change: What we owe to other animals', *Journal of Applied Philosophy* 36(4): 592-607.

Pianka, Eric (1970) 'On r- and K-selection', *The American Naturalist*, 104: 592-597.

Pinker, Steven (2011) *The better angels of our nature: A history of violence and humanity.* London: Penguin.

Plant, Michael (2022) 'The meat eater problem', *Journal of Controversial Ideas,* 2(2).

Plant, Michael (2016) 'The unproven (and unprovable) case for wild animal suffering'.

Available at: https://www.plantinghappiness.co.uk/the-unproven-and-unprovable-case-for-wild-animal-suffering/ [accessed 24/07/23]

Przybyla, Cyrille (2021) 'Space aquaculture: Prospects for raising aquatic vertebrates in a bioregenerative life-support system on a lunar base', *Frontiers in Astronomy and Space Sciences,* 8.

Purves, Duncan and Hale, Benjamin (2016) 'Non-identity for non-humans', *Ethical Theory and Moral Practice,* 19: 1165-1185.

Regan, Tom (2004) *The case for animal rights.* Berkeley: University of California Press.

Rethink Priorities (2022) *The welfare range table.* Available at: https://rethinkpriorities.org/publications/the-welfare-range-table [accessed 21/07/23]

Reznick, David, Bryant, Michael J., and Bashey, Farrah (2002) 'r- and K-selection revisited: The role of population regulation in life-history evolution', *Ecology,* 83(6): 1509-1520.

Roberts, Melinda A. (2022) 'The nonidentity problem', in Zalta, Edward N. and Nodelman Uri (Eds.), *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition). Available at: https://plato.stanford.edu/entries/nonidentity-problem/ [accessed 26/07/23]

Ross, W. D. (1930/2002) *The right and the good.* Oxford: Oxford University Press.

Rowe, Abraham (2020) 'Should longtermists mostly think about animals?'. Available at: https://forum.effectivealtruism.org/posts/W5AGTHm4pTd6TeEP3/should-longtermists-mostly-think-about-

animals#Could_human_welfare_dominate_animal_welfare_in_the_future_ [accessed 19/07/23]

Sanders, Bas (2018) 'Global animal slaughter statistics and charts', *Faunalytics.* Available at: https://faunalytics.org/global-animal-slaughter-statistics-and-charts/ [accessed 21/07/23]

Schneider, Michael (2001) 'Habitat loss, fragmentation and predator impact: spatial implications for prey conservation', *Journal of Applied Ecology* 38(4): 720-735.

Schopenhauer, Arthur (1851/2016) 'On the sufferings of the world', in Benatar, David (Ed.), *Life, death, and meaning.* Lanham, Md: Rowman and Littlefield.

Sebo, Jeff (2023a) 'The Rebugnant Conclusion: Utilitarianism, insects, microbes, and AI systems', *Ethics, Policy and Environment,* (Published online 26th April 2023).

Sebo, Jeff (2023b) 'Utilitarianism and nonhuman animals', in R.Y. Chappell, D. Meissner, and W. MacAskill (eds.), *An Introduction to Utilitarianism*. Available at: https://utilitarianism.net/guest-essays/utilitarianism-and-nonhuman-animals/ [accessed 7/24/2023]

Shiffrin, Seana (1999) 'Wrongful life, procreative responsibility, and the significance of harm', *Legal Theory,* 5(2): 117-148.

Shklovskii, Iosef and Sagan, Carl (1966) *Intelligent life in the universe*. New York: Delta.

Sidgwick, Henry (1907) *The methods of ethics.* London: MacMillan and Co.

Singer, Peter (2023) *Animal liberation now.* London: The Bodley Head.

Singer, Peter (1975) *Animal liberation.* New York: Avon Books.

Singer, Peter (1973) 'Food for thought', *The New York Review of Books.* Available at: https://www.nybooks.com/articles/1973/06/14/food-for-thought/ [accessed 21/07/23]

Sivula, Oskari (2022) 'The cosmic significance of directed panspermia: Should humanity spread life to other solar systems?', *Utilitas,* 34(2): 178-194.

Smilansky, Saul (2013) 'Morally, should we prefer never to have existed?', *Australasian Journal of Philosophy,* 91(4): 655-666.

Soryl, Asher A., Moore, Andrew J., Seddon, Philip J., and King, Mike R. (2021) 'The case for welfare biology', *Journal of Agricultural and Environmental Ethics,* 34(7).

Tännsjö, Torbjörn (2002) 'Why we ought to accept the Repugnant Conclusion', *Utilitas,* 14(3): 339-359.

Tarsney, Christian (2022) 'The epistemic challenge to longtermism', GPI working paper available at: https://globalprioritiesinstitute.org/christian-tarsney-the-epistemic-challenge-to-longtermism/ [accessed 19/07/23]

Taylor, Paul (1986) *Respect for nature: A theory of environmental ethics*. New York: Princeton University Press.

Thornton, Alex (2019) 'This is how many animals we eat each year', *World Economic Forum.* Available at: https://www.weforum.org/agenda/2019/02/chart-of-the-day-this-is-how-many-animals-we-eat-each-year/ [accessed 21/07/23]

Tomasik, Brian (2016) 'Habitat loss, not preservation, generally reduces wild-animal suffering'. Available at: https://reducing-suffering.org/habitat-loss-not-preservation-generally-reduces-wild-animal-suffering/ [accessed 21/07/23]

Tomasik, Brian (2014) 'Will space colonization multiply wild-animal suffering?'. Available at: https://reducing-suffering.org/will-space-colonization-multiply-wild-animal-suffering/ [accessed 19/07/23]

Tomasik, Brian (2013) 'Applied welfare biology and why wild-animal advocates should focus on not spreading nature'. Available at: https://reducing-suffering.org/applied-welfare-biology-wild-animal-advocates-focus-spreading-nature/ [accessed 24/03/23]

Tomasik, Brian (2009a) 'How many wild animals are there?'. Available at: https://reducing-suffering.org/how-many-wild-animals-are-there/ [accessed 19/07/23]

Tomasik, Brian (2009b) 'The importance of wild-animal suffering'. Available at: https://longtermrisk.org/the-importance-of-wild-animal-suffering/ [accessed 19/07/23]

Torres, Phil (2018) 'Space colonization and suffering risks: Reassessing the "maxipok rule"', *Futures,* 100: 74-85.

UNEP (2012) *One planet, how many people? A review of Earth's carrying capacity.* Available at: https://na.unep.net/geas/archive/pdfs/geas_jun_12_carrying_capacity.pdf [accessed 21/07/23]

Unruh, Charlotte Franziska (2021) 'The strings attached to bringing future generations into existence', *Journal of Applied Philosophy,* 38(5): 857-869.

Vallentyne, Peter (2005) 'Of mice and men: Equality and animals', *The Journal of Ethics,* 9(3/4): 403-433.

Velleman, David (2008) 'The Gift of Life', *Philosophy and Public Affairs,* 36(3): 245-266.

Virk, Kameron (2019) 'Tardigrades: 'Water bears' stuck on the moon after crash', *BBC News.* Available at: https://www.bbc.co.uk/news/newsbeat-49265125 [accessed 20/07/23]

Whittlestone, Jess (2017) 'The long-term future'. Available at: https://www.effectivealtruism.org/articles/cause-profile-long-run-future [accessed 19/07/23]

Wilks, Matti, Caviola, Lucius, Kahane, Guy, and Bloom, Paul (2021) 'Children prioritize humans over animals less than adults do', *Psychological Science*, 32(1): 27-38.

Williams, George C. (1957) 'Pleiotropy, natural selection, and the evolution of senescence', *Evolution*, 11(4): 398-411.

Williamson, Patrick (2021) 'A new argument against Critical-Level Utilitarianism', *Utilitas,* 33: 399-416.

Woods Hole Oceanographic Institution (2019) 'Bristlemouth dominance: How do we know?'. Available at https://twilightzone.whoi.edu/bristlemouth-dominance-how-do-we-know/ [accessed 19/07/23]

Zapffe, Peter Wessel (1933) 'The last messiah'. Available at: https://philosophynow.org/issues/45/The_Last_Messiah [accessed 24/07/23]

Zuber, Stéphane et al. (2021) 'What should we agree on about the Repugnant Conclusion?', *Utilitas,* 33: 379-383.