Convergence and Value in Environmental and Animal Ethics

Yunjie Zhang

Submitted in fulfilment of the requirements for the Degree of
Doctor of Philosophy

Philosophy
School of Humanities
College of Arts
University of Glasgow

November 2021

Abstract

This thesis concerns the plausibility of arguments that certain normative ethical theories converge on the answer to the practical question of normative ethics (e.g., What (morally) ought I to do?) and the explanatory question of normative ethics (e.g., What explains why certain actions are permissible, required, or forbidden?). I explore different strategies for arguing that normative ethical theories converge on the answer to these questions. I then turn to the prospects of convergence between certain normative ethical theories in environmental ethics and animal ethics. In environmental ethics, I consider the prospects of convergence between Anthropocentrism (i.e., human-centred ethics) and Non-Anthropocentrism. I argue that Bryan Norton's influential convergence argument fails to show that there is convergence between these two views. I then propose my own version of Anthropocentrism (i.e., "Broad Anthropocentrism") and argue that it exhibits some degree of convergence with Non-Anthropocentrism on the answer to both the practical and the explanatory question of normative ethics. In animal ethics, I consider the prospects of convergence between Utilitarianism and a moral rights-based approach. I argue that a prominent attempt to argue for convergence between these theories fails because it relies on implausible or incomplete versions of a moral rights theory. Finally, I shift focus from convergence on the answer to the practical and explanatory questions to convergence on the question of whether one ought to care about nature and animals. I argue that there is a large amount of convergence between Broad Anthropocentrism and Non-Anthropocentrism on the answer to the questions of "Should one finally care about nature?" and "Why should one finally care about nature?" I also argue that there is a fair amount of convergence between Utilitarianism and a plausible moral rights view on the answer to the questions, "Should I care about animals?" and "Why should I care about animals?"

Table of Contents

Acknowledgements

I would like to first thank my parents, 陈素 and 张建忠, for their ultimate support and love. Without their support and love, I couldn't have finished my thesis or even started my PhD studies.

I want to express my appreciation to my supervisors, Michael Brady and Robert Cowan, for agreeing to supervise my thesis, for their detailed criticism of my work and for their academic guidance. I would like to thank the faculty and my colleagues in the philosophy department, especially the ones who came to my seminar talks, gave me presentation suggestions, peer-reviewed my writings and talked to me in the hallway or common room. Your kindness and thoughtfulness really helped me feel comfortable and like I belonged.

I thank Professor Katie McShane for agreeing to reply to my criticism of her paper and for submitting a discussion paper with me to *Environmental Values*. Also, thanks to *Tsinghua Studies in Western Philosophy* for publishing one section of my thesis in their 2021 winter volume. I presented parts of my thesis in Stirling, Barcelona, Vienna, Ireland, Oslo and on Zoom. I am grateful to the members of those audiences for their valuable feedback. I also want to thank the Royal Institute of Philosophy for offering the 10-week course by Professor Nadine Elzein on all varieties of interesting philosophy topics. Next, I want to thank the English Writing Centre of Glasgow University for solving my writing problems and offering me English writing classes. Additionally, I thank my proofreader Andrew for helping me polish my thesis.

I feel exceptionally fortunate to have my friends, Manon Haag, Kiko, Seongmee Yun, Alicia Gardner, Amy Zhou and Jana. Without sharing good moments and sad moments with

you, I cannot imagine how I could've survived studying abroad. I finally want to thank my partner, Max. Meeting you was one of the greatest things that has happened to me.

Introduction

Broadly speaking, this thesis concerns convergence between normative ethical theories. More concretely, it concerns what it means for normative ethical theories to converge and whether certain normative ethical theories in environmental ethics and animal ethics converge. While the issue of convergence between normative ethical theories has generally been overlooked by moral philosophers in general, it has been debated in both environmental ethics and animal ethics.

My plan in this introduction is to briefly introduce the core themes and aims of my thesis, provide a terminological guide, and to outline the individual chapters.

1. Core Themes and Aims

The fact that convergence between normative ethical theories has not received much attention (even in environmental ethics and animal ethics) is somewhat surprising. This is because if one could show that there is a significant amount of convergence between normative ethical theories, one could allay a few important worries that moral philosophers have arising from moral disagreement.

First, as Derek Parfit notes, divergence between normative ethical theories seems to cast doubt on moral facts and so, if one could show that there is significant convergence between these theories, one could allay such a worry. He writes:

> Of our reasons for doubting that there are moral truths, one of the strongest is provided by some kinds of moral disagreement. Most moral disagreements do not count strongly against the belief that there are moral truths, since these disagreements depend on different people's having conflicting empirical or religious beliefs, or on their having conflicting interests, or on their using different concepts, or these disagreements are about borderline cases, or they depend on the false assumption that all questions must have answers, or precise answers. But some disagreements are not of these kinds. These disagreements are deepest when we are considering, not the wrongness of particular acts, but the nature of morality and moral reasoning, and what is

> implied by different views about these questions. If we and others hold conflicting views, and we have no reason to believe that we are the people who are more likely to be right, that should at least make us doubt our view. It may also give us reasons to doubt that any of us could be right. [1]

However, if we could show that normative ethical theories actually agree on a lot of important issues, we could show that there is less moral disagreement than previously thought and we will have less reason to doubt that there are moral truths.

A second reason to explore whether there is a significant amount of convergence between normative ethical theories is to defend against scepticism about moral knowledge or justified moral beliefs. The worry is that disagreement between moral philosophers gives us reason to doubt our own moral beliefs and thus makes it harder for our moral beliefs to be justified or count as knowledge. This is in part because knowing that we disagree with at least some philosophers and they are more likely to have true moral beliefs than us, gives us reason to doubt our own moral beliefs. But, if there are moral facts, as most people believe there are, then it will be important to know or have justified beliefs about what they are.

Given that whether normative ethical theories exhibit some significant amount of convergence is important for defending again certain kinds of scepticisms, it is also important to be clear about what convergence consists in and how it is possible. Therefore, one of the aims of this thesis is to get clear on what convergence between normative ethical theories involves and in what ways might it come about.

Another central aim of this thesis is to argue that some amount of convergence between normative ethical theories in environmental ethics and animals is possible. More specifically, I will argue that there is some convergence between seemingly opposing views in environmental ethics (Anthropocentrism and Non-Anthropocentrism) concerning when and why to treat nature well. I also argue that there is a surprising degree of convergence in

---

[1]Parfit (2011): 418-419.

both environmental ethics and animal ethics concerning when and why we should *care* about nature and animals. One of the important upshots of my arguments is that it may be easier to get agreement on environmental policies by showing opposing sides how there is a fair amount of agreement between plausible versions of their preferred views.

Finding convergence between opposing views in environmental ethics is more important than ever because of the imminent threat posed by climate change. The threat is not only to human and animal life but also other to parts of nature, e.g., ecosystems. If policymakers could more quickly and easily agree on environmental policies, then climate change could be combatted sooner. If it could be shown, and I think it can, that the major ethical theories concerning nature exhibit a high degree of convergence on how and why to protect nature, then policymakers should be more convinced to support policies focused on protecting nature. Policies that protect nature from the effects of climate change are important both because they ultimately benefit humans and non-human animals, but also because they benefit other parts of nature. In Chapters 3 and 5, I argue for a degree of convergence between the aforementioned opposing views. In particular, I argue that both views can agree that nature has an important kind of value (i.e., final value) which justifies protecting it.

Moreover, it is also vitally important to consider the morality of animal experimentation. This is because non-human animals (especially mammals) are routinely used in the development of coronavirus vaccines. Given that the coronavirus does not seem to be going away anytime soon, we can expect that there will continue to be more and more vaccines developed to fight the growing number of variants. This means that there will be an increasing number of experiments on mammals.[2] This means that it is urgent that we get as plausible a view as possible about how much and what kind of animal experimentations are

[2] EFPIA (2021).

morally permissible. I explore this issue directly in Chapter 4 when I consider whether a prominent version of Utilitarianism exhibit any meaningful convergence with a deontological view prominent in animal ethics (i.e., the Moral Rights View).

2. Terminology

2.1 Nature

Throughout this thesis, I will discuss "nature" and what prominent normative ethical views in environmental ethics entail about how we should treat nature. This is important to be clear about what I mean by "nature". I will not offer a precise account of what I mean by nature, but rather propose a disjunction of related entities that might plausibly be thought to be picked out by talk of "nature". My reason for this is twofold. First, environmental ethicists themselves are often imprecise about exactly what they have in mind what they talk about nature. Second, my arguments will be interesting if they work for any of the entities that I will list as candidates for the meaning of "nature".

By "nature", I will primarily have in mind the disjunction of the following: (1) ecosystems, (2) sets of ecosystems, (3) some set of the components of ecosystems, or (4) the biosphere.[3] As the concept of an ecosystem figures heavily in this understanding, it will be helpful to say a few things about it. Here are a few helpful definitions of ecosystems:

> An ecosystem is an interaction-structure of organisms and their inorganic environment, which is open and, to a certain degree, capable of self-regulation.[4]

> An ecosystem consists of living organisms in some abiotic environment. What makes it a system is the fact that there exist specific dynamic relationships between these constituents. What makes it cybernetic is the existence of coordination, regulation, communication, and control in these relationships.[5]

---

[3] For a similar approach, see Samuelsson (2008): 17.

[4] Klötzli (1993): 288.

[5] McNaughton and Coughenour (1981): 985.

> [An ecosystem is a] functional unit consisting of all the living organisms (plants, animals, and microbes) in a given area, and all the non-living physical and chemical factors of their environment, linked together through nutrient cycling and energy flow.[6]
>
> [An ecosystem is] a whole whose parts include all living and nonliving processes or objects (slow processes), and their associated biogeo- and physico-chemical, energetic, material, and informational parameters within a region of time and space: together with portions of the surroundings of these units.[7]

We can see that the core constituents of an ecosystem are the biotic (i.e., living) components (e.g., humans, plants, animals, microbes, etc.), the abiotic (i.e., non-living) components (e.g., water, air, soil, climate, sunlight, rocks, etc.) and the various interactions between them (e.g., processes of transferring energy, material, and information between various components).[8] Examples of ecosystems include: forests, wetlands, grasslands, lakes, and coral reefs.[9] Many ecosystems interact with other ecosystems and some ecosystems are *within* larger ecosystems.[10]

The second disjunct is some set of ecosystems, but not the complete set of them. The third disjunct for "nature" is a set of components of ecosystems, e.g., water, air, soil, etc. The fourth disjunct is "biosphere" or "ecosphere" which refers to the sum of all ecosystems. Jørgensen et al. claim that the set of all ecosystems "comprise all nature".[11]

2.2 Animals

When I speak of animals throughout this thesis, I will primarily have in mind only a small subset of non-human animals. In particular, I will focus on mammals that are over a year

---

[6] Gregorich et al. (2001): 112.

[7] Jørgensen et al. (1992): 5.

[8] Peacock (2008): 351.

[9] Harris (2018).

[10] Peacock (2008): 354.

[11] Jørgensen et al (1992): 5.

old.[12] The reason to focus on mammals over a year old is that all the main authors involved in this debate agree that these animals have certain properties or abilities that give them a certain moral status. For Utilitarians, what matters about these animals is either that they are capable of experiencing pleasure and pain or that they are capable of having desires that can be satisfied or frustrated. It seems obvious that animals other than mammals can experience pleasure and pain (e.g., birds).

However, the opposing normative ethical view, i.e., the Moral Rights View, requires more and more sophisticated abilities, which only humans and mammals over one-year-old have. These properties make these mammals what Regan (2004) calls "a subject of a life." Regan writes that

> To be the subject-of-a-life… involves more than merely being alive and more than merely being conscious… [I]ndividuals are subjects-of-a-life if they have beliefs and desires; perception, memory, and a sense of the future, including their own future; an emotional life together with feelings of pleasure and pain; preference- and welfare-interests; the ability to initiate action in pursuit of their desires and goals; a psychophysical identity over time; and an individual welfare in the sense that their experiential life fares well or ill for them.[13]

Thus, the reason to focus on mammals over the age of one year is that they have the abovementioned mental states (e.g., beliefs, desires, emotions, etc.), they have an emotional life, a persistent identity, and so on.

While it might be possible that other animals (including younger mammals) have the kinds of mental states and abilities that make them a subject of a life, this claim is much more controversial. Thus, in order to minimise making controversial assumptions about animal biology and psychology, I will focus on this small set of animals.

---

[12] Regan (2004): 78.

[13] Regan (2004): 243.

3. An Outline of the Thesis

The main purpose of Chapter 1 is to get clear on exactly what kinds of convergence is possible between two or more normative ethical theories, what is involved in these different kinds of convergence, and how these kinds of convergence are possible. Being clear about these issues is essential for the rest of the thesis where I consider particular arguments for why certain normative ethical theories in environmental ethics and animal ethics converge. A secondary purpose of Chapter 1 is to assess some of the strengths and weaknesses of the possible ways two or more normative ethical theories could converge. This will help us assess the plausibility of convergence arguments in Chapters 3-5.

The main purpose of Chapter 2 is to provide a detailed discussion of the different ways that something can be valuable (e.g., finally, instrumentally, extrinsically, intrinsically, inherently, and so on) and how they relate to each other. Being clear on exactly what these different kinds of values consist in and how they relate is essential for Chapters 3-5. This is because those chapters concern all involve normative ethical theories that make claims about what kind of value nature and non-human animals have. A second aim is to argue that the debate in environmental ethics between Anthropocentrism and Non-Anthropocentrism and the debate in animal ethics between Utilitarianism and the Moral Rights View are fundamentally debates about what has *final* value. However, these debates are mistakenly cast as being about intrinsic value. I show this by providing evidence that the authors in these debates are not concerned with intrinsic value as such, but only with final value. This difference is also important for Chapters 3 and 5. In both chapters, I argue that while Broad Anthropocentrism and Non-Anthropocentrism disagree about whether nature has intrinsic value, this disagreement is not important. This is because they agree that nature is finally valuable and what they really care about is final value.

The main purpose of Chapter 3 is to assess the prospects for convergence between two prominent normative ethical theories in environmental ethics (i.e., Anthropocentrism and Non-Anthropocentrism) on when and why to treat nature well. I do this in two parts. First, I criticise Bryan Norton's influential Convergence Hypothesis according to which his preferred version of Anthropocentrism (i.e., Weak Anthropocentrism) exhibits a large amount of convergence with Non-Anthropocentrism. Second, I propose my own alternative version of Anthropocentrism (i.e., Broad Anthropocentrism) and argue that it exhibits some degree of convergence on when to treat nature well and a large amount of convergence on one explanation of why one should treat nature well.

Chapter 4 assesses the prospects for convergence between two prominent normative ethical theories in animal ethics (i.e., Preference Utilitarianism and the Moral Rights View) concerning the morality of animal experimentation. I do this by critically examining Gary Varner's view that the Moral Rights View will exhibit at least some convergence with Preference Utilitarianism on the moral permissibility of animal experimentations that involve harming or killing non-human animals. I argue that Varner's argument relies on an incomplete version of the Moral Rights View and that a complete version would include principles that prevent the Moral Rights View from converging with Preference Utilitarianism.

In Chapter 5, I argue that there is a fair amount of convergence between the abovementioned views in environmental ethics and animal ethics on the questions of whether one should care about nature and animals and why one should care about nature and animals. In particular, I argue that Broad Anthropocentrism and Non-Anthropocentrism exhibit a large amount of convergence on the question of whether and why one should care about nature. In particular, they agree that one should care about nature for its own sake in virtue of agreeing that nature has value in its own right. I also argue that two versions of Utilitarianism (i.e., Hedonistic Utilitarianism and Preference Utilitarianism) and the Moral Rights View exhibit a

high degree of convergence on the question of whether and why one should care about animals. However, they disagree about how to care about nature (i.e., for its own sake or for the sake of something else).

Finally, in the Appendix, I argue that recent attempts to justify treating nature well based on our supposed relationship with nature do not work. First, I argue that we do not have the kind the personal relationships that the authors claim we have. Second, I argue that the relationships they appeal are not universal, i.e., many people don't have these relationships with nature. Third, I argue that these views are disguised versions of Anthropocentrism. This means that appeals to the claim that nature is instrumentally valuable to justify treating nature well. In particular, they appeal to the claim that having a certain relationship with nature can be conducive to living a good life.

Chapter 1: Normative Ethics and Convergence

As discussed in the introduction, this thesis concerns two debates about whether normative ethical principles can converge. The first debate concerns whether certain normative ethical principles concerning issues in environmental ethics can converge and the second debate concerns whether certain normative ethical principles concerning issues in animal ethics (i.e., animal experimentation) can converge. Given the focus of this thesis, it is important to have a clear understanding of what normative ethical principles are, what it means for them to converge, and how it is possible for them to converge. Thus, the focus of this chapter will be to clarify these core issues so that we can make progress on the aforementioned debates about convergence in environmental ethics and in animals ethics.

The plan of the chapter is this. First, I will explain what normative ethical theories are. Second, I will explain what it means for normative ethical theories to converge. Third, I will explain how normative ethical theories can converge.

1. What Are Normative Ethical Theories?

Normative ethics is primarily concerned with providing standards or principles of right action, i.e., what we morally ought (and ought not) to do.[14] Normative ethical theories answer two questions. First, they answer a *practical* question, i.e., "Which actions are right and which actions are wrong?" Second, they answer an *explanatory* question, i.e., "In virtue of

---

[14] A kind of normative ethical theory, i.e., virtue ethics, focuses on what makes someone a good person as opposed to what makes an action right or wrong. However, because I will only be concerned with normative ethical theories that focus on action in this thesis, I will not discuss virtue ethics in any detail. In addition, virtue ethics will not neatly fit into the taxonomy of normative ethical theories that I provide below. There are a small group of virtue ethicists who argue that whether an action is right or wrong is determined by an agent's motivations or what a virtuous person would characteristically do in a given situation. For example, see Hursthouse (1999) and Slote (2001). However, because these theories are quite contentious and not related to the normative ethical theories that I will discuss in the rest thesis, I will leave them out of this chapter. A related question is what makes something a good outcome or a good state of affairs. While this question is important for consequentialist normative ethical theories it is only important insofar as it can help answer the question of what one morally ought to do.

what are right actions right and wrong actions wrong?" In other words, they tell us what properties of actions are right-making and which properties of actions are wrong-making. That is, they tell us which properties are *morally-relevant* properties. This means that normative ethical theories answer the practical question (i.e., what morally ought I to do?) by answering the explanatory question, i.e., they tell us what actions are right or wrong by telling us what properties make actions right or wrong. Thus, normative ethical theories answer the practical question because they offer a certain answer to a moral metaphysical question, i.e., what properties make an action right or wrong? This is a moral metaphysical question because it concerns a metaphysical relationship between non-moral properties and moral properties. In particular, it is a question about what non-moral properties "ground" or "generate" moral properties.

How should we distinguish between different kinds of normative ethical theories? We can roughly differentiate normal ethical theories distinguishing which aspects of a person's action are most important for determining the rightness or wrongness of that action.[15] These aspects are:

(i) The result or outcome of the action. Julia Driver refers to theories that focus on the result or outcome of an action "teleological," by which she means that these such accounts define moral properties in terms of achieving some goal or aim.[16] Teleological theories

[15] This taxonomy is endorsed by Brady (2015). For a similar, but distinct, taxonomy, see: Baron, Pettit, and Slote (1997), 1. In saying that these are the "most important" aspects for determining the rightness or wrongness of an action, I do not mean to imply that these aspects are the only aspects that are morally relevant. Theories will differ on what else matters for determining the rightness or wrongness of an action.

[16] Driver (2005).

are also often referred to as "consequentialist" because the consequences are what matter most, morally speaking.[17]

(ii) The agent's mental states, e.g., the agent's motive (or reason) for acting or their intention in acting. This kind of theory is non-teleological because there is nothing beyond the action and the mental states involved in its production that affects the morality of the action. As I will explain further below, these kinds of theories are often referred to as "non-consequentialist" or "deontological" theories.

(iii) The act itself or the features that the action has are independent of the agent's mental states. These kinds of theories are also non-teleological because all that matters for the morality of an action is the action itself and its properties. These kinds of theories are also often referred to as "non-consequentialist" or "deontological" theories.

Now that I have outlined in broad form how different normative ethical theories can be distinguished, I want to provide more detailed characterisations of the most prominent views.

In order to elucidate the difference between the main normative ethical theories, I will focus on one single case. Consider the following case from Thomson (1976):

> **Organ Transplant:** David is a great transplant surgeon. Five of his patients need new parts—one needs a heart, the other need, respectively, live, stomach, spleen, and spinal cord—but all are of the same, relatively rare, blood-type. By chance, David learns of a healthy specimen with that very blood-type. David can take the healthy specimen's parts, killing him, and install them in his patients, saving them. Or he can refrain from taking the healthy specimen's parts, letting his patients die.[18]

What should David do? The answer depends on which normative ethical theory is correct. As we will see, different normative ethical theories will disagree both on what David is morally

---

[17] One exception to this is Adams (1976) who argues for "*motive* Utilitarianism." This would seem to put his consequentialist view in category 2. However, Adam is arguing about the goodness or badness of *motives* and not about the rightness or wrongness of *actions*. So, his view is not importantly different than the other normative ethical views that I consider in this chapter and the rest of the thesis.

[18] Thomson (1976): 206.

required to do and on why he should act as he is required to act. In other words, seeing how these theories disagree about Organ Transplant will help make it clear how difficult it can be to argue for certain kinds of convergence (more on this in section 3).

In what follows, I will outline the core features of different normative theories. It will be important to make the basic elements of these theories salient to the reader for multiple reasons. First, it will be important for understanding recent convergence arguments. Second, it will be important for understanding the full range of possible ways of arguing that two normative ethical theories converge. Third, it will help the reader see how difficult it will be to argue for certain kinds of convergence (e.g., what I will refer to as mass or total fundamental explanatory convergence). Fourth, Chapter 4 will directly engage with arguments concerning the possibility of converging some version of consequentialism (i.e., Utilitarianism) with a deontological view and so having the features of these views in mind will help make the arguments in that chapter clearer.

1.1 Consequentialism

Consequentialism is a group of normative ethical theories that fall into what I previously referred to as teleological views, i.e., the first category in the above taxonomy. These theories argue that whether an act is morally right depends only on the act's consequences. Or, as Derek Parfit categorises these views, "all that ultimately matters is how well things go."[19]

There are many different consequentialist views, but the most historically prominent form of consequentialism is Utilitarianism. John Stuart Mill, one of the most prominent earlier developers of Utilitarianism, writes:

> The creed which accepts as the foundation of morals, Utility, or the Greatest Happiness Principle, holds that actions are right in proportion as they tend to promote happiness, wrong as they tend to produce the reverse of happiness. By happiness is intended pleasure, and the absence of pain; by unhappiness, pain, and the privation of pleasure. To give a clear view of the moral standard set up by the theory, much more

[19] Parfit (2011): 417.

> requires to be said; in particular, what things it includes in the ideas of pain and pleasure; and to what extent this is left an open question. But these supplementary explanations do not affect the theory of life on which this theory of morality is grounded – namely, that pleasure, and freedom from pain, are the only things desirable as ends; and that all desirable things (which are as numerous in the utilitarian as in any other scheme) are desirable either for the pleasure inherent in themselves, or as means to the promotion of pleasure and the prevention of pain.[20]

While Mill is not explicit about in this quotation, he has in mind *overall* or *general* pleasure or pain, i.e., the pleasure or pain of all people.[21] Thus, according to Utilitarianism, an action is morally right just in case it maximises overall pleasure/happiness or it *tends* to maximise overall pleasure/happiness.

Recall the Organ Transplant case from earlier. If we know that all the patients involved are equal with regard to the quality of the life they will live, then Utilitarianism would require David to kill the innocent, healthy person and transplant his body parts into his five patients. This is because Utilitarianism requires that agents maximise overall utility (i.e., happiness or pleasure) and the only way to do this in the transplant case is to save the lives of five people at the expense of another person's life.[22]

1.2 Deontology

Unlike consequentialism, deontology doesn't hold that moral rightness is about the outcome of an action, whether the outcome is pleasure, happiness or maximised goodness. In section 1, I noted that deontological views focus on the mental states of agents and/or the actions themselves in order to determine the rightness or wrongness of an action. In this section, I will explain some of the core features of deontology and elaborate on how this group of views differs from consequentialism. I will start by outlining the core features of

---

[20] Mill (2003): 186.

[21] Mill (2003): 203.

[22] I'm not claiming that this is the most plausible version of Utilitarianism. My main purpose is rather to show that different kinds of normative ethical theories posit different right- and wrong-making properties and to show these theories yield conflicting verdicts on practical questions.

deontological theories and then I will explain some of the most prominent versions of deontology, i.e., Kantianism, Contractualism, and Rossian Pluralism.

1.2.i The Core Features of Deontology

Deontology is the study of what our moral duties are and thus what we are morally obligated to do. David McNaughton & Piers Rawling (2007) argue that there are three significant moral statuses that deontology concerns.

i) *Constraints*. As McNaughton & Rawling argue, deontological theories are usually concerned with preventing harm to innocent people.[23] In particular, they argue for certain constraints, i.e., duties not to perform certain kinds of actions. For example, deontologists tend to argue that we have duties to not kill, torture, or otherwise harm innocent people. Exactly how strict these duties are is a matter of debate. The main difference with consequentialism is deontological constraints or duties cannot be overridden whenever doing so would lead to some good state of affairs. Exactly how stringent these duties are is a matter of debate. Some philosophers (e.g., Kantians and Catholic theologians and ethicists) think that moral duties are *absolute*, i.e., there are no circumstances in which one can morally permissibly violate a duty.[24] Others argue that one can morally permissibly violate a moral duty if the situation is extreme enough, e.g., if it is necessary to prevent a catastrophe.

ii) *Duties of special relationship*. We have special duties to friends and family. Usually, these special duties require (or at least permit) us to treat our loved ones differently (i.e., better) than we treat strangers. For example, if one is required to save either one's spouse or a complete stranger from drowning, then, when all else is equal, one is required (or at least permitted) to save one's spouse instead of the stranger.

---

[23] McNaughton & Rawling (2007): 425.

[24] Examples of defenders of absolutism include: Aquinas (e.g., *Summa Theologica* II-II, Question 110, Article, 3), Kant (1996): 613 or AK 8:427, Gewirth (1981), Boyle (1991), Finnis (1991), and Anscombe (2008): 67.

iii) *Options*. While consequentialist views, such as Utilitarianism, require that one always maximise utility regardless of the costs to the agent of doing so, deontologists generally agree that there is a limit to what an agent can be morally required to sacrifice just to bring about more goodness (e.g., happiness or pleasure). Therefore, deontologists give more weight to an agent's *own* interests, projects, and wellbeing in determining what that agent is morally required to do. [25]

Following the above taxonomy, we can see that deontological theories see the action itself as being most important to determining the rightness or wrongness of actions. After all, duties are duties to either perform or not perform certain *actions*, they are not duties to bring about certain outcomes.

Returning to the Organ Transplant case, we can see that, in general, deontological theories will likely argue that David is morally required to not kill his innocent patient and transplant their body parts–even though this means allowing five other patients to die. The explanation for this is that David either has an absolute duty to not kill innocent people or has a non-absolute duty to not kill innocent people and the circumstances are not severe enough to override this duty. That is, he has a constraint or duty to not kill.

One influential deontological principle is the Doctrine of Double Effect. While this principle is not meant to be a complete version of deontology, it is endorsed by deontologists in order to explain our intuition that David ought not to kill the one patient in Organ Transplant to save the others. This principle consists of two claims. First, it is wrong to harm an innocent person as a means to benefitting other innocent people if one intends the harm—even if the benefit is proportional to the harm. Second, it is not wrong to harm an innocent person as a means to benefitting other innocent people if one merely foresees (but doesn't

[25] McNaughton and Rawling (2007).

intend) this harm—as long as the benefit is proportional to the harm. This view is a clear case in which an agent's mental states (e.g., intentions) are morally important. The Doctrine of Double Effect offers a clear verdict on Organ Transplant. If David were to choose to save the five, he would have to intend the death of the one healthy patient as a means of benefitting the saved innocents. Therefore, according to the Doctrine of Double Effect, David morally ought not to kill the one to save the five.

1.2.ii Kantianism

Perhaps the most famous deontological theory is due to Immanuel Kant. Kant's theory is rather complex because it fits into a larger philosophical system. However, for our purpose, we need only consider the broad thrust of the view. Very roughly, Kant argued that whether an action is right or wrong is a matter of whether or not that action follows from a maxim or principle (e.g., "Never tell a lie" or "Always keep your promises") and that maxim or principle is such that we could rationally choose that everyone (ourselves includes) follow that maxim or principle. If one could rationally choose that everyone follow a certain maxim and that maxim tells one to perform a certain action, then the action is right or permissible. If one could not rationally choose that everyone follow a certain maxim and that maxim tells one to perform a certain action, then that action is wrong.[26] Because Kant cares most about the maxim or principle on which one acts, his view fits into the second category in our taxonomy. That is, his view cares most about the mental states (i.e., the maxim on which they act) of the agents acting.

Kant formulates his view as follows:

> act only in accordance with that maxim through which you can at the same time will [without contradiction] that it become a universal law.[27]

---

[26] For similar summaries of Kant's view, see: Brady (2015): 18-19 and Korsgaard (1997): xvii-xxi.

[27] AK 4: 421.

He refers to this principle as the Categorical Imperative.[28] This is because it is a moral imperative (i.e., a command) that applies to one categorically (i.e., regardless of one's desires). The basic idea is similar to, but importantly different from, the idea encapsulated in the question, "What if everyone did that?" This question is meant to show that there would be bad *consequences* if everyone behaved as one did. However, being a deontologist, Kant is not concerned with consequences, but with the *maxim* on which one acts. What this question has in common with Kant's view is the idea of universalisability, i.e., the idea that what actions are permissible are those that everyone could perform without there being a kind of conflict.

Imagine that one is deciding whether or not one should make a false promise in order to get a loan (i.e., make a promise without intending to keep it).[29] That is, imagine that one is considering whether to act on the maxim, "Make a false promise if you have to in order to get a loan." Kant asks us to imagine that one is deciding whether to make this false promise in a world in which everyone else who finds themselves in need of a loan makes a false promise to repay it, i.e., everyone follows the above maxim. Kant then notes that it would not be rational to choose that everyone in this world act on that maxim and thus make false promises when they need a loan. Why? Because if everyone did this, then no one would believe the promises of those who are taking out loans. And, if no one believed that people taking out loans would repay them, then no one would give people loans.[30] So, one could not even get a loan in such a world and this would defeat the point of making the promise. Thus, because it would not be rational for one to choose that everyone make false promises in order to get a loan, it is morally wrong to make a false promise in order to get a loan.

---

[28] AK 4:420.

[29] AK 4: 422.

[30] AK 4: 422.

Turning to the case of Organ Transplant, it should be clear that it would be impermissible to kill the one to save the five for Kant. This is because killing one person to save five others is not universalizable in the sense mentioned above. If every doctor were to kill one healthy patient anytime doing so would allow them to save other patients, no one would trust doctors. That is, no one would go to the doctor for a check-up and thus it would not be possible to even kill one to save five.

Thus, for Kant the fundamental morally-relevant property is rational choosability (as a universal law). If some action is rationally choosable as an action that everyone performed, then it is morally permissible. If an action is not rationally choosable in this way, then it is morally wrong.

1.2.iii Contractualism

The final kind of deontological theory I will consider is Contractualism. Contractualism is a set of normative ethical theories that roughly hold that whether an action is right or wrong depends on an actual or hypothetical contract or agreement between people. Contractualism takes two forms along the two different lines of moral thinking started by Thomas Hobbes and Kant, respectively.[31]

The Hobbesian version of this view, contractarianism, holds that principles of normative ethics are created by agreements that rational agents come to when they are each trying to advance their own rational self-interest.[32] Normative ethical principles on this view are the rules that rationally self-interested agents mutually agree to follow as a means of furthering their own rational self-interest.[33] Right actions are those that conform to these agreed-upon principles and wrong actions are those that violate these principles.

---

[31] Hobbes (1668/1994) and Kant (1785/1997). So-called Kantian Contractualism was originally developed by Rawls (1971).

[32] Ashford and Mulgan (2018) and Cudd and Eftekhari (2018).

[33] Ashford and Mulgan (2018) and Cudd and Eftekhari (2018).

Hobbes starts with the idea that there are "laws of nature" that we figure out through our own reasoning. These laws concern our own self-interest or self-preservation. Hobbes writes:

> A law of nature (*lex naturalis)* is a precept or general rule, found out by reason, by which a man is forbidden to do that which is destructive of his life or taketh away the means of preserving the same.[34]

The first fundamental law of nature according to Hobbes is that people seek peace with each other as a means of self-preservation.[35] From the first law of nature, Hobbes argues, follows a second: to give up certain freedoms, on condition that everyone else do the same, for the purpose of peace between people.[36] The third law of nature is that people must follow the agreement that made with each other to give up certain rights or freedoms.[37] Finally, Hobbes argues that to violate an agreement is unjust.[38]

The Kantian version of this view, contractualism, also sees agreement amongst rational agents as being fundamental to normative ethical principles as well. However, on this view, it is not the agreement of rationally *self-interested* agents that matters. Rather, the correct normative ethical principles are those that are agreed on by agents who view themselves and each other as both free and worthy of equal moral respect and regard.[39] The most prominent version of contractualism, due to Scanlon, holds that the correct moral principles are those "principles for the general regulation of behaviour that no one could reasonably reject as a basis for informed, unforced, general agreement."[40] To reasonably

---

[34] Hobbes (1668/1994): Ch. 14, part 3.

[35] Hobbes (1668/1994): Ch. 14, part 4.

[36] Hobbes (1668/1994): Ch. 14, part 5.

[37] Hobbes (1668/1994): Ch. 15, part 1.

[38] Hobbes (1668/1994): Ch. 15, part 2.

[39] Southwood (2010): 6.

[40] Scanlon (1998): 156.

reject a principle is to provide a compelling objection to it, e.g., it fails to treat agents equally. Right actions are those that conform to these principles and wrong actions are those that violate them.

In the case of the "Organ Transplant", what should David do according to these theories? While contractarianism doesn't offer a straight path to answering this question, some think that rationally self-interested agents would agree to principles that forbade killing one innocent person to save five. The idea is that such a practice would make seeking medical help too risky for people.[41] Contractualists would also require that David not kill the innocent person and thus let the five patients die. This is because killing one innocent person in order to save the lives of five others is to fail to treat that one person as being morally equal to others. After all, if their lives matter, why shouldn't his? Moreover, if David were to kill the patient in order to help other people, he would be treating his patient as a kind of tool, i.e., a means to saving other people's lives. This makes it clear that if David were to kill the one to save the five, he would not be treating the one with moral respect or regard.

1.2.iv Rossian Pluralism

All of the normative ethical theories (both consequentialist and deontological) that we have been considering thus far are *monistic* theories. That is, they think there is only *one* fundamental morally-relevant property that makes actions right or wrong. These properties make actions right or wrong by themselves and not in virtue of instantiating any other right-making or wrong-making properties. For example, Utilitarianism thinks that the only fundamentally important morally-relevant property is utility maximisation while Kantians think that the only fundamentally important morally-relevant property is rational choosability (as a universal law). However, W.D. Ross argues that there are actually numerous morally-relevant properties none of which is more fundamental than each other and thus it is

[41] Rosenberg (1992).

considered a *pluralistic* theory. As we will see, Ross's view can be seen as a kind of hybrid between deontology and consequentialism.

Ross argued that we have the following duties:

Fidelity: One should keep one's promises and be honest and truthful.

Reparation: One should make amends when one has wronged someone else.

Gratitude: One should be grateful to others when they perform actions that benefit one and one should try to return the favour.

Non-injury (or non-maleficence): One should refrain from harming others either physically or psychologically.

Beneficence: One should be kind to others and improve their health, wisdom, security, happiness, and well-being.

Self-improvement: One should improve one's own health, wisdom, security, happiness, and well-being.

Justice: One should be fair and distribute benefits and burdens equably and evenly.[42]

Ross called these principles "*prima facie"* duties, but he is most often interpreted as thinking that these are actually claims about we have moral reason to do.[43] However, when at least one of these principles disallows an action and none of the other principles allow it, then that action is wrong. When at least one of these principles allows an action and none of the other principles disallow it, then that action is permissible. Finally, when at least one of these principles require an action and none of the other principles forbid, then that action is morally required.[44]

Ross's theory shows us one way in which one could argue for convergence between normative ethical theories. This is because Ross's view can be seen as a kind of hybrid between deontology and consequentialism. As I will argue in section 3, one way to converge

---

[42]Ross outlines these seven *prima facie* duties in Ross (1930): 19-22.

[43] Stratton-Lake (2003): xxxiv.

[44] Ross (1930): 19-20.

two normative ethical theories is to combine the morally-relevant properties of each in a certain way. Ross's view can be seen as a combination of deontological and consequentialist morally-relevant properties. It is deontological because: (a) it makes the prevention of harm and injustice important, (b) it rejects the claim that only the outcome matters, and (c) it says that certain *types* of actions are required, permitted, or forbidden (e.g., lies, promise-keeping, harms, etc.). However, it is partly consequentialist because the duties of Self-Improvement and Beneficence hold that one ought to increase goodness (either for yourself or others).

Ross's view also shows us that normative ethical principles can vary in their scope, i.e., the range of situations in which they apply. Most normative ethical principles are meant to be fully general, i.e., they are meant to apply in all situations and to be able to answer any normative ethical question. We can call such principles *full scope* principles. For example, Utilitarianism, Kantianism, and Contractualism are meant to be full-scope principles.[45]

However, some normative ethical principles only apply in limited situations and therefore cannot answer all normative ethical questions. Each of Ross's *prima facie* duties is a moral principle of this type as is the Doctrine of Double Effect. We can call such principles partial scope principles. For example, take Ross's *prima facie* duty of gratitude. Could this principle help David in Organ Transplant when he is trying to figure out whether or not to kill one patient in order to save five? The answer is clearly no. In fact, most normative ethical questions have nothing to do with showing gratitude. Thus, the *prima facie* duty of gratitude is a partial scope principle.

The distinction between partial and full scope moral principles is important because the rest of this thesis will primarily be concerned with arguments about whether certain partial scope normative ethical principles in animal ethics and in environmental ethics

---

[45] However, Scanlon (1998) does argue that his contractualism is only concerned with what he calls the "morality of right and wrong" which is concerned with what duties people owe to each other (6). Thus, his view does not directly bear on certain questions, e.g., how to treat non-human animals.

can converge. However, what I say about convergence in the next two sections will apply equally to full-scope and partial-scope normative ethical principles.

2. What Does It Mean For Normative Ethical Theories To Converge?

Now that we have a better sense of what normative ethical theories are about, how they work, and how they can conflict with one another on certain practical questions (e.g., what David is morally required to do in Organ Transplant), we can address the question of what it means for two or more of these theories to converge. In this section, I will do two things. First, I will explain what it means for two or more normative ethical theories to converge. Second, I will discuss strategies for getting two or more normative ethical theories to converge and discuss some limitations of each strategy.

2.1 Practical Convergence

Recall that normative ethical theories answer two questions, i.e., the practical question of "What actions are right and wrong?" and the explanatory question "In virtue of what are actions right or wrong?" Two or more normative ethical theories can converge on their answers to either or both of these questions.

Let's start by looking at what it takes for two or more normative ethical theories to converge on the practical question. Practical convergence is not mere compatibility. Two or more normative ethical theories are practically compatible to the degree that they do not disagree about which actions are required, permitted, or forbidden. But two or more normative ethical theories can be practically compatible merely because they don't apply to any of the same situations or actions. Consider the following simple (and clearly implausible) normative ethical principles, which I have created solely for purpose of making my point:

Monday Killing: It is morally wrong to kill innocent people on Mondays.

Tuesday Killing: It is morally permissible to kill innocent people on Tuesdays.

These two views are practically compatible, but only for a trivial reason, i.e., they don't apply in the same situations. Monday Killing only applies to situations on Mondays and Tuesday Killing only applies to situations on Tuesdays. Thus, they do not disagree on the moral status of any action, but this is only because they are never applicable to the same situations or actions.

However, in order for two or more normative ethical theories to exhibit practical convergence, they must apply in at least some of the same situations and to at least some of the same actions.

Two or more normative ethical theories converge when they agree (as opposed to failing to disagree) on answers to either the practical question or the explanatory question mentioned above. When normative ethical theories converge on the question of which actions are right or wrong, this is *practical convergence*. Practical convergence is a matter of two or more normative ethical theories agreeing on which actions are required, permitted, or forbidden. Practical convergence is a matter of degree. The degree of practical convergence depends on how many actions two or more different normative ethical theories agree on the moral status of. The more actions that two or more normative ethical theories agree are required, permitted, or forbidden, the more they exhibit practical convergence. What I will call *Total practical convergence* occurs when two or more normative ethical theories require, permit, and forbid all and only the same actions. What I will call *Mass practical convergence* occurs when two or more normative ethical theories require, permit, and forbid many or most of the same actions.[46]

What I will call *trivial practical convergence* occurs when two or more normative ethical theories require, permit, and forbid a small amount of the same actions. The reason

---

[46] While numerous philosophers have written about convergence between normative ethical principles or theories, none of them (to my knowledge), have focused on the fact that there can be varying degrees of convergence. For example, Norton (1991), Varner (1994), Parfit (2011), and Baumann (2021).

these theories converge is a kind of accident, because there is no morally-relevant explanation for why they convergence. What I will call *non-trivial practical convergence* occurs when two or more normative ethical theories require, permit, and forbid the same actions to a higher degree than trivial convergence involves, but to a lesser degree than mass convergence involves. Exactly how much convergence is involved in non-trivial convergence is vague and will vary depending upon what theories are compared. The important point is that it involves a higher amount of convergence than one would expect by accident. As we will see, non-trivial convergence *usually* involves some degree of explanatory convergence.

2.2 Explanatory Convergence

Like practical convergence, explanatory convergence (either non-fundamental or fundamental) requires more than mere compatibility. Two or more normative ethical theories exhibit explanatory compatibility to the degree that they do not disagree about why actions are required, permitted, or forbidden. But two or more normative ethical theories can be explanatorily compatible merely because they don't apply to any of the same situations or actions. Consider the following toy theories:

> *Monday Wrongness:* On Mondays, actions are wrong if and only if they constitute harming an innocent person.
>
> *Tuesday Wrongness:* On Tuesdays, actions are wrong if and only if they constitute a failure to maximise utility.

These two views exhibit explanatory compatibility, but only for a trivial reason, i.e., they don't apply in the same situations. Monday Wrongness only applies to situations on Mondays and Tuesday Wrongness only applies to situations on Tuesdays. Thus, they do not disagree on the moral status of any action or on the explanation of why those actions are wrong.

Two or more theories can also converge on the explanatory question, i.e., "What explains why right actions are right and wrong actions are wrong?" That is, they can agree on the right-making and wrong-making properties, i.e., the properties that make right actions

right and wrong actions wrong. However, there are two kinds of right-making and wrong-making properties and therefore two kinds of explanatory convergence. There are *derivative* or *non-fundamental* right-making and wrong-making properties. These properties make actions right or wrong but only because they are instances of some more fundamental right-making or wrong-making property. For example, consequentialists and deontologists can agree that some action is wrong because it constitutes the killing of an innocent person. However, they will disagree about why the fact that this action constitutes the killing of an innocent person makes this action wrong. Deontologists will argue that this particular killing is wrong because one has a moral duty to not harm innocent people and killing them constitutes a harm. Consequentialists will argue that this particular killing is wrong not because of a particular moral duty to not harm, but because it has bad consequences (e.g., it fails to maximise overall happiness).

There are also *non-derivative* or *fundamental* right-making and wrong-making properties. These properties make actions right or wrong by themselves and not in virtue of instantiating any other right-making or wrong-making properties. For example, according to Utilitarianism, the only fundamental right-making and wrong-making property is utility maximisation. If any other properties are right-making or wrong-making, it is only because these properties instantiate the property of maximising utility or the property of failing to maximise utility.

Non-fundamental explanatory convergence is a matter of two or more normative ethical theories agreeing on why actions are required, permitted, or forbidden. That is, it is agreement on what properties at least sometimes count as right-making or wrong-making. For example, most forms of deontology and most forms of consequentialism will argue that one ought not to torture children for fun. Or at the very least, they will agree that the fact that some act constitutes the torture of children for fun counts against performing that action.

Fundamental explanatory convergence is a matter of two or more normative ethical theories agreeing on the fundamental explanation of why actions are required, permitted, or forbidden. That is, it is agreement on what properties at least sometimes count as fundamental right-making or wrong-making.

3. How Can Normative Ethical Theories Converge?

One may ask how it is possible for two or more normative ethical theories to converge on either non-fundamental or fundamental right-making and wrong-making properties. After all, isn't what distinguishes normative ethical theories what they consider to be non-fundamental or fundamental right-making and wrong-making properties? I now turn to the third main question of this chapter: how is *any* convergence (i.e., practical or explanatory) possible between different normative ethical theories?

3. 1 Strategies for Practical Convergence

In this section, I will do two things. First, I will explain why most philosophers agree that there is at least some minimal level of practical convergence. Second, I will explain a few strategies that philosophers have employed to argue that there can be mass or even total practical convergence. In the next section, I will then turn to the question of trivial, non-trivial, mass, and total explanatory convergence.

The question of how it is possible for minimal practical convergence to occur is easy to answer. The first reason that there is some minimal level of convergence is that all these theories are trying to answer the same questions and as a matter of luck they will sometimes agree. For example, deontologists and consequentialists might agree that it is morally impermissible for one person to murder an innocent person. Why do they converge on this answer? Because it just so happens that this action both violates the deontological constraint against harming innocent people and, at the same time, fails to maximise happiness. Thus, sometimes a single action has both a deontological wrong-making property *and* a utilitarian

wrong-making property. However, had things been slightly different, that action might not have had the property of failing to maximise happiness. For example, perhaps the innocent person was hated by many people and his death would have brought them great joy. In this case, Utilitarianism would have required the murder.

A second reason that there is some minimal level of practical convergence between the major normative ethical theories is that such theories must entail or at least be compatible with what Mark Timmons (2012) calls our *considered moral beliefs*.[47] These are the deeply held and widely shared beliefs we have about which actions are morally required, permitted, or forbidden. In order for a normative ethical theory to be plausible it must entail or at least be compatible with some set of these beliefs.[48] For example, it is a widely agreed and deeply held belief that it is wrong to torture children for fun. If a normative ethical entailed this belief is false, it would count against the plausibility of this view. Thus, all normative ethical theories must be amended to avoid being incompatible with a large number of considered moral beliefs. Because all these theories must be compatible with a large number of considered moral beliefs, they will converge on the moral status of at least some actions. Because this degree of convergence is not significant and the particular cases on which different theories convergence is merely coincidental, I will refer to this level of convergence as "trivial".

Given that different moral theories posit very different right-making and wrong-making properties (e.g., harm, maximising pleasure/happiness, respecting autonomy, being permitted by a principle that no one could reasonably reject, and so on) it is not surprising that most deny that there is any more than *trivial* convergence between normative ethical theories. So how could there possibly be mass or total practical convergence?

---

[47] Timmons (2012): 14.

[48] Some consequentialists are sceptical of this claim because they think our moral intuitions are biased or otherwise unreliable. For example, Singer (2005) and the research he discussed in that paper.

There are two basic groups of strategies for arguing that two or more normative ethical theories exhibit mass or total convergence. One group of strategies, which I will explore in the next section, consist of arguments that these normative ethical theories exhibit some high degree of explanatory convergence. Because explanatory convergence consists in arguing that two or more theories agree (to some extent) on the morally-relevant properties, and the degree of agreement on morally-relevant properties determines whether there is practical convergence, showing that there is explanatory convergence shows that there is also practical convergence. A second group of strategies consists of different arguments that purport to show that these two or more theories exhibit mass or total practical convergence without arguing that they have any or much explanatory convergence.

Now that we know the two broad groups of strategies one has for arguing for practical convergence, I will spend the rest of this section discussing particular instances of the second strategy. In the next section, I will turn to particular instances of the first strategy.

Recall that the second group of strategies for arguing that there is mass or total practical convergence between two normative ethical theories is to argue that the right-making and wrong-making properties that each theory posit are actually largely or completely co-extensional, i.e., the set of actions with the first kind of morally-relevant property largely or completely overlaps with the set of actions with the second kind of morally-relevant property.

Of course, no one argues that the original version of different normative ethical theories exhibit mass or total convergence, because it seems quite obvious that the different morally-relevant properties that different normative ethical theories posit are quite distinct and thus have different extensions. For example, just consider again David's choice in Organ Transplant. The utilitarian will argue that one is morally required to kill the one to save the five because that will lead to overall happiness, but the Kantian will deny this because such

an action cannot be universalised in the right way. Thus, in order to show that distinct normative ethical theories exhibit mass or total convergence, philosophers have to alter one or both theories. But how can one alter such theories to make them exhibit mass or total convergence?

One strategy for getting two normative ethical theories to exhibit mass or total practical converge is to make both theories concerned with *similar* morally-relevant properties. For example, one could get deontological theories and consequentialist theories to exhibit mass or total convergence if one made consequentialism concerned with minimising harm to innocent people. In particular, if one made consequentialism focused on minimising harm to innocent people *and* made preventing harm *weightier* than producing benefits. Because Deontological theories tend to be concerned with preventing harm to innocent people, a consequentialist theory that required minimising harm (and put more weight on preventing harm), would likely make many of the same recommendations as many deontological theories.

In fact, Parfit takes this strategy in arguing that Kantianism and contractualism converge. He changes Kantianism into a form of contractualism. He argues that the following is the most plausible version of Kantianism:

> *Kantian Contractualism:* Everyone ought to follow the principles that everyone could rationally will (or choose) to be universal laws.[49]

When Parfit says that principles can be rationally willed or chosen as "universal laws," he means that such principles are either universally accepted or universally followed.[50]

Notice that Parfit has altered the focus of Kantianism. Recall that the version of Kantianism we discussed above said the following: whether an action is right or wrong is determined by which maxim or principle is such that *the person deciding how to act* could

---

[49] Parfit (2011): 407.

[50] Parfit (2011): 407.

rationally choose that everyone *in her world* act on that maxim (e.g., everyone falsely promises to pay back their loan). However, Parfit has altered Kantianism to be about which principles are such that *everyone* (and not just the person deciding how to act) could rationally choose them to be accepted or followed. Thus, the best version of Kantianism, according to Parfit, is a kind of contractualism.

Given that he thinks that the best version of Kantian ethics is a form of contractualism it is perhaps not surprising that he thinks that Kantian ethics converge with what he thinks is the best version of contractualism, i.e., T.M. Scanlon's contractualism.

Parfit interprets Scanlons's view as follows:

> *Scanlon's Formula:* Everyone ought to follow the principles that no one could reasonably reject. [51]

The move from Kantian Contractualism to convergence with Scanlonian Contractualism is rather straightforward. Parfit argues that the principles that everyone could rationally will (or choose) to be universal laws *just are* the principles that no one could reasonably reject.[52] Thus, Kantianism (or Kantian Contractualism) and Scanlonian Contractualism are co-extensional and thus they exhibit total practical convergence.

A second strategy for getting two normative ethical theories to exhibit mass or total practical convergence is to argue that one theory *mimics* the other. For example, take the following simple version of Divine Command Theory:

> *Divine Command Theory:* One is morally required to do whatever God commands one to do.

One could argue that Divine Command Theory exhibits mass or total convergence with another normative ethical theory if one argued that God's commands would largely or totally mimic the requirements of some other normative ethical theory.

---

[51] Parfit (2011): 411.

[52] Parfit (2011): 411-412.

However, arguing that Divine Command Theory mimics some other normative ethical theory is not to argue for explanatory convergence. First, one can avoid explanatory convergence by insisting that God *could* have commanded different actions, but, for some reason, choose to command actions that largely or completely overlap with the actions required by some other normative ethical theory. Second, one can avoid claiming that this involves explanatory convergence even if God makes the commands he does because he thinks the other normative ethical theory is correct. This is because one could argue that God's commands are *fundamental* right-making properties. Let me explain. One option is that God commands some actions *because* it's morally required and the only reason that it is morally required is that that action has a non-divine right-making property. In such a case, it is the non-divine right-making property that fundamentally explains the rightness or obligatoriness of the action. A non-divine case would be one in which a parent commands their child to tell the truth. The parent's command is not what makes telling the truth morally required, but the parent commands it because it is morally required.

In the case I'm thinking about, it is the fact that God *commands* something that *makes* it morally required. As I'm thinking about it, God could command any action and simply in virtue of being commanded by God that action is required. It doesn't matter whether there are any other right-making properties present in the action. However, when God commands an action that is already morally-required *before* he makes the command, that action can have two fundamental right-making properties, one divine (i.e., God's command) and one non-divine (e.g., it saves innocent lives). In other words, in this case, there are *two* sources of one's requirements to perform the right actions, i.e., the other normative ethical theory *and* God's commands.

To be clear, it is not uncommon for there to be more than one source of a requirement to do something, so there is nothing suspicious about this second option. For example,

perhaps Amy is required to help her child do her homework because she is her mother. But she can also be required to help her child do her homework because she *promised* her spouse or her child to help her with her homework. So, if she doesn't help her, she violates her duty as a mother and she breaks her promise to her spouse or child.

3.2 Strategies for Explanatory Convergence

But what about (fundamental or non-fundamental) explanatory convergence? How could that be possible? I will consider eight different strategies for converging two or more different normative ethical theories at the explanatory level. Along the way, I will briefly consider some limitations of each strategy. As I noted above, any strategy for converging two or more normative ethical theories at the explanatory level is also a strategy converging them at the practical level.

As with practical convergence, there can be total, mass, trivial, and non-trivial explanatory convergence. This can occur with either non-fundamental or fundamental morally-relevant properties. Total explanatory convergence occurs when two or more normative ethical theories agree on all the morally-relevant properties. Mass explanatory convergence occurs when two or more normative ethical theories agree on many or most of the morally-relevant properties. Trivial explanatory convergence occurs when two or more normative ethical theories agree on some small portion of the morally-relevant properties. Non-trivial explanatory convergence occurs when two or more normative ethical theories agree on a number of the morally-relevant properties that is greater than the number of morally-relevant properties involved in trivial explanatory convergence and less than the number of morally-relevant properties involved in mass explanatory convergence. The difference between trivial and non-trivial explanatory convergence will often not be clear. This is because most views only posit a small number of morally-relevant properties. However, if there were theories that posited a large number of morally-relevant properties,

then the difference between trivial and non-trivial explanatory convergence would be much clearer.

In this section, I will focus on fundamental explanatory convergence because any strategy for fundamental explanatory convergence is also a strategy for both non-fundamental explanatory and practical convergence. Before discussing strategies for fundamental explanatory convergence, I will briefly discuss two options for non-fundamental explanatory convergence that do not involve fundamental explanatory convergence.

3. 2. i Strategies for Non-Fundamental Explanatory Convergence

There are two strategies for arguing for non-fundamental explanatory convergence that do not involve arguing for fundamental explanatory convergence. These two strategies are parallels of arguments for practical convergence that do not involve fundamental explanatory convergence.

One strategy for getting two normative ethical theories to exhibit mass or total non-fundamental converge (without mass or total fundamental explanatory convergence) is to make both theories concerned with *similar*, but distinct, fundamental morally-relevant properties. Take the following toy examples:

> *Golden Rule Wrongness:* A property is a wrong-making property of an action if it is a property that you wouldn't want actions that affect you to have.
>
> *Platinum Rule Wrongness:* A property is a wrong-making property of an action if it is a property that others wouldn't want actions that affect them to have.

Given that there is probably large agreement amongst you and other people concerning the kinds of properties that you both wouldn't want action that affect you to have, there will be at least mass non-fundamental explanatory convergence between these two theories. For example, most people don't want to be physically or psychologically harmed, robbed, imprisoned, killed, etc., and so the property of being a harming, a robbing, an imprisonment, and a killing, are all non-fundamental wrong-making properties.

A second strategy for getting two normative ethical theories to exhibit mass or total non-fundamental convergence parallels one of the strategies for getting mass or total *practical* convergence. In particular, one can argue that one theory *mimics* the morally-relevant properties of the other. For example, take the following toy theory:

> *Divine Disapproval Theory:* A property is a wrong-making property of an action if God disapproves of that property.

One could argue that Divine Disapproval Theory exhibits mass or total non-fundamental explanatory convergence with another normative ethical theory if one argues that the properties that God disapproves of largely or totally mimic the non-fundamental wrong-making properties of some other normative ethical theory. For example, perhaps God disapproves of harm, stealing, imprisonment, killing, and so on. Divine Disapproval Theory will exhibit mass or total non-fundamental explanatory convergence with any normative ethical theory that holds that most or all of those properties are wrong-making.

However, as we saw above, arguing that Divine Disapproval Theory mimics some other normative ethical theory is not to argue for *fundamental* explanatory convergence. First, one can insist that God could have disapproved of different properties, but, for some reason, disapproved of the same properties that some normative ethical theories posit as non-fundamental wrong-making properties. Second, as we also saw above, one can avoid claiming that this involves fundamental explanatory convergence even if God disapproves of the properties he does because he thinks the other normative ethical theory is correct about the non-fundamental wrong-making properties. This is because one could argue that God's disapproval *also makes* those properties non-fundamental wrong-making properties in addition to whatever other fundamental wrong-making properties already made those properties non-fundamental wrong-making properties. Let me explain.

Imagine that some version of Utilitarianism is true and a non-fundamental wrong-making property is "causes pain." The reason causing pain is a wrong-making property is that it contributes to the failure to maximise overall pleasure. Now imagine that God likes Utilitarianism and commands people to not cause harm. In this case, there is non-fundamental explanatory convergence between Divine Disapproval Theory and this version of Utilitarianism. However, there is no fundamental explanatory convergence because there are now *two* fundamental explanations of why causing pain is a non-fundamental wrong-making property. First, some version of Utilitarianism entails this is true. Second, God commands people to not cause harm.

3. 2. ii Strategies for Fundamental Explanatory Convergence

One strategy for arguing that two or more normative ethical theories exhibit fundamental explanatory convergence is to argue that the fundamental morally-relevant properties that different theories posit are actually the *same* properties. To my knowledge, no one has defended explanatory convergence by arguing that the morally-relevant properties are two or more distinct normative ethical theories are the same. Parfit uses the metaphor of "climbing the same mountain from different sides" as a metaphor for this thesis that all normative ethical theories, once they are properly understood, exhibit some amount of practical and explanatory convergence.[53] While he doesn't think that any two normative ethical posit the same fundamental right-/wrong-making properties, he does argue that the principles that everyone could rationally will (or choose) to be universal laws *just are* the principles that no one could reasonably reject.[54] However, he is not arguing that the property of being "allowed/ disallowed by principles that everyone could rationally will (or choose) to be universal laws"

[53] Parfit (2011): 419.

[54] Parfit (2011): 411-412.

is the same property as being "allowed/disallowed by principles that no one could reasonably reject."[55] As I mentioned above, he just thinks that these two properties are co-extensional.

This strategy would most clearly show that there is no fundamental disagreement between moral philosophers. However, it would imply that different moral philosophers were climbing the same mountain from the *same* side without knowing it. But this is implausible. One reason to think this is that *if* philosophers were doing this, then their different normative ethical theories wouldn't disagree about which actions are right and which are wrong. But it is generally agreed that all normative ethical theories disagree about the morality of some large set of actions.

A second option for getting two or more normative ethical theories to exhibit fundamental explanatory convergence is to create a conjunctive hybrid of the theories. On this view, an action is right only if it has the fundamental right-making properties of *both* theories and an action is wrong only if it has the fundamental wrong-making properties of both theories. For example, take the following normative ethical theories:

> *No Harm:* An action is morally wrong if and only if it causes an innocent person harm, otherwise it is morally permissible.
>
> *Utilitarianism:* An action is morally wrong if and only if it fails to maximise happiness, otherwise it is morally permissible.

A conjunctive hybrid of these two views would say the following:

> *No Harm + Utilitarianism:* An action is morally wrong if and only if it both causes an innocent person harm *and* fails to maximise happiness, otherwise, it is morally permissible.

If this view were correct, then many fewer actions would be wrong, because causing harm or failing to maximise happiness would now no longer be sufficient for moral wrongness. Only actions that had both properties would be wrong.

---

[55] While Parfit (2011) does not explicitly say this, the fact that he mentions them as distinct criteria for determining if an action is right or wrong is pretty good evidence that he doesn't think they are the same property (412-413)

No one, to my knowledge, has argued for such a view. This is not surprising for a few reasons. First, the hybrid theory can't explain the rightness or wrongness of any of the cases that No Harm and Utilitarianism disagree on. This is a large explanatory disadvantage. Second, it is also more ontologically complex and therefore, everything else being equal, less plausible than either No Harm or Utilitarianism. Third, it weakens the normative power of both morally-relevant properties. Let me explain. Both morally-relevant properties remain fundamental in the sense that their ability to *make* actions right or wrong is not derived from any other property. However, both are now only *conditionally* morally-relevant, i.e., both are such that they can only make an action right or wrong *on the condition* that the other property is present in the action as well. But most normative ethical theories are thought to supply unconditional and fundamental morally-relevant properties.[56]

Parfit sometimes speaks as if this is the kind of view that he endorses. He focuses on three distinct normative ethical principles and then argues that they exhibit some kind of explanatory convergence. The three theories he discusses are:

> *Universal Follow Rule Consequentialism:* Everyone ought to follow the principles of which it is true that, if they were universally followed, things would go best.[57]
>
> *Kantian Contractualism:* Everyone ought to follow the principles that everyone could rationally will to be universal laws.[58]
>
> *Scanlon's Formula:* Everyone ought to follow the principles that no one could reasonably reject.[59]

According to Universal Follow Rule Consequentialism, an act is wrong if and only if it is disallowed by a principle which, were it universally followed, things would go best.

---

[56] Exceptions to this include Rossian pluralism and particularism. Particularism is the view that there are no general normative ethical principles, i.e., principles that determine what is right or wrong in *all* cases. For particularists, whether an action is right or wrong is *context-sensitive*. See Dancy (2004) for a classic defence of particularism.

[57] Parfit (2011): 405.

[58] Parfit (2011): 407.

[59] Parfit (2011): 411.

According to Kantian Contractualism, an act is wrong if and only if it is disallowed by a principle that everyone could rationally will to be a universal law. Finally, According to Scanlon's Formula, an act is wrong if and only if it is disallowed by a principle that no one could reasonably reject.

It appears that Parfit might defend a conjunctive hybrid of these three theories, because he writes:

> *Triple Theory:* An act is wrong if and only if, or just when, such acts are disallowed by some principle that is: (1) one of the principles whose being universal laws would make things go best, (2) one of the only principles whose being universal laws everyone could rationally will, and (3) a principle that no one could reasonably reject.[60]

He also offers the following shortened version:

> TT: An act is wrong just when such acts are disallowed by some principle that is optimific, uniquely universally willable, and not reasonably rejectable.

The idea is that an act is wrong just in case it is the kind of act that is disallowed by a principle with the following three properties. First, the disallowing principle is such that if everyone followed that principle things would go best, i.e., the principle is optimific. Second, the disallowing principle is such that it would be rational for everyone to will that it be a universal law. Third, the disallowing principle is such that no one could reasonably reject it. Thus, it sounds as if Parfit is conjoining the wrong-making properties of Universal Follow Rule Consequentialism (more on this view below), Kantian Contractualism, and Scanlonian Contractualism.

However, Parfit argues that all three of these morally-relevant properties are co-extensional.[61] So, there is a way in which Triple Theory is a hybrid view. But it differs from No Harm + Utilitarianism in that it is a hybrid of three co-extensional properties and No Harm + Utilitarianism is a hybrid of two properties with different extensions.

---

[60] Parfit (2011): 413.

[61] Parfit (2011): 411-412.

It also differs from No Harm + Utilitarianism in that it requires changing one of the morally-relevant properties (i.e., being a principle whose being universally followed would make things go best) into a non-fundamental morally-relevant property (more on this below). As I will explain further below, Parfit doesn't treat the above three theories are being equally explanatorily fundamental. While he thinks that Kantian Contractualism and Scanlonian Contractualism are co-extensional and that both are fundamental explanatory properties, he doesn't think that the morally-relevant property discussed in Universal Follow Rule Consequentialism is a fundamental explanatory property. Rather, he thinks that one ought to follow the principles mentioned in Universal Follow Rule Consequentialism *because* they are the only principles that can be rationally willed to be universal laws. So, the truth of Kantian Contractualism *explains* the truth of Universal Follow Rule Consequentialism and therefore the latter does not provide a *fundamental* explanatory property while the former does.[62]

Finally, Parfit's view also entails that, even though Kantian Contractualism and Scanlonian Contractualism are co-extensional and fundamental explanatory properties, they are both now only *conditional* explanatory properties. That is, they both are such that they can only make an action right or wrong *on the condition* that the other property is present in the action as well. But, as noted above, most normative ethical theories are thought to supply unconditional and fundamental morally-relevant properties. Therefore, I think that this is another way in which Parfit fails to show that there is perfect convergence between these three views.

A fourth option for getting two or more normative ethical theories to converge is to create a disjunctive hybrid of the theories. Consider a case of creating a hybrid from just two theories. On this view, an action is right only if it has the right-making properties of *either* (or both) theories and an action is wrong only if it has the wrong-making properties of *either* (or

---

[62] For a related argument, see Baumann (2021).

both) theories. For example, taking No Harm and Utilitarianism as our example, we get the following disjunctive hybrid view:

> *No Harm or Utilitarianism:* An action is morally wrong if and only if it either: (1a) causes an innocent person harm *or* (2a) fails to maximise happiness (or both 1a and 2a are satisfied). An action is morally permissible if and only if it either: (1b) it doesn't cause an innocent person harm *or* (2b) it succeeds in maximising happiness (or both 1b and 2b are satisfied).

If this view were correct, then many more actions would be wrong, because causing harm or failing to maximise happiness would now both be sufficient for moral wrongness.

Ross's Pluralism might be considered a form of disjunctivism. For example, in cases where is there no conflict between *prima facie* duties, an action is wrong if and only if it violates at least of the duties and an action is right if and only if it is allowed or required by at least one of the duties. In other words, in cases without conflicts between *prima facie* duties, an action is wrong if and only if it violates the *prima facie* duty to: keep one's promises and be honest and truthful, or make amends when one has wronged someone else, or be grateful to others when they perform actions that benefit one and one should try to return the favour, or refrain from harming others either physically or psychologically, or be kind to others and improve their health, wisdom, security, happiness, and well-being, or improve one's own health, wisdom, security, happiness, and well-being, or be fair and distribute benefits and burdens equably and evenly, or some combination of these.

In cases in which there are conflicts between *prima facie* duties, we get the following views of right and wrong action. An action is wrong if and only if it violates at least one of the duties and the duty (or duties) it violates is weightier than the duty (or duties) that allows or required the action. An action is right if and only if it is allowed or required by a duty and the duty (or duties) that allows or requires it is weightier than the duty (or duties) that forbid it.

On this way of converging two normative ethical theories both morally-relevant properties remain unconditionally morally-relevant and fundamentally morally relevant. While disjunctive hybrid theories are popular in some fields (e.g., perception), they are not concerning moral theory. This is likely for a few reasons. First, most normative ethical theories posit only a single unconditional, fundamental morally-relevant property and are thus ontologically simpler than hybrid views. Second, disjunctive hybrid views call out for a unifying explanation, i.e., why do these two properties have in common such that they are both unconditional and fundamental morally-relevant properties? Without such an explanation, such hybrid views seem rather *ad hoc*. Finally, such views must explain how one can resolve the apparent conflict between the morally-relevant properties. For example, if an action fails to cause an innocent harm, then it is permissible (via 1b), but the same action might fail to maximise happiness and thus be wrong (via 2a).

A fifth option for getting two normative ethical theories to converge is to argue that at least one of the theories has multiple fundamental morally-relevant properties and the other theory includes at least one of these fundamental morally-relevant properties. For example, consider the following simple normative ethical theory:

> *Beneficence:* One should be kind to others and improve their health, wisdom, security, happiness, and well-being.

Beneficence exhibits partial fundamental explanatory convergence with Rossian Pluralism because it just is one of the principles that constitute Rossian Pluralism. However, it does not exhibit mass or total fundamental explanatory convergence with Rossian Pluralism, because Rossian Pluralism claims that there are numerous other fundamental morally-relevant properties.

There are a few potential drawbacks to this strategy. First, in cases like the one above, in which one normative ethical theory posits only one fundamental morally-relevant property

and the other posits eight, there is not much convergence. This is, at best, a trivial (as opposed to a mass or total) fundamental explanatory convergence. Second, there is still a sense in which there is a large disagreement between the two theories. This is because one theory says that *all there is* to morality is beneficence and the other theory says that there is much more to morality than beneficence.

A sixth option is to argue that the morally-relevant property that one theory posits is actually an instance of the morally-relevant property that another theory posits. This is what Parfit argues concerning Rule Consequentialism and Kantian Contractualism. He argues that the best version of consequentialism is Universal Follow Rule Consequentialism":

> Everyone ought to follow the principles of which it is true that, if they were *universally followed*, things would go best.[63]

Parfit then argues that these principles are the only principles whose being universal laws everyone could rationally will. In other words, Kantian Contractualism (along with some ancillary premises) entails Rule Consequentialism. Given that the principles of Rule Consequentialism are the only principles that whose being universal laws everyone could rationally will, Kantian Contractualism and Rule Consequentialism are completely co-extensional and exhibit total explanatory convergence.

However, he argues that these two properties are not on the same level of explanation. He writes, "Everyone ought to follow the principles whose being universal laws would make things go best, *because* these are the only principles whose being universal laws everyone could rationally will."[64] In other words, what *makes it the case* that one ought to follow the principles whose being universal laws would make things go best is that these principles are the only principles whose being universal laws everyone could rationally will. As Parfit himself argues:

---

[63] Parfit (2001): 405.

[64] Parfit (2011): 411.

> When acts are wrong . . . that is not merely or mainly because such acts are disallowed by one of the optimific principles [the principles specified by Rule Consequentialism]. These acts are also wrong because they are disallowed by one of the only set of principles whose being universal laws everyone could rationally will.[65]

Notice that what we have here is the claim that the *fundamental* morally-relevant property is being "allowed/disallowed by principles whose being universal laws everyone could rationally will" and the property of being "allowed/disallowed by the principles of which it is true that, if they were *universally followed*, things would go best," is a *non*-fundamental morally-relevant property. That is, the latter morally-relevant property is only morally-relevant because it is an instance of the former morally-relevant property.

Does this count as total explanatory convergence? Yes and no. If Parfit's arguments work, then he has shown that Kantian Contractualism and Rule Consequentialism agree that the property of being "allowed/disallowed by the principles of which it is true that, if they were *universally followed*, things would go best," is a morally-relevant property. That is, when an action is right or wrong, this can be explained by the fact that this action is allowed or disallowed by the principles of which it is true that, if they were *universally followed*, things would go best. But, this property is always a *non*-fundamental morally-relevant property.

However, Kantian Contractualism and Rule Consequentialism still disagree about what properties are *unconditionally* and *fundamentally* morally-relevant. Kantian Contractualism holds that the only unconditional, fundamental morally-relevant property is being allowed/disallowed by principles whose being universal laws everyone could rationally will and Rule Consequentialism holds that the only unconditional, fundamental morally-relevant property is being allowed/disallowed by the principles of which it is true that, if they

---

[65] Parfit (2011): 417.

were *universally followed*, things would go best. Thus, while there is a kind of total explanatory convergence, there is no *fundamental* explanatory convergence at all.

That being said, Parfit's argument does show us there are two ways in which some property can be a non-fundamental morally-relevant property. First, a non-fundamental morally-relevant property can be co-extensional with a fundamental morally-relevant property. This is what Parfit argues is true about the properties of being allowed or disallowed by the principles of which it is true that, if they were *universally followed*, things would go best. This means that this kind of non-fundamental morally-relevant property is similar to a fundamental morally-relevant property in that both are *always* morally-relevant and thus both always provide some explanation of the rightness or wrongness of an action.

Second, a non-fundamental morally-relevant property can fail to be co-extensional with a fundamental morally-relevant property and thus will be morally-relevant in certain cases. For example, causing harm is often, but not always a morally-relevant property for utilitarians. Some actions don't involve any harm and so the rightness or wrongness of the action can be explained without saying anything about harm. And, even if the action does involve harm if the harm is outweighed by pleasure and the action is morally required, then the property of causing harm will not explain why the action is wrong, because it will not be wrong.

The downside to this strategy is that it requires at least one of the theories to change the morally-relevant property it claims was unconditional and fundamental to conditional and non-fundamental. Even then, this strategy is insufficient for any amount of fundamental explanatory convergence. This is because one of the morally-relevant properties is not a fundamental morally-relevant property and so this isn't a case in which there is *fundamental* explanatory convergence.

A seventh, option for getting two normative ethical theories to converge is to take certain elements of each and create a new theory from these elements. For example, let us take No Harm and Utilitarianism as our test theories again. To create a new theory from these two, one can first start by taking harm to innocent people as the value/disvalue that is most morally important. We can then look at the ultimate aim of Utilitarianism and see that it is to make things go best.[66] Combining what No Harm sees as the most important moral value and what Utilitarianism sees as the broad aim of morality, we get the following theory:

Minimise Harm: One ought to act always so as to minimise harm to innocent people.

Importantly, Minimise Harm is not equivalent to Utilitarianism because minimising harm is not equivalent to maximising benefit (e.g., happiness or pleasure). For example, one can increase pleasure or happiness without thereby reducing harm. For example, imagine giving more money to a billionaire. They would likely give them some more pleasure or happiness, but it wouldn't constitute preventing any harm to that person.

There are a few drawbacks to this strategy. First, one faces the explanatory challenge of why one chooses which elements of which theory. Second, neither theory gets to keep its unconditional fundamental morally-relevant property.

An eighth, and final, option is to argue that the supposed fundamental morally-relevant properties of two or more normative ethical theories are actually just non-fundamental morally-relevant theories that share a common fundamental explanation. On this view, both normative ethical theories were correct about what the non-fundamental morally-relevant properties are—although they mistook them for fundamental morally-relevant properties. For example, consider the following toy normative ethical theories:

No Harm*: An action is morally wrong if it causes an innocent person harm.

Beneficence*: An is morally permissible if it benefits an innocent person.

Now imagine that while these two views were posited as fundamental wrong-making and right-making properties, there is actually an explanation of why those two views are partly correct. For example, consider the following toy theory:

> Divine Attitude Theory: An action is morally wrong if God disapproves of it and an action is permissible if God approves of it.

Now imagine that God disapproves of all and only those actions that cause harm to innocent people and that God approves of all and only those actions that benefit innocent people. When now have a more fundamental explanation for these two theories, which offer a more unified account of the fundamental morally-relevant properties.

The first obvious drawback of this theory is that it allows two or more normative ethical theories to converge only by changing what those theories posited as fundamental morally-relevant properties into *non-fundamental* morally-relevant properties. Some might deny that this amounts to a true convergence between two normative ethical theories. Second, proponents of either or both normative ethical theories might disagree about the fundamental moral relevance of the fundamental morally-relevant property that was posited to converge the two views.

3.3 The Best Prospects for Convergence

It seems to me that the most plausible place for convergence between normative ethical theories is at the level of non-fundamental right- and wrong-making properties. However, as we will see this kind of convergence need not entail a large amount of practical convergence.

Recall that non-fundamental right-and wrong-making properties are properties that help explain the moral status of actions at a non-fundamental level. These properties are right- and wrong-making only because they instantiate fundamental right- and wrong-making properties. For example, causing physical harm is often a non-fundamental right- or wrong-making property according to Utilitarianism. This is because causing physical harm often

leads to a failure to maximise pleasure. But not always. Sometimes causing someone physical harm causes other people a great deal of pleasure and in those cases causing physical harm can be a right-making property.

The reason that mass non-fundamental explanatory convergence between different normative ethical theories is at least plausible is that normative ethical theories are supposed to be able to do two things. First, they are supposed to entail that the kinds of action we generally think are right (e.g., donating to charity, helping friends, keeping promises, etc.) and wrong (e.g., acts of harming, killing, deceiving, torturing, stealing, and so on) are actually right and wrong. Second, they are supposed to do this by *explaining* why these kinds of actions are generally right or wrong by providing fundamental right- and wrong-making properties that explain why actions with certain properties tend to be right and actions with certain other properties tend to be wrong. For example, consider the following properties of actions: causing harm, causing death, being an act of deception, being an act of torture, being an act of theft, and so on. All these properties tend to lead to the failure to maximise overall pleasure or happiness, they all often constitute violations of duties (e.g., Ross's *prima facie* duties), maxims that told us to perform these actions are not rationally choosable as universal laws, they are the kinds of actions that would be disallowed by principles that no one could reasonably reject, and so on.

However, even mass convergence on these non-fundamental right- and wrong-making properties is not enough to entail practical convergence, especially between versions of consequentialism and deontology. This is because versions of consequentialism tend to allow for the weighing up of all the consequences of an action while versions of deontology tend to disallow this weighing up. Let me explain by returning to Organ Transplant. According to Utilitarianism, David is morally required to kill the one healthy patient in order to save the five sick patients. However, Utilitarians think that the fact that David must kill someone

counts against the permissibility of his action. However, the fact that David could save more lives by killing the healthy patient counts for the permissibility of his action. Moreover, once the amount of pleasure of happiness generated by the fact that five people will be saved is weighed against the amount of pain caused by the killing of one person, we see the more overall pleasure will result from the killing than from refraining from killing. Unlike with Utilitarianism, many versions of deontology (i.e., those which endorse constraints) will not allow good consequences to matter at all for determining what David is morally required to do. The fact that five people will be saved and there will be a positive balance of pleasure simply does not matter for these deontologists. This partly explains why consequentialists and deontologists will disagree about what David is morally required to do in Organ Transplant.

In fact, as we will see in Chapters 4, it is implausible that there is much practical (or explanatory) convergence between Utilitarianism and a form of deontology (i.e., the Moral Rights View) concerning animal experimentation. However, we should not be pessimistic about all prospects for convergence. The fact that consequentialist and deontological views might have trouble exhibiting non-trivial amounts of practical and explanatory convergence doesn't mean that other normative ethical views will have the same issues. In fact, as we will see in Chapters 3 and 5, there is a surprising amount of convergence that is possible between views in environmental ethics.

4. Conclusion

The purpose of this chapter was to clarify some of the core ideas that underlie the rest of this thesis. To that end, I first explained what normative ethical theories are and what questions they are meant to answer. I then distinguished different forms they can take (e.g., consequentialist vs deontological), what values they are concerned with (e.g., pleasure, happiness, harm, and so forth), and explained concrete explains of each of the forms of

normatively ethical theories (e.g., Utilitarianism, Kantianism, Contractualism, and Rossian Pluralism). Next, I explained what it means for normative ethical theories to converge. I distinguished: (a) practical and explanatory convergence, (b) non-fundamental and fundamental explanatory convergence, and (c) trivial, mass, and total convergence. Finally, I explained several strategies for accomplishing practical and explanatory convergence and discussed the limitations of these strategies.

In the next chapter, I will address another core issue of what kinds of value there are and what kinds of value environmental and animal ethicists care about. In Chapter 3, I turn to the question of convergence in environmental ethics. In particular, I consider the prospects for convergence concerning the question of when and why we should treat nature well. I start by considering an influential argument from Bryan Norton according to which the two most prominent normative ethical views in environmental ethics (Anthropocentrism and Non-Anthropocentrism) exhibit mass or total convergence despite not exhibiting any explanatory convergence. I argue that this view doesn't show that there will be much practical convergence between Anthropocentrism and Non-Anthropocentrism. I then argue that a more nuanced view of Anthropocentrism (i.e., Broad Anthropocentrism) can exhibit at least non-trivial practical convergence and total non-fundamental explanatory convergence with Non-Anthropocentrism. The reason that Broad Anthropocentrism and Non-Anthropocentrism can exhibit non-trivial practical convergence is because they exhibit total non-fundamental explanatory convergence.

In Chapter 4, I turn to the prospects of practical and explanatory convergence in animal ethics concerning animal experimentation. I consider the plausibility of at least non-trivial convergence between Utilitarianism and a form of deontology (i.e., the Moral Rights View). Gary Varner argues that Utilitarianism will exhibit at least non-trivial practical convergence with a principle that is part of the Moral Rights Views (i.e., the worse-off

principle). However, I argue that the worse-off principle is not, by itself, a version of the Moral Rights View that anyone would accept. Moreover, any plausible version of the Moral Rights View will include additional principles that would prevent the Moral Rights View from exhibiting even non-trivial convergence with Utilitarianism.

In Chapter 5, I turn from the question of practical convergence to the question of convergence on the question of how one ought to feel. In particular, I argue that there is convergence on whether and why one ought to care about nature and non-human animals. In particular, I argue that Broad Anthropocentrism and Non-Anthropocentrism exhibit total convergence on whether to finally care about nature and they exhibit total non-fundamental explanatory convergence as well. I argue that they exhibit total convergence on whether to finally care about nature *because* they exhibit total non-fundamental explanatory convergence. I then argue that Utilitarianism and the Moral Rights View exhibit total convergence on whether to care about non-human animals and total non-fundamental explanatory convergence. I argue that they exhibit total convergence on whether to care about non-human animals because they exhibit total non-fundamental explanatory convergence.

The Appendix addresses a recent argument that we can increase agreement about environmental policies without appealing to convergence if we appeal to important relationships that humans have with nature. I argue that this strategy will not work because it is implausible that humans have any relationships with nature that are sufficient for requiring that we treat nature well.

Chapter 2: Kinds of Value

1. Introduction

A core issue throughout most of this thesis is what kinds of value exists in the world and which entities have which kinds of value. The purpose of this chapter is to clarify what kinds of value exist, how they differ from one another, which entities can have which kinds of value, and how the answers to this question bear on the two main debates I engage with in this thesis (i.e., Anthropocentrism vs Non-Anthropocentrism and Utilitarianism vs the Moral Rights View (MRV)).

This chapter has the following structure. First, I briefly outline the two main debates I'm concerned with in this thesis. I also specify what kinds of value is involved in each of these debates. Second, I outline what kinds of value philosophers think exist and how they are related to each other. Third, I will argue that once we understand how these different kinds of value are related, we can see that the core of the disagreement between Anthropocentrism and Non-Anthropocentrism, on the one hand, and Utilitarianism and MRV, on the other, is about what has *final* value. This helps us focus in on what matters most for the debates and it helps us avoid focusing on different kinds of value, e.g., intrinsic and inherent, because doing so might distract us from what really matters to proponents of all these views.

2. The Debates

In this thesis, I'm concerned with a particular debate about animal ethics, i.e., animal experimentation. The debate between Utilitarians and proponents of MRV is a debate about the extent to which it is morally permissible to engage in animal experimentation. In particular, it is about engaging in experiments that cause harm to non-human animals.[67] On

[67] This debate only concerns sentient animals, i.e., animals that are capable of feeling pleasure and pain.

the Utilitarian view, animal experimentation is morally permissible if and only if it has the right consequences overall, i.e., if and only if it leads to the maximisation of overall happiness/pleasure or desire satisfaction in the world. On the Utilitarian View, animals do not have intrinsic value. The only thing that has intrinsic value for them is the pleasure/happiness or desire satisfaction that agents are required to maximise. Animals, like humans, are only valuable as containers of potential pleasure/pain, happiness/unhappiness, or desire satisfaction/frustration. This means that any kind of experimentation can be morally permissible or even required according to Utilitarianism as long as it produces enough pleasure, happiness, or desire satisfaction.

On the Moral Rights View, animal experimentation is either never morally permissible or rarely morally permissible—depending upon how strict one views moral rights as being. This is because animals have moral rights which limit the ways in which they can be morally permissibly treated in the same way that humans have moral rights that limit the ways in which they can be morally permissibly treated. The source or foundation of the moral right that these animals have is their "inherent" value. This inherent value comes from the fact that animals are what Tom Regan calls, "subjects of a life."[68] That is, they have beliefs, desires, emotions, consciousness, the ability to initiate action, a continuous identity over time, and so on.

Thus, the debate about the moral permissibility of animal experimentation is partly a debate about what kind of value animals have. If they have inherent value, then it looks like proponents of MRV are right about animal experimentation, i.e., they are right that we should severely limit animal experimentation. If animals lack inherent value and have value only insofar as they are potential containers of happiness, pleasure, and desire satisfaction, then Utilitarianism is right about animal experimentation.

---

[68] Regan (2004): 243.

In environmental ethics, there is debate about how we're morally required to treat nature. While it is agreed that we're morally forbidden from treating nature any way we please, there is a substantial disagreement about why nature matters morally. And, the reason nature matters morally determines exactly how we're morally required to treat it. According to Anthropocentrism, only humans are intrinsically valuable and therefore only humans are morally important in their own right. This means that nature is valuable only insofar as it can benefit humans and thus it is not intrinsically valuable. Thus, the question of how we're morally required to treat nature is fully determined by how our treatment of nature affects humans. That is, Anthropocentrists claim that nature is only instrumentally valuable, i.e., valuable as a means of supporting the interests of humans. According to Non-Anthropocentrism, however, it is false that only humans have intrinsic value. In particular, Non-Anthropocentrists argue that nature itself is also intrinsically valuable. They argue that this means that how we're morally required to treat nature is determined by the effects our treatment has on nature itself and not just on how it affects humans. Thus, like the debate between the Utilitarian View and the Moral Rights View concerning animal experimentation, the debate between Anthropocentrism and Non-Anthropocentrism in environmental ethics is partly a debate about what kind of value exists and what entities have what kind of value.

Given that different kinds of values and which entities have them is at the core of the debates that my thesis concerns, it is essential to see how these kinds of value are related to each other. First, it is important because it will help us understand precisely what the views I'm discussing are saying. Second, it is important because it helps see exactly what these different views are disagreeing about.

In what follows, we will discuss the following kinds of value: objective value, subjective value, final value, instrumental value, extrinsic value, intrinsic value, and inherent

value. As we will see, although the debates I examine in this thesis are often put in terms of what has intrinsic value, what really matters for these debate is actually what has final value.

3. Objective vs Subjective Value

In this section, I will distinguish between objective and subjective value, indicate which kind of value is the focus of the debates I'm interested in, and whether or not these kinds of values are compatible.

The kind of value that is at stake in the Anthropocentrism/Non-Anthropocentrism and Utilitarianism/MRV debates is objective (i.e., mind-independent or assessor-independent) value. To say that something has mind-independent or assessor-independent value is to say that it is valuable in virtue of something other than the assessments (e.g., beliefs or other attitudes) of persons. That is, the source of the value is not the beliefs, perceptions, desires, etc. of any individual person or group of people. Thus, something can be objectively valuable even if no one believes it is or wants it to be valuable, and it can be valuable even if no one exists to assess it as valuable. For example, the COVID-19 vaccine is objectively valuable because it saved the lives of millions of people.[69] Even if no one believed that the vaccine is valuable, it would still be valuable. Even if no one believed that it actually saved millions of lives, it would still be valuable, because, as a matter of fact, it saved millions of lives.

Objective value is to be contrasted with subjective value (i.e., mind-dependent or assessor-dependent value).[70] When something is subjectively valuable, it is valuable just because it is valued or assessed as being valuable by some individual or group of people. In other words, the fact that someone or group values X or believes that X is valuable is the

---

[69] I'm assuming here that human lives are objectively valuable and that something that is an effective means of protecting human lives is also objectively valuable. I thank an anonymous referee for helping me be clear about this.

[70] See O'Neill (1992).

source of X's value. If everyone ceased valuing X or believing that X is valuable, then X would cease to have any value at all.

Thus, the question of whether something has objective or subjective value is a question about what kind of properties give things value. Objectivists think that the value of nature is generated by properties other than the attitudes of assessors while subjectivists think that the value of nature is solely generated by the attitudes of assessors. Of course, something could be both objectively and subjectively valuable if the source of its value were both its objective (i.e., mind-independent or assessor-independent) properties and people's valuing it. However, I will only be focused on objective value as that it is the kind of value at stake in the Anthropocentrism/Non-Anthropocentrism and Utilitarianism/MRV debates.[71]

4. Final Value and Instrumental Value

In this section, I will define both final value and instrumental value and explain how they are different. This section will also help us understand how the final/instrumental value distinction differs from another common distinction in value theory: intrinsic/extrinsic value.

When something is finally valuable, it is valuable on its own, for its own sake, in its own right, as such, or as an end. That is, it is valuable, but its value doesn't come just from the fact that it is an effective means to bringing about something else of value.[72] Common examples of things that have final value are the following: pleasure, happiness, knowledge, and achievements.[73] Under normal circumstances, it is valuable to feel pleasure, to be happy, to know things, and have achievements. Moreover, the value of these things doesn't seem to be exhausted by the fact that they can bring us other things of value. For example, perhaps being happy is also valuable because it can make us kinder to strangers and knowledge is

---

[71] There are some who have argued that the value of nature is subjective (e.g., Callicott (1989) and Elliot (1992)) and some have questioned the idea of the objective value of nature (e.g., Svoboda (2011)).

[72] Of course, something can be both finally and non-finally valuable at the same time, as we will shortly see.

[73] See Hurka (2001).

valuable because it lets us make well-informed decisions. However, happiness would still be valuable even if it didn't make us kinder to strangers and knowledge would still be valuable even if the knowledge was not practically relevant. This is because happiness and knowledge are valuable *in their own right* or *for their own sake*.

Another way to make the point is to note that finally valuable things are worth pursuing for their own sake. That is, it makes sense for one to pursue happiness, knowledge, achievement, and pleasure even if nothing else of value comes about as a result of achieving these ends.

We can also get a better sense of final value by looking at its opposite, i.e., instrumental value. Something is instrumentally valuable when it is valuable as a means to something else that has value. That is, something is instrumentally valuable when it is a casual means to producing or acquiring something else that has value.

Common examples of things which tend to only have instrumental value include: money, sources of nutrition (e.g., food and drink), household appliances (e.g., dishwasher, washing machine, dryer, and so on), and perhaps even health. Why are household appliances valuable? Because they are effective means of making our lives easier. That is, they are valuable because they are effective means to something else that is valuable: our well-being. Why is money valuable? Because it is an effective means of getting things we want. In fact, it is one of the most effective means of getting other things we want. But money is not valuable for its own sake or in its own right. Notice that treating money like it is valuable in its own right is exactly the problem with misers like Ebenezer Scrooge in Charles Dickens' *A Christmas Carol*.[74] Scrooge values money for its own sake, i.e., he treats it like its finally valuable. But notice that we think that this is the wrong attitude to have toward money. The

[74] Dickens (2003).

explanation of why this is a mistake is simple: money is not finally valuable, it's only instrumentally valuable and therefore to treat it as if it is finally valuable is a mistake.

Importantly, when something is valuable as a means to something else, the something else can be either itself instrumentally valuable or finally valuable. For example, money is instrumentally valuable because it can be used to purchase household appliances, but household appliances are themselves only instrumentally valuable. On the other hand, money is also instrumentally valuable when one can pay someone to give one knowledge, e.g., to teach one how to speak another language or understand economics. This knowledge is plausibly valuable in its own right because it is knowledge—although it might also be instrumentally valuable.

Another way to make the point is to note that instrumentally valuable things are *not* worth pursuing for their own sake. They are only worth pursuing if getting them would result in getting something else of value. That is, it makes sense for one to pursue money, try to get household appliances, or eat food only if something else of value comes about as a result of achieving these ends. If you couldn't buy anything with money, it would make no sense to pursue it.

To get a better sense of the difference between final value and instrumental value, consider the following question: Why should I pursue health? One might respond, "because being healthy is required for engaging in certain pleasant activities (e.g., traveling, exercising, socialising, etc.)." In other words, if one has poor health, then one cannot engage in these pleasant activities. But then the following question arises, "Why should I try to be able to engage in pleasant activities?" To which one might respond, "They're pleasant." But then the following question arises, "Why should I pursue pleasure or having pleasant experiences?" At this point, one might naturally be confused and respond, "Having pleasant experiences is

worth pursuing for its own sake." That is, it looks like a certain stage in the questioning, one can justifying reject the question, "But why should I pursue that?"[75]

This kind of back-and-forth could be had regarding the value of many things, e.g., "Why is helping the needy valuable?", "Why is the Mona Lisa valuable?", "Why is knowledge valuable?", "Why is knowing how to cure diseases valuable?" In response to these questions, one will eventually get to a point at which it no longer makes sense to ask the question, "But why is that valuable?" For example, one might answer that helping the needy is valuable because it involves reducing human suffering. If someone were to ask, "Why is reducing human suffering valuable," you might again reject the question by responding, "It just is."

This line of questioning shows us two things. First, some things ought to be pursued because something else is worth pursuing and the first thing is a way of getting the second thing. For example, health is worth pursuing because being able to have pleasurable experiences is worth pursuing. Second, some things are worth pursuing for their own sake. For example, pursuing pleasurable experiences seems to be worth pursuing in its own right or for its own sake. That is why one can justifiably reject the question, "Why should I pursue pleasant experiences?"

Why are certain things worth pursuing for their own sake? Because they are finally valuable, i.e., valuable for their own sake. This answer relies on two plausible claims. First, if something is valuable, then, all else being equal, it seems worthy of pursuing. Second, if something is valuable *for its own sake*, then, all else being equal, it is worthy of pursuing *for*

---

[75] A related line of questioning is also used to bring out the distinction between derivative and non-derivative value (see Zimmerman and Bradley (2019)). The question of how the derivative/non-derivative value distinction maps onto the distinctions between final/instrumental value and intrinsic/extrinsic value is controversial. Fortunately, for our purposes, we can set this debate aside. For more on the relationship between these distinctions in value, see Zimmerman and Bradley (2019)).

its own sake. On the other hand, things with instrumental value are also valuable and therefore, all else being equal, worth pursuing. However, they are only worth pursuing for the sake of something else that has value.

How instrumentally valuable something is depends upon: (1) how effectively or efficiently it can cause or lead to something else that is valuable and (2) how valuable the end for which it is a means is. For example, kindness is probably a more reliable and faster way of getting other people to like you, but you might also be able to get some people to like you by being cruel to them. In the case of getting someone to like you (which we can assume is valuable), kindness is more instrumentally valuable than cruelty. Generally speaking, more money is more instrumentally valuable than less money. For example, with more money, I can purchase greater pieces of art. So, a large sum of money is more instrumentally valuable than a small sum because the former is an effective means to something more valuable than the latter.

5. Extrinsic vs Intrinsic Value

A distinction that is related to, but importantly different from, the final/instrumental distinction is the extrinsic/intrinsic value distinction.[76] The final/instrumental value distinction is about whether something has value in its own right or only because it can cause or bring about something else. The extrinsic/intrinsic value distinction concerns what kind of properties the value that something has supervenes on. What does it mean for some property (or properties) to supervene on another property (properties)? At its core, supervenience is a claim about metaphysical dependence. In particular, A-properties supervene on B-properties if and only if no two things can differ concerning A-properties without differing concerning

[76] Korsgaard (1983) was the first to emphasise that final and instrumental value are contrasting kinds of value and that intrinsic and extrinsic value are contrasting kinds.

B-properties. In this way, A-properties are dependent on B-properties. The concept is often explained in part by the slogan "there cannot be an *A*-difference without a *B*-difference."[77]

For example, thin moral properties (e.g., right/wrong, good/bad, optional/non-optional) are thought to supervene on certain non-moral, but morally-relevant, properties (e.g., causing/reducing harm, respecting autonomy, increasing/decreasing pleasure, lying/telling the truth). Take the case of betraying a friend. Grant that it is morally wrong to betray a friend. In this case, the property of wrongness supervenes on the property of being an act of betrayal. To say that wrongness in this case supervenes on the property of being an act of betrayal means that there could not be two acts that differ on whether they are wrong without differing concerning whether they are acts of betrayal.

Many philosophers also think that mental states supervene on (physical) brain states. This means that no two people can be in the same mental state without also being in the same brain state. For example, if being in the mental state of pain supervenes on the brain state of C-fibres firing, then there are no two people who differ with regard to whether they are experiencing pain without differing with regard to whether they have C-fibres firing.[78]

Thus, to say that someone is extrinsically or intrinsically valuable is to make a claim about the kind of properties on which that thing's value supervenes. With this in mind, we can explore the ways in which something can be extrinsically valuable and then contrast that with intrinsic value.

Something is extrinsically valuable when it is valuable in virtue of some set of its external or extrinsic properties. While it is controversial exactly what counts as an extrinsic property, for our purposes, we can assume that extrinsic properties are just relational

---

[77] For more on supervenience, see McLaughlin (1995) and McLaughlin and Bennett (2021).

[78] Kripke (1980): 98.

properties.[79] Relational properties are properties that something has in virtue of being related to something else in some way. For example, whether a person is tall or rich is a relational property because it depends upon the height and wealth of the people around them. I'm rich only if I have more money than some portion of my community and I'm tall only if my height is greater than the average height of people of my age and gender in my community.

Something can be extrinsically instrumentally valuable, e.g., money. As we saw above, the value of money is based on the fact that it is a means to getting other things and the fact that it is a means is based on a relational property of money, i.e., the fact that people are willing to trade things for money. Thus, in the case of instrumental value, the relevant relation is *causal*. That is, something is instrumentally valuable when it is causally related to something valuable in the right way, i.e., it is an effective or efficient causal means to that thing.

Something can be extrinsically valuable in virtue of *contributing* (e.g., as a part) to something that is valuable. For example, a particular brush stroke on the Mona Lisa is valuable in virtue of helping to make up the whole painting, which is itself finally valuable. But the particular brush stroke is not valuable because it *causes* the painting to exist. After all, the single brush stroke doesn't cause the painting to exist at all. Rather, it is one part, among many, that is necessary for the painting to exist.[80]

Something can be extrinsically valuable in virtue of *signalling* or providing evidence of something that is valuable.[81] For example, imagine that a star basketball player injures his leg during a game. Imagine further that the player is brought to the hospital to be X-rayed. Finally, imagine that that X-ray indicates that he doesn't have a fracture or any serious

---

[79] Ronnow-Rasmussen (2015) writes, "Value theorists tend to assume we have an intuitive grasp of what properties are internal or external (30)."

[80] For more on this kind of value, see Bradley (1998): 110 and 120-121.

[81] For more on this kind of value, see Bradley (1998): 110 and 118-119.

damage. Some think that this X-ray has value in virtue of signalling or indicating something valuable, i.e., that the star basketball player is not seriously injured and can keep playing.

As I will argue further in Chapter 5, something can have extrinsic value in virtue of providing the opportunity or materials for something of value to come into existence. Raz (2011) call this kind of value, "facilitative value".[82] To see what I have in mind, consider a case from Zimmerman and Bradley.[83] Imagine that you allow a struggling, talented artist, to use your home as their art studio while you are on vacation. Imagine further that, given their talent, the art that they will create there will be finally valuable. However, if they didn't have this place to work, they would not have been able to create their masterworks of art. It's not clear that the workspace you provide to them is valuable in its own right, i.e., finally valuable. And, if it is, it has nothing to do with the artist and their work. Moreover, your house doesn't cause the artist to create masterpieces of art. So, it's instrumentally valuable. Whether they painted was up to them. The value that your house seems to have is as a space that provided the artist *the opportunity* to create things with final value. The same holds for any workspace that provides the opportunity for final value to be created, e.g., the workspaces of fiction and non-fiction writers, scientists, medical doctors, and so on. Thus, these spaces seem to be extrinsically valuable, i.e., valuable in virtue of some set of their relational properties.

The same is true of museums. Art museums provide visitors with the opportunity to experience beauty and wonder, which are plausibly finally valuable experiences. Natural history and science museums provide visitors the opportunity to gain knowledge about important facts (e.g., human history, life on earth, and scientific theories). Notice that museums do not cause visitors to have this knowledge. After all, one could go to the

---

[82] Raz (2011): 220.

[83] Zimmerman and Bradley (2019).

museums and ignore everything around as a small child might. Nor is it clear that museums are finally valuable.

Finally, something can even be extrinsically *finally* valuable. For example, the pen that Abraham Lincoln used to sign the Emancipation Proclamation seems to be finally valuable in virtue of some of its relational properties.[84] Notice that this pen is certainly valuable and that at least some of its value comes from what it was previously used for and not just what it can be used for now. In particular, some of its value seems to come from the role in played in an important historical event. But the property of being used by a particular person for a particular signing is a relational property and so the final value of Lincoln's pen must be based on extrinsic (i.e., relational) properties.

It makes sense to think of Lincoln's pen as being finally valuable, i.e., valuable in its own right, even in virtue of a relational property. This is plausible at least in part because this a relational property that the pen can never lose. Because what gives it value is the role it played in the past and facts about the past cannot change, it will always be the case that Lincoln's pen played this role in history. And, it will therefore always be true that it is valuable. This distinguishes Lincoln's pen from the unremarkable pens on my desk. These pens only have value as an effective means to writing. However, this property could easily go away, e.g., if the pen runs out of ink (and the ink can't be replaced) or if I accidentally break the pen.

Fourth, something is intrinsically valuable when it is finally valuable in virtue of some set of its non-relational properties (i.e., what some people call internal or intrinsic

[84] Kagan (1998): 285.

properties).[85] For example, some think that pleasure or happiness are intrinsically valuable. This is because they seem valuable for their own sake (i.e., finally valuable), but not in virtue of any set of their relational properties.

What are intrinsic properties? While I cannot hope to settle the debate here, I will provide some common characteristics of intrinsic properties. None of these characteristics is meant to be fully convincing. However, putting them together will help one get a better grasp of what intrinsic properties are like.[86]

First, if P is an intrinsic property of x, then any sentence ascribing P to x is entirely about x.[87] For example, if "good" is an intrinsic property of being a pleasure, then the sentence, "Pleasure is good" is entirely about pleasure. On the other hand, the property "being well-known" is not an intrinsic property. For example, the sentence "*Squid Game* is well-known" is about the Netflix show *Squid Game* and the large amount of people who are familiar with the show. Second, if P is an intrinsic property of x, then x has P in virtue of how x, and nothing else, is.[88] That is, x has P in virtue of the way that x itself is. For example, triangles have the property of being three-sided in virtue of the way that triangles themselves are and the way that triangles are is not the way that anything else is. So, three-sidedness seems like an intrinsic property of triangles. But it is not true that things have extrinsic properties in virtue of the way that they, and nothing else, are. For example, if Jessica is rich,

---

[85] For more on intrinsic value, see Beardsley (1965), Elliot (1992), O'Neill (1992), Lemos (1994), Zimmerman (1999), Ronnow-Ramussen and Zimmerman (2006), Ronnow-Ramussen (2015), and Zimmerman and Bradley (2019). For discussion of intrinsic value in environmental ethics, see: Callicott (1989), Hargrove (1992), Regan (2004), Svoboda (2011), and Rea and Munns (2017). For discussion of intrinsic value in animals ethics, see: Verhoog (1992) and DeVries (2008).

[86] For a more in-depth examination of the various ways of defining intrinsic properties, see Marshall and Weatherson (2018).

[87] Lewis (1983): 197.

[88] Lewis (1983): 197.

then it is false that she is rich in virtue of the way that she, and nothing else, is. Whether she is rich is, in part, about how other people are, i.e., how much money they have.

Finally, If P is an intrinsic property of x, then every duplicate of x has P—regardless of the duplicate's surroundings.[89] For example, a take triangle with the inside angles of 57, 86, and 37. Any duplicate of this triangle will also have the inside angles of 57, 86, and 37. However, it is false that if P is an extrinsic property of x, then every duplicate of x has P. For example, imagine that x is P stands for "Tim is tall." Will all duplicates of Tim be tall? No. Imagine that Tim is 193cm in a country in which the average height for males is 178cm. But now imagine that Tim's duplicate is in a different country where the average height for males is 210cm. In that country, Tim is no longer tall. So, height must be an extrinsic property.

6. Inherent Value

Finally, being inherently valuable means being finally valuable in virtue of having certain kinds of properties. For example, Regan (2004) argues that something is inherently valuable only if it is the "subject of a life." By this he means that the thing has certain mental states (e.g., desires, beliefs, emotions), the capacity for certain experiences (e.g., it can experience pleasure and pain), and certain abilities (e.g., the ability to initiate action in pursuit of a goal).[90] Thus, inherent value is something that is only had by certain kinds of creatures, e.g., most humans and some non-human animals. However, while certain things like fine art might be finally or intrinsically valuable, they cannot be inherently valuable because they lack the above properties.

One seemingly puzzling feature of Regan's conception of inherent value, is that everything that has inherent value has it *equally*, i.e., inherent value is not a matter of degree. He writes:

---

[89] Lewis (1983): 197.

[90] Regan (2004): 243.

> …Inherent value is thus a categorical concept. One either has it, or one does not. There are no in-betweens. Moreover, all those who have it, have it equally. It does not come in degrees.[91]

This seems to importantly distinguish inherent value from all other kinds of value that we have discussed so. In the next section, however, I will argue that Regan is really thinking of being inherently valuable as being a kind of final value. Once we understand, we will see that there is a sense in which his claim that everything that has inherent value has it equally is no longer puzzling.[92]

A remaining question is whether being inherently valuable is a way of being intrinsically valuable. In particular, we can ask if the property of being a subject of life is an intrinsic property of an animal. The answer seems to be yes. First, if P is an intrinsic property of x, then any sentence ascribing P to x is entirely about x. The sentence, "My dog is the subject of a life" seems to be entirely about my dog. Second, if P is an intrinsic property of x, then x has P in virtue of how x, and nothing else, is. This characterisation is less clear. On the one hand, the fact that my dog is a subject of a life is true in virtue of the kind of creature he is, but not in virtue of the way that *he himself* is. On the other hand, the particular way in which he is a subject of a life (e.g., the beliefs, desires, emotions, and so on that he has) is something he has in virtue of the way that he, and nothing else, is. Finally, any duplicate of my dog should also be the subject of life. So, it looks like at first glance, that being inherently valuable is a way of being intrinsically valuable.

7. What Kind of Value Matters?

Getting clear on the kinds of value that exist matters for getting clear about what the participants in the Anthropocentrism/Non-Anthropocentrism and Utilitarianism/MRV debates

---

[91] Regan (2004): 240-241.

[92] For more on the inherent value of non-human animals, see: Warren (1987), Russow (1988), Edward (1993), Frey (2002), and Francione (2008).

really care about. In this section, I will show that authors in both debates have conflated intrinsic and final value. This is important, because whether or not we can show that there is any convergence between these theories will depend upon what the theories actually hold—at least when they are properly understood. Moreover, when assessing the plausibility of these views, it matters whether they are on the hook for defending the view that nature/non-human animals have final value vs the view that nature/non-human animals have intrinsic value. As we saw above, to have intrinsic value is to be finally valuable in virtue of some set of non-relational properties. So, having to defend the claim that something has intrinsic value involves defending the view that that thing has final value *and* that its final value comes from some set of its non-relational properties.

7.1 Final vs Intrinsic Value in Environmental Ethics

Now that we have seen that there is a difference between final value and intrinsic value, we can see that many authors in environmental ethics and specifically in the Anthropocentrism/ Non-Anthropocentrism have conflated final and intrinsic value. Consider the following representative quotations:[93]

> Something is intrinsically valuable if it is valuable in and for itself—if its value is not derived from its utility, but is independent of any use or function it may have in relation to something or someone else. In classical philosophical terminology, an intrinsically valuable entity is said to be an "end-in-itself," not just a "means" to another's ends.[94]

> The principle of intrinsic value states that, regardless of what kind of entity it is in other respects, if it is a member of the Earth's community of life, the realization of its good is something *intrinsically* valuable. This means that its good is prima facie worthy of being preserved or promoted as an end in itself and for the sake of the entity whose good it is. Insofar as we regard any organism, species population, or life community as an entity having inherent worth, we believe that it must never be treated as if it were a mere object or thing whose entire value lies in being

[93] In addition, see Callicott (1984): 299.

[94] Callicott (1989): 131.

> instrumental to the good of some other entity. The well-being of each is judged to have value in and of itself.[95]
>
> One of the most important ethical issues raised anywhere in the past few decades has been whether nature has an order, a pattern, that we humans are bound to understand and respect and preserve. It is the essential question prompting the environmentalist movement in many countries. Generally, those who have answered ‘yes’ to the question have also believed that such an order has an intrinsic value, which is to say that not all value comes from humans, that value can exist independently of us: it is not something we bestow. On the other hand, those who have answered 'no* have tended to be in an instrumentalist camp. They look on nature as a storehouse of Resources' to be organised and used by people, as having no other value than the value some human gives it.[96]

These environmental ethicists contrast intrinsic value with instrumental value. For example, Callicott writes, “an intrinsically valuable entity is said to be an "end-in-itself," not just a "means" to another's ends.” Notice that Taylor says that something that is intrinsically valuable “must never be treated as if it were a mere object or thing whose entire value lies in being instrumental to the good of some other entity.” Thus, he views intrinsic value as being the opposite of instrumental value. Finally, Worster also contrasts intrinsic value with instrumental value.

But now that we know the difference between intrinsic value and final value, we can see that these authors (and those that follow them) are making a mistake. What Non-Anthropocentrists think is that nature is *finally* valuable, i.e., valuable for its own sake and not a means to serving human interests. Anthropocentrists disagree and argue that nature is only instrumentally valuable. However, notice that there is no mention in the above quotations (and other Non-Anthropocentrists texts) of intrinsic/non-relational properties and how/why these particular properties matter. Why is this? Because what really matters for Non-Anthropocentrists is that nature is *finally* valuable. This is because what Non-Anthropocentrists care about is establishing that nature has a value that is independent of the

---

[95] Taylor (1981): 198.

[96] Worster (1985): xi.

instrumental value has in virtue of being able to serve human interests. But all they need to establish this is to show that nature is finally valuable. They don't also have to show that nature has value in virtue of its non-relational properties. And so it is at best misleading to argue that the debate is about intrinsic value.[97]

7.2 Final vs Intrinsic and Inherent Value in Animal Ethics

We can see a similar conflation in the animal ethics literature. Consider the following representative quotations:

> (1) To say that animals have intrinsic value is to say that they do not only have instrumental value, i.e., that they are not only valuable as means to an end, but also in themselves or for their own sake. This sense seems to be derived from the sense in which the term 'intrinsic value' is used in the consequentialist tradition (see e.g., Moore, 1922). It seems to refer primarily to the type of value that attaches to animals (and not to duties we might have on the basis of this value). (2) To say that animals have intrinsic value is to say that they ought never to be treated solely as means. This sense is reminiscent of the meaning of 'intrinsic value' in the Kantian deontological tradition.[98]
>
> [T]hose concerned about animals and nature began to use the term "intrinsic value"; by using this term they sought to indicate that it is not only humans who have moral status. Other natural entities are said to have moral status as well: they have a "value of their own" independent of their utility-value for humans.[99]
>
> Something has intrinsic value if its existence is a good thing in itself, apart from its role as a means to other goods.[100]
>
> The idea of intrinsic value has been defined in a variety of ways, including "what is desired for its own sake," "what would be good even if it existed in isolation from everything else," "what ought to exist for its own sake," and "what is valued or preferred in itself." Differ though these characterizations do, each attempts to articulate the difference between (a) something's being of positive value only as a means to something else and (b) something's being good independently of its being a means to something else.[101]

---

[97] For an argument that nature has *extrinsic* final value, see Tenen (2020).

[98] DeVries (2008): 376.

[99] Verhoog (1992): 148.

[100] O'Neil (2000): 185.

[101] Regan (2004): 142.

First, notice that these authors always contrast "intrinsic value" with instrumental value. DeVries writes, "To say that animals have intrinsic value is to say that they do not only have instrumental value." Regan characterises intrinsic value as being "good independently of its being a means to something else." As we saw above though, the natural contrast of instrumental value is final value and something can be good/valuable independently of its being a means to something else if it is *extrinsically* finally valuable. Second, notice that the authors say nothing about the internal or non-relational properties of animals. Rather, they just argue that animals are valuable "for their own sake" and that they have a "value of their own".

What proponents of MRV really care about is showing that animals shouldn't be treated as if they are valuable only as a means to serving human interests. As Regan writes, "To borrow part of a phrase from Kant, individuals [e.g., certain non-human animals] who have inherent value must never be treated merely as means to securing the best aggregate consequences."[102] But all proponents of MRV need to establish to show that animals shouldn't be treated as "mere" means is to show that animals are *finally* valuable. This is because if something is finally valuable, then it is plausible that one shouldn't treat that thing just as a tool for serving human interests.

These authors also fail to distinguish between intrinsic and inherent value.[103] Even if being inherently valuable is just a way particular way of being intrinsically valuable, it matters that we distinguish them. For example, if the main reason we ought to treat humans and non-human animals with respect is that they have intrinsic value, then we need to explain why we should always value non-human animals and humans over other things that are

---

[102] Regan (2004): 249.

[103] Regan (2004) is an exception (235). Although, Regan's main point is that the inherent worth that some non-human animals have is not reducible to the intrinsic value of *their experiences* (235). In other words, some non-animals have inherent worth, but the source of this inherent worth is *not* the fact that they have or could have intrinsically valuable *experiences*.

intrinsically valuable (e.g., knowledge, achievement, pleasure, etc.). It seems like the value that certain non-human animals and humans have is importantly different than the kind of value that knowledge has. But if we just talk about the intrinsic value of animals giving us reasons to treat them well, we cannot clearly distinguish these differences.

Moreover, even though I indicated that being inherently valuable might be a way of being intrinsically valuable, the reasons I gave were not particularly strong. So, if it turns out that being inherently valuable is not a way of being intrinsically valuable, then the above authors (and many others) have been making an even bigger mistake in talking about intrinsic value when they should be talking about final value.

Finally, I will briefly argue that part of what Regan means when he says that something has inherent value is that it has final value. To see this, recall that Regan makes the following claim about inherent value:

> …Inherent value is thus a categorical concept. One either has it, or one does not. There are no in-betweens. Moreover, all those who have it, have it equally. It does not come in degrees.[104]

Here, Regan lists two features of inherent value: (1) something either has it or it doesn't (i.e., it's an on/off property) and (2) everyone who has the property has it equally. These features need not go together. For example, whether or not something is good is an on/off matter, but once something is good, it can be more or less good.

The property of having *final* value is such that something has it or it doesn't. It's either true or false that X is worth pursuing for its own sake. Moreover, anything that is worth pursuing for its own sake has this property *equally*. Notice that being worth pursuing for its own sake has two elements: (1) being worth pursuing and (2) *what makes it* worth pursuing (e.g., itself or something else). When two things are worth pursuing *for their own sake*, they equally possess the second element, i.e., it's equally true of both of them that *what makes*

---

[104] Regan (2004): 240-241.

them worth pursuing is themselves. That is, the explanation of why they are each worth pursuing is the same: they have value in their own right.

As an analogy, take the case of dining at either of two nice restaurants. Imagine that both are worth dining at because they both have delicious food. Regardless of whether the food is equally delicious, both restaurants have the property of being worth dining at *because the food is delicious* equally. The explanation of why one should eat at either place is the same: the food is delicious. If there's a difference in the deliciousness of the food, it doesn't follow that there is a difference in the explanation of why one should eat there.

However, it doesn't follow from this, that something that is finally valuable is not worth pursuing *more* than something else that is finally valuable. But the difference in the degree that something is worth pursuing is reflective of a difference in the degree of value, not in a difference of the degree "for its own sake". Thus, it seems to me that what Regan thinks is distinctive about something that is inherently valuable is that it is *finally* valuable.

7.3 Non-Relational/Relational Properties and Moral Relevance

We've seen in this section that what Non-Anthropocentrists and proponents of MRV really care about is final value instead of intrinsic value. In other words, what matters for them is the fact that something has final value, regardless of whether the properties that make it finally valuable are relational or non-relational. In this subsection, I want to add additional reason to think that this is the correct conclusion by showing that it is not controversial to hold that relational properties are relevant for moral properties (e.g., moral reasons, duties, obligations, etc.). I will focus on obligation for the sake of expediency, but what I say can be translated to moral reasons and duties.

It is uncontroversial that what seems to be non-relational properties matter for what obligations people have toward us. For example, part of the explanation of why people have an obligation to not harm or kill other humans is because those humans are conscious,

sentient, and rational beings. However, many of our obligations are based on or impacted by relational properties. Perhaps the most obvious case of this is that of the special obligation we have to our nearest and dearest. That is, we have certain obligations to our friends and family, and perhaps even to our colleagues and co-nationals.[105] For example, under normal circumstances, it seems plausible that I have a moral obligation to feed, clothe, provide shelter and emotional support to *my* child, but not to other people's children. And, given the choice between helping my spouse and a stranger, where all else is equal, I'm obligated to help my spouse. Likewise, I might be obligated to help my aging parents, but not your aging parents. But the source of these obligations seems to importantly involve my *relationship* to them. That someone is my child or parent or conational is a relational, not an internal or intrinsic, property that I have. This is not to deny that perhaps some intrinsic property of me is involved in my having an obligation to them, but it should be clear that my relationship to them is important.

Another way that I can have an obligation in virtue of my relational property is when I promise to do something for someone. The fact that I promised to help someone move obligates me to help them move. But the fact that I promised is a relational property. There is no internal or intrinsic about the property of having promised to do something.

Finally, sometimes one's epistemic or physical position or physical ability determines whether one has a particular moral obligation. For example, imagine that I'm the only person who knows that a doctor is about to accidentally give a poisonous liquid to a patient and I know that I'm the only person who knows this. It looks like, in virtue of being the only one who knows this fact, I'm morally obligated to intervene by telling the doctor to stop. Relatedly, I might be the only person close enough to a child drowning in a pond to reach them in time to save their life. Perhaps everyone else is too far away to reach the child in

---

[105] For an overview of this issue, see Jeske (2019).

time. So, the fact that I'm the only person close enough to save the child makes it the case that I'm morally obligated to save the child. Finally, imagine that an innocent person has fallen on train tracks and I'm the only person around that is strong enough to remove them from the tracks in time for them to avoid being hit by a train. Grant also that I can remove them without any risk to myself. It seems like I might be morally obligated to remove this person from the tracks. However, all these properties (i.e., being the only one around who knows something morally important, the only one around who is close enough to save someone, and the only one around who is strong enough to save someone) are relational properties. After all, these properties all have to do with how I compare to the people around me.

Whether an obligation is determined by a non-relational or a relational property seems to make no difference to the obligatoriness of some action. What matters morally is that someone has a moral obligation, not the kind of properties (i.e., relational or non-relational) in virtue of which one has that moral obligation. Likewise, what matters for Non-Anthropocentrists and proponents of MRV is that nature and animals are finally valuable and not the kind of properties (i.e., relational or non-relational) in virtue of which they are finally valuable.

8. Conclusion

In this chapter, I have distinguished the following kinds of value: objective value, subjective value, final value, instrumental value, extrinsic value, intrinsic value, and inherent value. I argued that once we distinguish these different kinds of value, we can see that the participants in the Anthropocentrism/Non-Anthropocentrism and the Utilitarianism/MRV debates really care about is final value. In particular, they care about *objective* final value. This means that debates about whether things have *intrinsic* or *inherent* value are somewhat misleading.

It is important for this thesis to show that being intrinsically valuable is a way of being finally valuable. This is because, as we will see in Chapter 3 and Chapter 5, this fact allows us to see that a version of Anthropocentrism and Non-Anthropocentrism can exhibit both practical and derivative explanatory convergence because the two views agree about what has final value. Moreover, it is important to see that what the philosophers engaged in the Anthropocentrism/Non-Anthropocentrism debate really care about is final value, because while the view of Anthropcentrism I sketch and Non-Anthropocentrism agree that nature is finally valuable, they disagree about whether it has intrinsic value. However, if what they care about most is final value, that the latter disagreement is not important.

Chapter 3: Converging Anthropocentrism and Non-Anthropocentrism?

1. Introduction

How should we decide which environmental policies are morally correct? As we saw in the last chapter, there are two broad approaches to answering practical and policy questions in environmental ethics: Anthropocentrism and Non-Anthropocentrism. According to Anthropocentrism, only humans (and human well-being, preferences, desires, interests, etc.) matter morally in their own right and not merely as a means for something else, because they are intrinsically valuable.[106] Thus, anything else (e.g., nature) only matters morally insofar as it is a means to benefiting humans.[107] Therefore, Anthropocentrism holds that that actions and policies concerning nature are morally required is solely determined by how those actions and policies affect humans (and their well-being, preferences, desires, interests, etc.). Non-Anthropocentrism holds that other creatures (e.g., non-human animals, plants, ecosystems, etc.) matter morally in their own right because at least some non-human entities are intrinsically valuable.[108] Therefore, Non-Anthropocentrism holds that that actions and policies concerning nature are morally required is determined by not only by how they affect humans but also by how they affect nature.

---

[106] Defenses of some version of Anthropocentrism include: Passmore (1980, 2010) and Norton (1984, 1991). Even though I argued in Chapter 2 that the Anthropocentrism/Non-Anthropocentrism debate is really about final, as opposed to intrinsic, value, I talk about these views in terms of claims about intrinsic value to stay true to the original positions. However, when I introduce my preferred version of Anthropocentrism (i.e., Broad Anthropocentrism), I will focus on final value because I argue that this version of Anthropocentrism and Non-Anthropocentrism exhibit derivative explanatory convergence because they agree that nature has final value.

[107] Callicott (2009): 142 and Stenmark (2009): 82.

[108] Defenses of some version of Non-Anthropocentrism include: Leopold (1949), Naess (1973), Routley and Routley (1979), Regan (1981), Norton (1984, 2002, 2009), Rolston (1988, 2012), Callicott (1989), Naess and Sessions (1995), and Sterba (2011).

Given their disagreement about what has intrinsic value and therefore what kinds of entities are morally important in themselves, it looks like Anthropocentrism and Non-Anthropocentrism are unlikely to agree on which actions involving nature are right or wrong. That is, it seems implausible that Anthropocentrism and Non-Anthropocentrism would recommend or forbid all or most of the same actions. In fact, it also seems implausible to that they would exhibit much practical convergence even if we accepted altered Anthropocentrism and Non-Anthropocentrism to be claims about final, as opposed to intrinsic, value. This is because, on this view, Anthropocentrism holds that only humans and their interests have final value and Non-Anthropocentrism holds that nature also has final value.

However, Bryan Norton argues that there is reason to think that Anthropocentrism and Non-Anthropocentrism can exhibit practical convergence. In particular, he argues for what he calls the "convergence hypothesis," i.e., Anthropocentrism and Non-Anthropocentrism, when properly understood, will recommend, require, and forbid all or most of the same environmentally responsible actions and policies.[109] Norton argues that the reason that Anthropocentrism and Non-Anthropocentrism appear as though they will issue conflicting recommendations is that people have assumed an implausible version of Anthropocentrism. He argues that philosophers have understood Anthropocentrism as only caring about the "felt desires" of humans, which are, roughly, unreflective desires. However, he argues that we should understand Anthropocentrism as concerning "considered desires." Norton's notion of considered desires is complex, but, for now, one can think of them as the desires a well-informed and rational person *would have* after a process of deliberation. Norton refers to this version of Anthropocentrism as "Weak" Anthropocentrism.[110] Norton thinks that once one has an appropriately nuanced version of Anthropocentrism in hand, it will not conflict with

---

[109] Norton (1991): 243.

[110] Norton (1984): 134.

Non-Anthropocentrism. Norton's strategy for arguing for practical convergence is the kind that doesn't appeal to any explanatory convergence at all. In fact, as we will see, his strategy is rather complicated.

I argue in this chapter that Norton's convergence hypothesis is implausible. I object to Norton's convergence hypothesis in two ways. First, I argue that Weak Anthropocentrism cannot be converged with Non-Anthropocentrism because the former allows humans to have considered desires that are inconsistent with treating nature in the way that Non-Anthropocentrists require. In particular, I argue that even the nuanced version of Anthropocentrism is likely to recommend actions and policies that conflict with the recommendations of Non-Anthropocentrism. As I will explain below, Norton thinks that if Anthropocentrism holds that what matters morally are the rationally considered or endorsed desires, then those rationally considered desires will always align with the recommendations of Non-Anthropocentrism. I argue that there can be rational disagreements about what one ought to do between the nuanced version of Anthropocentrism and Non-Anthropocentrism. Second, I argue that, even if Norton is right about mass or total practical convergence, many Anthropocentrists and Non-Anthropocentrists will be unsatisfied with his argument. This is because what they really care about are the properties that make actions right or wrong.

Finally, I will argue for a set of more modest convergence claims. First, I will argue that a certain version of Anthropocentrism (what I will call "Broad Anthropocentrism") exhibits some degree of practical and non-fundamental explanatory convergence with Non-Anthropocentrism. This is because, I will argue, both views agree that certain parts of nature are finally valuable, i.e., value for their own sake. Second, I will argue that this version of Anthropocentrism will exhibit some degree of convergence with Anthropocentrism on the answer to the question of *how* valuable nature is.

One important question that I want to address before moving on is why the question of convergence matters in environmental ethics. Anthropocentrism and Non-Anthropocentrism are the two most common normative ethical theories endorsed by people in charge of environmental policy and so if their two preferred views agreed on what to do, it will be easier to get agreement on which environmental policies to support. If there is general agreement on specific environmental policies, then it will likely be easier to get policies that protect nature enacted. Thus, the question of convergence is quite important.

The chapter will proceed as follows. First, I will explain the standard versions of Anthropocentrism and Non-Anthropocentrism and why it seems implausible to claim that they would converge. Second, I will explain Norton's more nuanced view of Anthropocentrism (i.e., Weak Anthropocentrism) and why he thinks that this version of Anthropocentrism allows for the convergence of Anthropocentrism and Non-Anthropocentrism. Third, I argue that Norton's nuanced version of Anthropocentrism fails to show that Anthropocentrism and Non-Anthropocentrism exhibit much convergence. Fourth, I argue that even if it works, it would not satisfy many proponents of Anthropocentrism and Non-Anthropocentrism. Fifth, I argue for a weaker version of the Convergence Hypothesis. I argue that what I call Broad Anthropocentrism exhibits a non-trivial amount of convergence with Non-Anthropocentrism about which actions are right and why they are right. Finally, I argue that Broad Anthropocentrism and Non-Anthropocentrism will likely exhibit some non-trivial amount of convergence on the answer to the question of how valuable nature is.

2. Standard Accounts of Anthropocentrism and Non-Anthropocentrism

In this chapter, I'm going to argue that Norton fails to show that plausible versions of Anthropocentrism and Non-Anthropocentrism exhibit either mass or total practical convergence. I'm then going to argue that there is at least non-trivial practical convergence between Anthropocentrism and Non-Anthropocentrism once we have a more plausible

version of Anthropocentrism on the table. In order to do all this, we need to have a clear idea of what Anthropocentrism and Non-Anthropocentrism hold.

Anthropocentrism and Non-Anthropocentrism are both partial scope normative ethical theories because they only apply to actions that involve nature. The fundamental difference between these two views is how they answer the *fundamental explanatory* question, i.e., "What are the fundamental right-making and wrong-making properties?" As its name suggests, Anthropocentrism holds that only certain properties of humans (e.g., their well-being, preferences, desires, interests, etc.) fundamentally matter when answering questions about how to treat nature. That is, if it is morally permissible to do something to nature, the fundamental explanation is that that action either benefits humans or at least does not adversely affect humans. If it is morally wrong to do something to nature, the fundamental explanation or wrong-making property is that that action adversely affects human beings. The underlying assumption of Anthropocentrism is that only humans have intrinsic value and therefore only humans matter morally in their own right.[111]

Non-Anthropocentrism holds that nature matters morally in its own right. In particular, Non-Anthropocentrism holds that nature is *intrinsically* valuable and in virtue of the fact that it is intrinsically valuable it is morally wrong to treat it in certain ways, e.g., destroying it. The reason it is morally wrong to harm nature is because it is intrinsically valuable and the reason why humans might be morally required to preserve it is also because it is intrinsically valuable.

As we saw in Chapter 2, one of the core disagreements between Anthropocentrism and Non-Anthropocentrism is about *what* has intrinsic value. Anthropocentrism holds that only humans are intrinsically valuable while Non-Anthropocentrism holds that nature is *also* intrinsically valuable. Given the importance of the intrinsic value to the debate between

---

[111] Norton (1984): 133.

Anthropocentrism and Non-Anthropocentrism, it is important to briefly recap what intrinsic value is.

In order to move from facts about the value of something to claim about what is morally right or wrong, Anthropocentrism and Non-Anthropocentrism both implicitly endorse what Regan (2004) calls "The Respect Principle," i.e., "We are to treat those individuals who have inherent [or intrinsic] value in ways that respect their inherent [or intrinsic] value."[112] The basic thought behind this principle is that when an object has certain properties, others are morally required to treat that object in certain ways.

In what ways is one required to treat something with intrinsic value? Regan (1981) focuses on how we should *not* treat intrinsically valuable entities. For example, he argues that we are required to not destroy, damage, interference with, or meddle with things that are intrinsically valuable.[113] It seems plausible to claim that one shouldn't destroy or damage things that are intrinsically valuable. For example, take the case of masterpieces of art. It seems to like we are required to not destroy, damage, or alter these works. In fact, even if a person technically owns one of these works, it still seems to make sense to claim that they are forbidden from destroying, damaging, or otherwise altering them.

Not only are we required to refrain from destroying, damaging, or altering things that are intrinsically valuable. It's also plausible that we are required or at least have reason to prevent them from being destroyed, damaged, or altered. For example, if I notice that a master work of art was in danger of being thrown away, stepped on, or painted over, I would have reason to intervene so as to prevent these things from happening.

---

[112] Regan (2004): 48. Norton (1984) thinks that what Regan means by "inherent" value is the same as what others mean by "intrinsic" value (137, fn. 11). I disagree. For a discussion of the difference between inherent value and intrinsic value, see Chapter 2. Regan (1981) refers to this principle as the Preservation Principle (31).

[113] Regan (1981): 31.

Moreover, one would also have reason to destroy, damage, or alter things that are only instrumentally valuable in order to prevent intrinsically valuable things from being destroyed, damaged, or altered. For example, if one could save a masterwork of art from getting wet by covering with a copy of a magazine, then one should do so—even if the instrumentally valuable magazine would get destroyed or altered. One would also have reason to use instrumentally valuable things to benefit things that are intrinsically valuable. For example, if I could prolong the life of a masterwork of year, but I had to destroy a piece of bad art (e.g., for the frame), then I would have reason to destroy the bad art in order to protect the masterpiece.

Finally, Regan regards the requirement to treat intrinsically value entities with respect as being a *prima facie* requirement in the sense that Ross thought we had certain *prima facie* duties.[114] That is to say that the Respect Principle holds that we have *pro tanto* reason to treat intrinsically valuable things with respect (e.g., by not destroying or damaging them).

We can now see how Anthropocentrism and Non-Anthropocentrism in conjunction with the Respect Principle tell us how we ought, generally speaking, to treat nature. Anthropocentrism says that only humans have intrinsic value and therefore are due respect. As we just saw, part of this respect means one has reason (and is generally permitted) to destroy, damage, or alter instrumentally valuable things in order to protect or benefit what is intrinsically valuable. Thus, it is not surprising that Anthropocentrists think that we can destroy, damage, or alter nature in order to protect or benefit humans. Non-Anthropocentrism says that nature also has intrinsic value and so is also due respect. This means that we have *pro tanto* reason to not destroy, damage, or alter nature. This holds even in certain cases in which doing so would benefit human beings.[115]

---

[114] Regan (1981): 32.

[115] Regan (1981): 32.

Given this understanding of the relevant views, it looks like Anthropocentrism and Non-Anthropocentrism are unlikely to exhibit much practical convergence. This is in part because Non-Anthropocentrism seems to be a direct denial of an important implication of Anthropocentrism: nature can be destroyed, damaged, and altered if it can be used to protect or benefit human beings. Thus, according to Anthropocentrism, anytime destroying, damaging, or altering nature can help humans, we ought do destroy, damage, or alter nature, but Non-Anthropocentrism only allowed for this kind of treatment of nature when doing so is necessary for sufficiently important goals. For example, if one had to cut down a single tree in order to save a million human lives, Non-Anthropocentrism would certainly permit it.

One remaining question is this: "What entity or entities have intrinsic value according to Non-Anthropocentrism?" According to individualism, the primary bearers of value are the individual entities that make up nature (i.e., each individual animal, plant, and so on). On this view, the value of nature *as a whole* is derived from the value of each of these entities.[116] According to holism, the primary bearer of value is nature as a whole and the value of each individual entity is derived from its contribution to the whole of nature.[117] I raise this question to clarify the possible views one might have, however, I will now set this question aside. This is because both Norton's and my arguments are independent of this debate. As I noted in the Introduction, I will use the word "nature" to refer to nature as a whole (e.g., the collection of all environments, ecosystems, species, and individual organisms) and parts of nature (e.g., environments and ecosystems).[118]

---

[116] Examples of such views include biocentrism (i.e., each living being is valuable in its own right) and animal rights (i.e., each animal is valuable in its own right). Defenses of individualistic views include: Singer (1990), Varner (1998), and Agar (2001).

[117] McShane (2009): 411. Examples of holistic views include: Leopold (1949), Katz (1985), and Callicott (1989).

[118] For a similar view, see Samuelsson (2008): 17-18.

3. Norton's Argument for Convergence

In this section, I will explain how Norton argues that, despite the fact that Anthropocentrism and Non-Anthropocentrism seem quite divergent on practical questions, these two views can come to exhibit total or mass practical convergence.[119] His main strategy is to provide a more nuanced version of Anthropocentrism, which he calls Weak Anthropocentrism, and argue that there is practical convergence between Weak Anthropocentrism and Non-Anthropocentrism. He calls the claim that Weak Anthropocentrism and Non-Anthropocentrism will exhibit mass or total practical convergence, "The Convergence Hypothesis."

Before doing this, however, it is important to distinguish two versions of the Convergence Hypothesis. Sometimes Norton seems to be making an empirical claim about what policymakers will do, while other times, he is clearly making a claim about what follows from Anthropocentrism and Non-Anthropocentrism. For example, in his 1991 book, *Toward Unity Among Environmentalists*, he writes: "Environmentalists believe that policies serving the interests of the human species as a whole, and in the long run, will serve also the "interests" of nature, and vice versa."[120] There are two empirical claims in this statement. First, there is a claim about what environmentalists believe. Second, and more importantly, there is the claim that the environmental policies that serve the interests of both current and future humans will also serve the interests of nature and vice versa. Call this the Empirical Convergence Hypothesis. I will not be concerned with this version of the Convergence Hypothesis, which has already received a fair amount of criticism.[121]

Other times, Norton is clearly making a philosophical claim about what two different views will entail. For example, he writes:

---

[119] McShane (2009) interprets him as arguing for total practical convergence (418, fn. 32).

[120] Norton (1991): 240.

[121] Callicott (2009) and Stenmark (2009).

> [P]rovided anthropocentrists consider the full breadth of human values as they unfold into the indefinite future, and provided nonanthropocentrists endorse a consistent and coherent version of the view that nature has intrinsic value, all sides may be able to endorse a common policy direction.[122]

While he hedges his claim here by writing that all sides "may" converge, he is more confident in his claims in other places. For example, he writes:

> The convergence hypothesis is a general, empirical hypothesis *about policy*—it claims that policies designed to protect the biological bequest to future generations will overlap significantly with policies that would *follow* from a clearly specified and coherent belief that nonhuman nature has intrinsic value [my emphasis].[123]

Despite the fact that he calls his hypothesis "empirical," it is clear that he is making a philosophical claim about what policies are entailed by Anthropocentrism. In particular, he's claiming that the policies entailed by Anthropocentrism will overlap with those entailed by Non-Anthropocentrism.

In his 2002 book, *Searching for Sustainability*, he writes:

> [I]f reasonably interpreted and translated into appropriate policies, a nonanthropocentric ethic will advocate the same policies as a suitably broad and long-sighted anthropocentrism.[124]

> According to that hypothesis [i.e., the Convergence Hypothesis], the interests of humans and the interests of nature differ only in the short run. If we recognize the extent to which the human species is an integral part of the community of life, long-term human interests *coincide* with the "interests" of nature. To protect the fullness of life is to protect the far-distant future of the human species and its evolutionary successors, and vice versa. Since the survival of our culture depends upon the survival of the ecosystems on which we, in turn, depend, the conception of the world one adopts is less important than the longsightedness with which it is applied in environmental management. [my emphasis][125]

These are clearly philosophical claims as they involve certain normative commitments. The first claim is a claim about what policies it would be *reasonable* for Non-Anthropocentrism to

---

[122] Norton (2009): 36.

[123] Norton (2009): 46.

[124] Norton (2002): 11.

[125] Norton (2002): 28.

support and a claim about how these policies would agree with a certain version of Anthropocentrism. The second is about what *interests* people and nature have. Call this the *Philosophical Convergence Hypothesis*.

In this chapter, I will only be concerned with the Philosophical Convergence Hypothesis. According to this claim, a particular version of Anthropocentrism (i.e., Weak Anthropocentrism) exhibits total or mass practical convergence with Non-Anthropocentrism.[126] Moreover, I will only have in mind the convergence on policies on which it seems that there is a conflict between human interests and nature's "interests." This is because Non-Anthropocentrists don't deny that nature has instrumental value because it can serve humans' needs.[127] And, when some environmental policy or action benefits humans without "harming" nature, then Non-Anthropocentrists and Anthropocentrists will agree that that action is permissible.

In order to show that Anthropocentrism and Non-Anthropocentrism can converge so that they recommend the same actions and environmental policies, Norton suggests a novel and more sophisticated version of Anthropocentrism that he calls "Weak Anthropocentrism." First, he views Weak Anthropocentrism as holding that only humans have intrinsic value and therefore only their interests have value. While he does not give an explanation of what he means by human "interests", he does argue that at least sometimes these interests are constituted by human desire satisfaction.[128] In other words, it is in humans' interest to have their desires satisfied. If the core of Weak Anthropocentrism is the value of human interests and the relevant human interests consist in desire satisfaction, then we need to analyse these desires and ask what they amount to. As we will see, humans might desire to treat nature in

---

[126] Norton (1984) and (2002): 11.

[127] McShane (2009): 408.

[128] Norton (1984): 134. Norton often explicitly talks about "preferences", but what he really means is desires (164).

exactly the same way as Non-Anthropocentrists think we're morally required to treat it (more on this below). If this is true, then there will likely be mass or total practical convergence between Weak Anthropocentrism and Non-Anthropocentrism even though there will still not be fundamental explanatory convergence.

Norton suggests that there are two kinds of human desires: felt desires and considered desires. A felt desire is a desire that doesn't arise out of reflection on one's needs or interests, but rather tend they are just the desires that come to us. Lee (2008) argues that these desires "represent survival interests and basic desires."[129] An example of a felt desire would be the strong basic desire for sensory pleasure, e.g., the strong desire to eat ice cream upon seeing a particularly delicious-looking cone of ice cream. One could have this felt desire even if it would be irrational for one to have it. For example, one might be lactose intolerant and thus eating the ice cream would lead to one suffering a terrible stomachache.

A considered desire, on the other hand, is a bit more complicated. According to Norton, these are the desires a human individual *would* express after careful deliberation, including a judgment that the desire is consistent with a rationally adopted world view - a worldview which includes fully supported scientific theories and a metaphysical framework interpreting those theories, as well as a set of rationally supported aesthetic and moral ideals.[130]

First, it is essential to note that considered desires are the desires that a person *would have* were they to go through a certain reflective or deliberative process.[131] Norton doesn't think that anyone has actually gone through this process and so considered desires are hypothetical desires. Second, while Norton provides some reason for thinking that one's considered desires will agree with the values of Non-Anthropocentrists, he doesn't provide

---

[129] Lee (2008): 12.

[130] Norton (1984): 164. Norton seems to be conflating preferences with desires and needs (and perhaps even interests). I mention this to show a possible limitation of his view.

[131] Norton (1984): 134.

any argument for why one needs to have a world view with "fully supported scientific theories and a metaphysical framework interpreting those theories."

While Norton's view of considered desires is somewhat opaque, I think a plausible way of glossing it is as follows. When a person desires something, then she must, in some sense, see or think of that object as being good or having some other evaluative property. A considered desire is a desire that a person would have if: (1) that person were fully informed about the all the non-evaluative facts that bear on the goodness of the object of desire, (2) that person had rationally supported moral and aesthetic beliefs, and (3) that person was reasoning correctly (e.g., without any bias) based on all the relevant non-evaluative facts. Whether or not anyone actually has desires that satisfy these conditions is not important for our purposes.

Before moving on to say why Norton thinks that these kinds of desires would lead one to prefer the same actions as Non-Anthropocentrism, it is essential we clarify one misleading aspect of Norton's account. Norton talks as if people only have unreflective desires. But that is clearly false. Every fully developed rational agent has reflective desires, i.e., actual desires based on *some amount* of reflection and deliberation. For example, my desire to donate money to one charity rather than another is such a desire. What is rarer is a reflective desire that satisfies the conditions mentioned above. Thus, we should distinguish between *partially considered* and *fully considered* desires. Partially considered desires are actual desires that are based on *some amount* of reflection and deliberation involving at least some of the relevant non-evaluative facts. Fully considered desires are *hypothetical* desires that meet the above conditions, e.g., they are based on all the relevant non-evaluative facts, unbiased reasoning, and rationally supported moral beliefs.

We can now distinguish Norton's Weak Anthropocentrism from what he calls "Strong" Anthropocentrism. According to Weak Anthropocentrism, how a person ought to treat nature is determined by their fully considered desires. According to Strong Anthropocentrism, how

a person ought to treat nature is determined by their felt/unreflective and partially considered desires.[132]

Now the important question for Norton is why we should think that a properly idealised person would have desires (i.e., fully considered desires) that would lead to them treating nature well in all or most of the same situations in which Non-Anthropocentrism would? Norton suggests that the idealised people could endorse an ideal of living in harmony with nature.[133] He argues that this ideal would be inconsistent with "the wanton destruction of other species or ecosystems even if the human species faces imminent extinction."[134] Even if we grant that some would have a fully considered desire to live in harmony with nature, it doesn't follow that many or most people would have that desire. After all, Norton only indicates that people *could* endorse this ideal, not that they *would* do so. Moreover, I find it rather implausible that even if people adopted an ideal of living in harmony with nature that they'd be willing to allow human extinction to protect nature. Finally, as we will see in the next section, it's plausible that at least some people would fail to have the fully considered desire to either live in harmony with nature or to treat nature well.

Fortunately, we can imagine three plausible reasons why Norton's idealised people would have fully considered desires to treat nature well or desires to protect nature from being destroyed or damaged. First, if one took the time to deliberate about what to do with nature, one would think about one's long-term needs or interests. In order for most people to live comfortable and healthy lives, nature must be healthy. For example, it would be bad for

---

[132] Here I'm providing what I think is a more plausible version of Strong Anthropocentrism than Norton (1984) presents. He writes, "Strong anthropocentrism, as here defined, takes unquestioned felt [desires] of human individuals as determining value." But it makes little sense to count unquestioned desires as mattering morally but not partially considered ones. Thus, I'm treating Strong Anthropocentrism as the more plausible view that both unconsidered and partially considered desires count morally.

[133] Norton (1984): 135.

[134] Norton (1984): 136.

the person in the long run if nature was self-destructing, e.g., through global warming, or if nature didn't provide certain essentials for the person, e.g., clear air. Thus, after rationally reflecting what is needed for one to live a long and comfortable life, Norton's rationally idealised people would desire to protect nature. Second, the idealised person might be risk-averse and think that we should be very careful about what we do to nature because it could have catastrophic results and, on the basis of being risk-averse, desire to treat nature well. Third, after rational reflection, the rationally idealised person might see that future generations have rights or interests that give him reason to protect nature for them. Given these thoughts and ideals, it would not be surprising if certain idealised persons had fully considered desires that were similar to the desires of a Non-Anthropocentrist. This is because the idealised person would take into account all the ways that interfering with nature now could plausibly negatively affect his life in the long term as well as the lives of future generations.

Thus, while there might be a large difference between what actions and policies Strong Anthropocentrism and Weak Anthropocentrism morally permit, there seems to be less of a difference between what actions and policies Weak Anthropocentrism and Non-Anthropocentrism morally permit. However, as I will argue below, it doesn't follow from the fact that Weak Anthropocentrism and Non-Anthropocentrism seem to make similar recommendations in some cases that it will make the same recommendations in *all* or *most* cases. Thus, as I will argue shortly, appealing to Weak Anthropocentrism is not sufficient for reconciling Anthropocentrism and Non-Anthropocentrism.

Before showing what is wrong with Norton's view, it will be helpful to clarify the practical convergence strategy he is appealing to. The strategy that Norton appeals is one that doesn't depend on explanatory to convergence. However, it is not the same as the two main strategies that we saw in Chapter 1. Recall that those two strategies were the following. The

first strategy was for one or both normative theories to adjust what they considered to be either a fundamental or derivative explanatory so that both theories proposed *similar* but not identical morally-relevant properties. For example, one could get some practical convergence between a consequentialist view and a deontological view by making the former concerned with minimising harm to innocent people and weighing minimising harm more than producing benefits. This would lead to some practical convergence because deontological theories tend to be concerned with preventing harm to innocent people. The second strategy was to show that one theory "mimicked" another one. For example, if God's commands gave one duties to act and God only commanded what Utilitarianism required, then God's Commands would overlap with Utilitarianism, but there wouldn't be explanatory convergence between there would just be two independent fundamental morally-relevant properties (God's commands and maximising utility).

Norton's strategy for convergence is a complicated view including elements of each of these strategies. First, the fundamental morally-relevant property, i.e., whether an action is right or wrong, depends on whether it is consistent with the content of people's fully considered desires. Norton thinks that people's fully considered desires include a desire to live in harmony with nature, a desire to preserve their lives in the long term or a desire to protect the lives of future generations. Thus, there is convergence because people desire to certain things (e.g., living in harmony with nature and protecting future generations) that overlap with the view that says that nature has intrinsic value.

Norton's strategy doesn't claim that the morally relevant properties of the two theories being converged are similar, so he isn't employing the first strategy. Nor is he claiming that one theory mimics another. In fact, he denies that nature has intrinsic value and so he couldn't argue that Weak Anthropocentrism mimics Non-Anthropocentrism.[135] Rather, his view is that

[135] Norton (1984): 142, 148.

idealised versions of people will have the fully considered desires to perform certain kinds of actions (e.g., actions that constitute living in harmony with nature, protect one's life in the long term, and protect future generations) and these actions are the kinds of actions that will be required by Non-Anthropocentrism.

4. Against Norton's Convergence Hypothesis, Part 1

In this section, I argue that Weak Anthropocentrism cannot be used to show that Anthropocentrism and Non-Anthropocentrism exhibit total or mass practical convergence. This is because Weak Anthropocentrism permits fully considered desires that are inconsistent with treating nature in the way that Non-Anthropocentrism requires.

The core problem is that people can have fully considered desires that are incompatible with treating nature in the way that Non-Anthropocentrism requires. For example, someone might grow up in an isolated community that teaches all the most up-to-date science and even plausible metaphysical views, but which also teaches that nature or the environment has no moral value at all. If everyone that a person knows (including all the smartest people one knows) claims that nature has no value, then it seems rational for one to believe that nature has no value. Notice that this view is also consistent with holding all the correct scientific and metaphysical views (at least all the non-normative metaphysical views) and it is supported by good evidence. A person in this social and epistemic situation can critically analyse the plausibility of various views about the value of nature and still end up thinking that nature is only valuable as a mere tool for satisfying felt and partial considered desires.

Importantly, Norton never claims that rationally adopting certain moral views entails holding the correct moral views. For example, Norton himself is sceptical of the claim that

nature has intrinsic value and so, he is not thinking that a person with fully considered desires will necessarily come to believe that nature has intrinsic value.[136]

Moreover, he seems to tactically assume there are multiple world views that one could rationally adopt. For example, he writes:

> Species preservationists can argue that their world view—its ontology, its epistemology, and its value of harmony with nature—is *a* rational response to the world . . . [my emphasis].[137]

So, it also looks like multiple sets of metaphysical, epistemic, and moral views can be rationally supported for Norton.

If it is true that there can be a variety of rationally supported moral ideals that give rise to fully considered desires, then there will be a large range of cases in which a person's fully considered desires will recommend actions that are morally forbidden by Non-Anthropocentrism. Or so I will now argue.

Before getting to my argument, it is important to recall two things. First, Norton seems to allow that different world views, and, in particular, different moral and aesthetic ideals, can be rationally adopted. So, he is not arguing that everyone's fully considered desires would be the same. Second, Norton does not provide any argument that limits the moral ideals that a person can rationally adopt. That is, he provides no reason to think that one could not rationally adopt a moral view on which nature has only the kind of instrumental value that Strong Anthropocentrists think it has.

The main problem with Norton's argument is that there can be rational disagreement between epistemic peers, i.e., people with the same evidence who are equally intelligent and rational, on moral issues. And, this seems to be true even after everyone has gone through the

---

[136] Norton (1984): 137. It is important to emphasise that Norton is sceptical about the claim that nature has intrinsic value/ This is because one might suspect that he is assuming that nature has intrinsic value and that sufficiently well-informed people would rationally come to believe that nature has intrinsic and that *this* is why Weak Anthropocentrism and Non-Anthropocentrism exhibit total or mass convergence.

[137] Norton (1987), 211.

kind of rational deliberation that Norton thinks is important for specifying fully considered desires.

One piece of evidence for this is the epistemic peer disagreement between moral philosophers about just about any moral issue. I take it that many of these philosophers have rationally adopted the same or similar scientific and metaphysical views, but they also rationally adopt different moral views. For example, it seems plausible to me that many moral philosophers have rationally adopted some form of consequentialism while others have rationally adopted some version of non-consequentialism or deontology.

Perhaps Norton will object that these philosophers haven't gone through as rigorous a process of self-criticism or deliberation as the one is he imagining. However, Norton provides no argument for thinking there will be a crucial difference between the views of moral philosophers, who have spent a lifetime defending some moral view, and the kind of idealised person who has engaged in rational deliberation and self-criticism.

If at least some of these moral philosophers meet the essential criteria for having fully (or nearly fully) considered desires—or criteria that are not far off from this essential criteria —then it is possible to have fully considered desires that conflict with the requirements of Non-Anthropocentrism. This is because most normative ethical theories either do not entail anything about the moral status of nature or they entail that the nature has no moral status. So, if these moral principles are rationally held by moral philosophers, it would be rational for these philosophers to doubt the moral importance of nature for its own sake.

If we look through the normative ethical principles that we discussed in Chapter 1, we will see nature doesn't seem to have any of the properties that these theories think is morally-relevant. Various versions of Utilitarianism are primarily concerned with pleasure or desire satisfaction and happiness. But nature is not capable of experiencing pleasure, having desires, or being happy. While we didn't discuss it in Chapter 1, Kantians are often concerned with

protecting or respecting autonomy or rationality. But nature is not autonomous or it cannot be rational. Contractualism is concerned with the agreement between well-informed and suitably motivated people living together. But nature can't be rational or well-informed and it cannot agree (actually or hypothetically) with humans. Deontological theories are concerned with preventing harm to innocent people. While nature can be harmed, e.g., by being destroyed, deontology is concerned with the kind of harm that only conscious and sentient creatures suffer, e.g., pain, manipulation, humiliation, etc.

Finally, it's not clear that any of Ross's *prima facie* duties apply to nature as well. Recall that his duties are:

> *Fidelity*: One should keep one's promises and be honest and truthful.
>
> *Reparation:* One should make amends when one has wronged someone else.
>
> *Gratitude:* One should be grateful to others when they perform actions that benefit one and one should try to return the favour.
>
> *Non-injury (or non-maleficence):* One should refrain from harming others either physically or psychologically.
>
> *Beneficence:* One should be kind to others and improve their health, wisdom, security, happiness, and well-being.
>
> *Self-improvement:* One should improve one's own health, wisdom, security, happiness, and well-being.
>
> *Justice:* One should be fair and distribute benefits and burdens equably and evenly.

One cannot lie or fail to keep a promise to nature, one cannot make reparations to nature, one cannot be grateful to nature,[138] self-improvement doesn't apply to nature, nature is also not the kind of thing that one could be fair toward or disturb benefits to. It might be that one could exhibit some degree of beneficence to nature, but it's clear that the kind of beneficence that Ross has in mind is only available to sentient beings. Notice that much of beneficence is concerned with wisdom, happiness, and well-being and it is clear that nature cannot have

[138] Manela (2018).

wisdom, happiness, or well-being. Perhaps it can be healthy or more secure, but, again, Ross was thinking of the kind of health and security that can only be enjoyed by sentient creatures. For example, health involves freedom from suffering, stress, disease, and so on. Likewise, when Ross is discussing Non-injury, he's imagining physical injury that is unpleasant to experience.

Given that it is far from clear that nature has any of the morally-relevant properties that the most prominent normative ethical views posit, it seems that the people that rationally hold these views would be rational in either disbelieving that nature has any moral status or being sceptical about it. And, if they can rationally disbelieve or doubt that nature has any moral status, then it would be rational for them to not have a desire to protect nature for its own sake—only when doing so would help humans.

Of course, this doesn't mean that it is rational for *everyone* to doubt the moral importance of nature. After all, there are certainly moral philosophers who think that nature is worthy of moral consideration or at least that it should be treated well—even for anthropocentric reasons.

Thus, if it is rational to doubt whether nature has a moral status in its own right, it is likely that at least some of the rational and well-informed agents that Norton discusses will not have desires that line us with the recommendations or requirements of Non-Anthropocentrism. In fact, in at least some of these cases, it is likely that people will have desires that align with Strong Anthropocentrism. This is because they will rationally believe that nature doesn't have a moral status of its own, but humans do, and so nature can and should be used to help humans.

There is even disagreement amongst Non-Anthropocentrists about what has moral value. Biocentric Non-Anthropocentrists claim that only the *living* parts of nature (including

humans) have value and thus ought to be protected.[139] Ecocentric Non-Anthropocentrists claim that in addition to living parts of nature, certain non-living parts of nature (e.g., species and ecosystems) have value and thus ought to be protected.[140] Within each of these views (i.e., Biocentric Non-Anthropocentrism and Ecocentric Non-Anthropocentrism) there is disagreement about *how* valuable certain entities are. As Stenmark (2009) argues, some biocentrists think that all living creatures have *equal* value,[141] some think that humans and certain non-human animals have equal value,[142] and some think that humans have more value.[143] The same holds for ecocentrists. Some think that ecosystems have the most value. For example, Callicott (1989) writes, "In every case the effect upon ecological systems is the decisive factor in the determination of the ethical quality of actions."[144] Some ecocentrists hold that humans have a greater value than other parts of nature (e.g. wildlands).[145]

The people having these disagreements are lifelong environmental ethicists and defenders of versions of Non-Anthropocentrism. If they can rationally disagree about what entities have value and how much, then it seems like there can be rational peer disagreement about these issues. Of course, these ethicists are not Norton's idealised person who has engaged in rational deliberation and self-criticism. However, as I noted above, they are likely the closest thing we have to such idealizations. And, if there can be such rational disagreement about what kinds of entities have value, then it's quite plausible that there will not be total or mass practical convergence between Weak Anthopocentrists and Non-

---

[139] Stenmark (2009): 83.

[140] Stenmark (2009): 83.

[141] Taylor (1986): 129.

[142] Regan (2004): 240.

[143] Grenholm (1997): 93-94.

[144] Callicott (1989): 21.

[145] Rolston (1988): 226.

Anthropocentrists. After all, there isn't even total or mass convergence between certain Non-Anthropocentrists.

The fact that Weak Anthropocentrism could recommend actions that conflict with the actions that Non-Anthropocentrism recommends should not be surprising. First, Norton only provides a few cases to indicate that Weak Anthropocentrism and Non-Anthropocentrism recommend the same action and a few is clearly not sufficient for showing that the views will *always* or *mostly* recommend the same action.[146] Second, it is widely accepted that there are many moral questions about which there is *rational* peer disagreement among philosophers. That is, there are people who are just as smart and well-informed who disagree what about has value and how much value it has. Therefore, it is not surprising that at least some people's fully considered (and so rational) desires would conflict with desires that Non-Anthropocentrists endorse.

5. Against Norton's Convergence Hypothesis, Part 2

In this section, I argue that even if Norton is correct that Weak Anthropocentrism and Non-Anthropocentrism exhibit total or mass convergence, many Anthropocentrists and Non-Anthropocentrists are unlikely to be satisfied with his argument, because it doesn't resolve a deeper and more important debate about what properties are right-making and wrong-making. That is, even if there is mass or total practical convergence between Weak Anthropocentrism and Non-Anthropocentrism, proponents of each view will think this convergence is, in a sense, superficial. This is because if there is practical convergence without any non-fundamental or fundamental explanatory convergence, then an important disagreement between the views remains.

To be clear, Norton does not claim that there is any explanatory converge between Weak Anthropocentrism and Non-Anthropocentrism. So, I'm not arguing that any particular

---

[146] I thank Robert Cowan for pointing this out to me.

argument of Norton's is flawed. Rather, I want to point out an important *limitation* of his argument. The idea is that, even if there is practical convergence, an important kind of convergence is ignored.

There are two broad ways of getting two normative ethical principles to converge. The first involves showing that there is some kind of explanatory convergence and the second involves merely showing that the two principles agree about the moral status of actions even though they disagree about the reasons why the actions have the moral statuses they have. Even if Norton's argument were successful, all it could show is that Anthropocentrism and Non-Anthropocentrism exhibit practical convergence without any kind of explanatory convergence. This is because their explanations of why treating nature in certain ways are either right or wrong are very different. After all, humans' fully considered desires and the intrinsic value of nature are quite different properties. However, what proponents of normative ethical principles really care about is *why* actions are right or wrong. For example, take Utilitarians who think that an action is morally right if and only, and because, it maximises pleasure. What they care about is whether an action maximises pleasure and that is it. If Utilitarianism and some form of deontology agree that some action is morally obligatory, they still disagree on the most important question, i.e., "Why is it morally obligatory?"

In addition, Weak Anthropocentrism and Non-Anthropocentrism do not exhibit *any* explanatory convergence. Some normative ethical principles exhibit some degree of explanatory convergence because they agree that some property partly explains why an action is right or wrong, but they disagree about *how much* that property contributes to making it right or wrong. For example, imagine that Margaret tells Elizabeth a lie in order to hurt Elizabeth's feelings. Two normative ethical theories might agree that it was morally wrong for Margaret to say what she said because it was a lie, but they might disagree about how

much the fact that the statement was a lie contributed to the wrongness of the statement. One theory might say that it was sufficient for making it morally wrong for Margaret to say what she said, but the second theory might say that it contributed a little bit to making it morally wrong, but was not sufficient on its own for making it morally wrong. So, both of these principles agree that lying contributes to making an action wrong, but they disagree about how much it contributes.

However, Weak Anthropocentrism and Non-Anthropocentrism do not exhibit any kind of explanatory convergence. In fact, from the Non-Anthropocentric point of view, it is at least sometimes irrelevant whether humans desire that nature be treated with respect. Likewise, from the Weak Anthropocentrism's point of view, it is irrelevant whether nature has intrinsic value. All that matters are the fully considered desires of people. In fact, as I noted earlier, Norton himself is sceptical of the claim that nature has intrinsic value.[147]

We can get a better sense of the disagreement and why mere practical converge will not satisfy either Weak Anthropocentrists or Non-Anthropocentrists, by considering a disagreement in another normative domain, i.e., aesthetics. Imagine that two art critics both think that a painting is beautiful. The first critic thinks that the painting is beautiful because of the gracefulness of its lines and brush strokes and while the second critic thinks that the painting is beautiful because of its evocative colour palette. Imagine further that the first critic doesn't think that the colour palette at all contributes to making the painting beautiful while the second critic doesn't think the lines or brush strokes at all contribute to making the painting beautiful. So while there is a kind of superficial convergence on the question of whether the painting is beautiful, there is a deep disagreement remaining between the critics. It is this kind of deep disagreement that also remains between Weak Anthropocentrism and Non-Anthropocentrism.

---

[147] Norton (1984).

Norton might reply that there is *some* amount of explanatory convergence because the reason that Norton's hypothetical agent desires to treat nature well is that they believe that nature has intrinsic value. Recall that fully considered desires are not merely desires that issue from deliberation and which are consistent with one's rationally adopted world view. They must also be judged to be consistent with a world view that has "fully supported scientific theories and a metaphysical framework interpreting those theories, as well as a set of rationally supported aesthetic and moral ideals."[148] That is, this world view is made up of beliefs, values, concepts, etc. which must be informed by reasonable scientific theories which are consistent with a plausible metaphysics. Finally, one must also have rationally held aesthetic and moral views. If one has rationally held moral views, then one might essentially endorse some version of Non-Anthropocentrism.

However, this response will not work. On one interpretation of this response, Norton is just claiming that Non-Anthropocentrism is true and that his hypothetical rational agent endorses it. But then it makes little sense to claim that the reason that nature has any moral standing is that hypothetical rational agents prefer that nature be treated well. Or, at the very least, it is unclear that this is a kind of explanatory convergence.

But perhaps the idea is a bit more complicated. Perhaps the idea is this. First, Non-Anthropocentrism actually consists of two claims: (1) nature is intrinsically valuable and (2) because nature is intrinsically valuable, people are morally required to respect it and therefore treat it well. Second, Weak Anthropocentrists argue that rational and well-informed agents will agree with (1), i.e., that nature is intrinsically valuable, but they will disagree with (2), i.e., because nature is intrinsically valuable, people are morally required to respect it. Instead, Weak Anthropocentrism will argue that rational and well-informed agents will think that

---

[148] Norton (1984): 134. Norton seems to be conflating preferences with desires and needs (and perhaps even interests). I mention this to show a possible limitation of his view.

nature is intrinsically valuable and on this basis they will desire to treat it with respect. And, therefore, the only reason that nature is worthy of respect is that rationally and well-informed agents would desire that it be treated this way.

Norton might then claim that both Weak Anthropocentrism and Non-Anthropocentrism agree with the following statement: One is morally required to treat nature well *because* it is intrinsically valuable. But this statement hides an essential disagreement between Weak Anthropocentrism and Non-Anthropocentrism. For Non-Anthropocentrism, the fact that nature is intrinsically valuable *makes it the case* that one is morally required to treat nature well. However, for Weak Anthropocentrism, the fact that nature is intrinsically valuable *doesn't* make it the case that one is morally required to treat nature well. Rather, it is only the fact that certain people would *desire* to treat nature with respect that it makes it the case that one is morally required to treat nature with respect. The reason why rational and well-informed people have this desire is irrelevant for making it the case that people are morally required to treat nature well. If they desired that nature be treated with respect because they think nature is aesthetically pleasing, it would equally follow that one is morally required to treat nature well. Thus, while Weak Anthropocentrism and Non-Anthropocentrism agree in some sense that the fact that nature is intrinsically valuable, they don't agree that this fact is a fundamental right-making or wrong-making property. Therefore, there is no fundamental explanatory convergence between them.

On a second interpretation, Norton is claiming that regardless of whether Non-Anthropocentrism is true, the *evidence* that his hypothetical rational agent would have would indicate that it is true. But now we have the question of whether or not the truth of Non-Anthropocentrism *explains* the evidence that Non-Anthropocentrism is true. If it does, then Non-Anthropocentrism is true and Norton faces the same problem he faces with the first interpretation. If the truth of Non-Anthropocentrism doesn't explain why the evidence

supports Non-Anthropocentrism, then it must be mere luck or accident that the evidence supports it. If this is true, then it is unclear in what way Weak Anthropocentrism and Non-Anthropocentrism exhibit any explanatory convergence.

Therefore, even if Norton had successfully argued that Weak Anthropocentrism and Non-Anthropocentrism exhibit total or mass convergence, proponents of both views would still be unsatisfied. So, Norton doesn't resolve the fundamental disagreement between these views, i.e., the explanation of why certain actions concerning nature are morally right or wrong.

6. Broad Anthropocentrism and Convergence

In the last few sections, I argued that Weak Anthropocentrism and Non-Anthropocentrism do not exhibit total or mass practical convergence nor do they exhibit any explanatory convergence. In this section, I will sketch a new version of Anthropocentrism, what I will call "Broad" Anthropocentrism. I will argue that this version of Anthropocentrism allows that some parts of nature are finally valuable. I will then argue that Broad Anthropocentrism and Non-Anthropocentrism both agree that some parts of nature are finally valuable. And, because they agree about the fact that nature is finally valuable, they will agree on how to treat nature (or parts of nature) *and* they will also agree, to some degree, on *why* we ought to treat nature well, i.e., because it is finally valuable.

6.1 Extrinsic Final Value

In order to show how Broad Anthropocentrism exhibits non-trivial practical and explanatory convergence with Non-Anthropocentrism, I need to show that there is reason to think that nature is finally valuable. However, in order to do this, I need to show how something that started off as being instrumentally valuable can come to be *extrinsically finally valuable*.

Recall from Chapter 2, that many philosophers have come to see that something can go from being instrumentally valuable to being finally valuable in virtue of some set of its

external or relational properties. One prominent case of this is the pen that Abraham Lincoln used to sign the Emancipation Proclamation. Notice that this pen is certainly valuable and that at least some of its value comes from what it was previously used for and not just what it can be used for now. In particular, some of its value seems to come from the role it played in an important historical event that resulted in the saving and bettering of millions of innocent human lives. But the property of being used by a particular person for a particular signing is a relational property and so the final value of Lincoln's pen must be based on extrinsic (i.e., relational) properties.[149] Of course, the pen also had instrumental value for Lincoln because he could use it to write with. Moreover, it currently has instrumental value because it can be used to remind people of the value of human freedom.[150] But the fact remains that, even if it couldn't be used for these purposes now, it would still be valuable. So, it looks like it is finally valuable, i.e., valuable in its own right, in virtue of some relational properties. So, it looks like there are things that are extrinsically finally valuable.

The case of Lincoln's pen shows us something important about how instrumental value can give rise to final value. There are two kinds of instrumental value: potential and actual. First, something can be instrumentally valuable in virtue of its *potential* to bring about something else of value. In other words, this thing hasn't brought about anything of value yet, but, at the moment, it has the potential to. This is the kind of value that any of Lincoln's pens had *before* he signed the Emancipation Proclamation. Any one of these pens has the *potential* to be used to sign the proclamation. Second, something can be instrumentally valuable in virtue of *actually* having brought about something with final value. The pen that Lincoln actually used to sign the Emancipation Proclamation actually helped to bring something with final value about (i.e., the freeing of American slaves). So, Lincoln's pen actually brought

---

[149] Kagan (1998): 285.

[150] Kagan (1998): 286.

about something of final value. Moreover, it will always be true that Lincoln's pen brought about something with final value. That is, it is a permanent feature of Lincoln's pen that it played this crucial role in a value event in human history. This helps explain why we can appropriately value the pen for its own sake, i.e., because this pen will always have the property of being of playing an important role in freeing American slaves.

The value of Lincoln's pen is also increased by the fact that it is the *only* pen that played a crucial role in freeing the slaves. Had he used 100 different pens, each to write a small part of his signature, it is unclear that any one of these pens would be as valuable as the one pen he actually used. So, the value of Lincoln's pen is partly due to its uniqueness.

Moreover, Lincoln's pen not only brought about something of final value. It brought about something with *tremendous* final value, i.e., freeing millions of human lives from slavery. So, for example, Lincoln's pen is importantly different from the pen I used to help gain knowledge by taking notes in class. Knowledge is finally valuable, but most individual pieces of knowledge are not tremendously valuable. So, while my pen played a role in bringing about something with final value, it didn't play an important role in bringing about something with tremendous final value.

Finally, Lincoln's pen helped to bring about something that benefited all human beings and made the world a better place. Of course, it benefitted the slaves that were freed and anyone who cared about them. But it also benefited, albeit indirectly, everyone by making the world a more just place. It even benefited the previous slave owners by preventing them from continuing to engage in seriously immoral behaviour.

So, Lincoln's pen is a case in which something that started out as an object with instrumental value went on to have final value. But the reason for this is that it actually brought about something that was not only finally valuable, but tremendously valuable in the sense that it benefited all humans. So, it looks like something can be finally valuable if the

following features obtain: (1) the finally valuable object played an important role in actually bringing about something else with final value, (2) the finally valuable object *uniquely* played an important role in actually bringing about something else with final value, (3) the thing it brought about was tremendously valuable, and (4) the thing it brought about benefited all human beings.[151]

The question I will answer in the next subsection is this: can Anthropocentrism claim that nature (or some parts of nature) is valuable in a way similar to Lincoln's pen (i.e., extrinsically finally valuable)? I will argue that they can.

6.2 Broad Anthropocentrism

Many philosophers, including Norton, seem to interpret the core value claim of Anthropocentrism as being that nature is only valuable *as a means* to serving current or future human interests. But this is not the core of Anthropocentrism. The core value claim of Anthropocentrism seems to be that whatever value nature has it has in virtue of its relationship(s) with humans or human interest. But the relationship need not be *only* instrumental. That is, nature need not be seen only as a tool for benefitting current or future human interests. In fact, I will argue that there is a broader view of Anthropocentrism, which shows the value of nature (or its parts) can derive from the important role it has played in our lives and in the lives of our ancestors. In particular, I argue that nature has uniquely played a crucial role in sustaining the lives of all human beings (ourselves included!). Thus, just as Lincoln's pen has extrinsic final value in virtue of playing a unique role in bringing about something of tremendous, final value, which benefitted all humans, so too, nature has played a unique role in bringing about something of tremendous, final value, which benefitted all humans.

---

[151] Perhaps something weaker would work for this condition. Perhaps the finally valuable thing only has to benefit a large group of humans. My arguments later on will work just as well even if this condition is weakened.

Nature has played a crucial role in the sustaining of human life. It's undeniable that many ecosystems provided the resources necessary for humans to survive. One of the main reasons our ancestors were able to survive is the ecosystems that they lived in provided them with the necessary means of survival. These ecosystems provided hospitable conditions for life. For example, they were not too hot or too cold, they had shaded areas to protect people from overexposure to the sun, the air quality was good enough, and they didn't have too many natural disasters (e.g., earthquakes, typhoons, hurricanes, tornadoes, flooding, snowstorms, etc.). These ecosystems provided necessary means for survival. For example, they provided breathable air (unlike in the ocean or around active volcanoes), drinkable water (via rain or lakes, ponds, rivers, and streams), food (via plant and animal life), herbs or plants for medicinal use, soil in which new crops can be grown, materials for building shelter or tools, materials for staying comfortable (e.g., firewood), places for shelter (e.g., in caves), and water for bathing.

Moreover, nature has provided many of these very same benefits to us. Of course, we have more help in our lives from human innovation and technology. But we still rely on nature for oxygen, water, food, medicine, soil to grow crops, materials for building shelter and tools, and so on.

While there is not enough space to describe all the ways that nature helps in these ways, a few details will help bring out exactly partly ecosystems have helped to sustain human life. Forests can store carbon, which prevents more carbon dioxide from entering the atmosphere and contributing to global warming.[152] Wetlands have filter water to remove sediments, which can be harmful to both plants and animals.[153] Coral reefs and mangroves

---

[152] Wildlife Conservation Society (2020).

[153] Rogers (2019).

help protect coastal areas from waves, storms, and floods.[154] Up to 35% of the world's food crops rely on animals (e.g., birds, bees, butterflies, beetles, and so on) for pollination.[155] In these ways, and more, nature (e.g., ecosystems and nature areas like forests) have benefited and continue to benefit all humans.

Given its crucial role in helping to sustain human life, it's clear that nature helped sustain something with final value (i.e., human life).[156] Lincoln's pen was mainly responsible for *bringing about* as opposed to *sustaining* something of final value. However, it is not clear that this difference is important. In fact, it is often thought to be morally better to sustain a life than to create a new one. And while we have no duty to bring about new lives, we do have a duty to save lives—under normal circumstances.[157] Moreover, it should be clear that Lincoln's pen also played a role in sustaining the lives of the slaves that were freed as well.

Nature's role in helping to sustain human life is also unique. That is, nothing else is responsible for making the same kind of contribution to the sustaining of human life that nature has. This is not to say that nothing else contributed to the sustaining of human life. After all, advances in medicine and medical technology have clearly helped sustain human life. However, they have helped in different ways. They didn't create oxygen (although they might have helped to deliver it to a patient), they didn't provide climates and weather that are suitable for human life, they didn't provide fertile soil, they didn't create water, they didn't create plants or animals for eating, and so on.

Nature played a unique role in helping to sustain something of *tremendous* value, i.e., the lives of innumerable humans being—including all currently-living humans. So, what

---

[154] NOAA (2021).

[155] United States Department of Agriculture (nd).

[156] Insofar as one thinks that certain non-human animals' lives are finally valuable, then nature also played a crucial role in sustaining even more things with final value (e.g., see Regan (2004)).

[157] Roberts (2011).

nature is responsible for sustaining is not just something with final value (e.g., knowledge), but something with tremendous, final value.

Finally, the finally valuable thing that nature played a unique role in sustaining benefited all human beings. All of us have benefited from the climate, food, water, shelter, medicine, and materials that nature has provided that have helped us remain alive. Moreover, nature has benefited us by providing the same kind of benefits to our ancestors and insofar as a benefit to them is a benefit to us, nature has doubly benefitted us. To claim that nature has been all humans is not to claim that nature has benefited all humans equally. Nor is it a claim that nature has provided an *overall* benefit to people. Perhaps nature has benefited people in some ways and harmed them in others. For example, perhaps nature provided the means of eating and drinking for someone, but that person was eventually harmed by a tornado. All I'm claiming is that nature has provided *some* benefit to all humans. This is true of Lincoln's signing of the Emancipation Proclamation. While it was an overall benefit for most people, it paved the way for new forms of mistreatment of black Americans.[158]

Thus, like nature seems to share the properties with Lincoln's pen that account for why the latter went from being potentially instrumentally valuable to finally valuable. First, nature played an important role in sustaining something of final value. Second, it played a unique role in doing this. Third, the finally valuable thing it sustained was tremendously valuable. And, fourth, the finally valuable thing it sustained benefited all humans. Thus, insofar as Lincoln's pen is extrinsically finally valuable, so too is nature. However, because the source of nature's final value is solely based on its previous or current relationship with humans, this view is completely Anthropocentric.[159]

---

[158] Forde and Bowman (2017).

[159] Tenen (2020) argues that nature (or some of its parts) can have extrinsic final value, but he does not focus solely on cases in which the source of the final value is nature's relationship with humans. For example, he notes that the age and rarity of parts of nature might give it extrinsic final value.

To clarify, none of what I've written here implies that nature is also potentially instrumentally valuable or actually instrumentally valuable in ways that fail to make it finally valuable. This is also true of Lincoln's pen. Nature is instrumentally valuable in this way because of the joy, wonder, and awe that it has brought and can bring to people just as Lincoln's pen is instrumentally valuable because has reminded (and continues to remind) people of the value of freedom.

7. Converging Broad Anthropocentrism and Non-Anthropocentrism

In the last section, I argued that a form of Anthropocentrism (i.e., Broad Anthropocentrism) can claim that nature is extrinsically finally valuable. In this section, I will argue that Broad Anthropocentrism and Non-Anthropocentrism exhibit a non-trivial amount of both explanatory and practical convergence.

7.1 Derivative Explanatory Convergence

Recall that Broad Anthropocentrism holds, among other things, that nature (or parts of it) is extrinsically finally valuable. What matters for my argument that Broad Anthropocentrism and Non-Anthropocentrism exhibit a few kinds of convergence is that the former holds that some parts of nature are *finally* valuable. This is important because Non-Anthropocentrism *also* holds that nature is finally valuable. This is because, as I noted above, when something is intrinsically valuable, it is finally valuable in virtue of some set of non-relational properties. This means that Anthropocentrism and Non-Anthropocentrism agree that nature (or at least some parts of it) is finally valuable.

Because Broad Anthropocentrism and Non-Anthropocentrism agree that nature is finally valuable, they exhibit some amount of derivative explanatory convergence, i.e., they agree about a non-fundamental explanation of why we should treat nature well. This is because, the *reason why* people should treat certain parts of nature well is that these parts of nature are finally valuable, i.e., valuable for their own sake.

Because Broad Anthropocentrism and Non-Anthropocentrism agree that nature is finally valuable, they will also agree in many cases that one should treat nature well. This is because both Broad Anthropocentrists and Non-Anthropocentrists can agree on something like the following principle proposed by Regan (1981):

> *The Preservation Principle*: If X is finally valuable, then one has a *pro tanto* moral reason to refrain from destroying, interfering with, or meddling with X.[160]

To be clear, this principle concerns what one has *pro tanto* reason to refrain doing and thus is kind of *prima facie* duty, to use Ross's terminology from Chapter 1. This principle, or something like it, is how Broad Anthropocentrists and Non-Anthropocentrists can link their claim about what kind of value nature has to claims about what one morally ought to do concerning nature.

This means that Broad Anthropocentrism and Non-Anthropocentrism exhibit some non-trivial amount of practical convergence as well—at least as long as they endorse some principle like The Preservation Principle. Moreover, this amount of practical convergence will not be the kind of trivial accidental convergence mentioned above. It will not be non-accidental because, at least some of the time, they will agree on why one should treat nature well. As I argued in Chapter 1, the most plausible strategy for arguing for practical convergence is by arguing for derivative explanatory convergence. The key to getting practical convergence between a version of Anthropocentrism and Non-Anthropocentrism was to show how the former could some to agree with the latter that nature has final value.

One might object that Broad Anthropocentrism and Non-Anthropocentrism disagree about something rather important, i.e., they disagree about whether nature is *intrinsically* valuable. However, I think that it is quite clear that what matters for how we should treat

---

[160] Regan (1981) argues for this principle as if it concerns only objects with "inherent" value. However, it is clear from what he writes that he has final value in mind. For example, he writes that if X has "inherent" value it has value "in its own right" and "not merely as a resource to be used in the name of human interests" (31).

nature has little to do with the kind of properties (relational vs non-relational) that give nature its final value. What seems to matter is how one should treat X is whether X is finally or instrumentally valuable and *how* valuable X is.

First, if X is only valuable as a causal means to Y and there are many other causal means to Y, then it is not clear why one should be concerned with preserving or not interfering with X. If, however, X is finally valuable, then it seems like it matters how you treat X. This is because X matters for its own sake or in its own right.

Second, something can be instrumentally *very* valuable. For example, think about the cure for a deadly virus, a film that most people find very entertaining, or a rescue plan for hostages. It seems very important to protect these things even though they are only instrumentally valuable. On the other hand, some things can be intrinsically *not very* valuable. For example, knowledge is thought to be intrinsically valuable, but knowledge of a random fact, such as knowing that one's favourite restaurant purchased 1,250 napkins last week, doesn't seem to be very valuable. Achievements are also supposed to be intrinsically valuable. But some achievements don't seem to be very valuable, e.g., holding the record for the world's longest hair or fingernails.

Also, as I argued in in Chapter 2, it is fairly clear that when many Non-Anthropocentrists claim that nature has intrinsic value, all they mean is that it has final value. For them, what really matters is showing that nature (or its parts) is not only *instrumentally* valuable. That is, they want to show that nature is not only valuable because of how it can be used to help human beings. They do not argue that nature's value must come from some set of its non-relational properties.

It is still true that Broad Anthropocentrism and Non-Anthropocentrism don't exhibit any amount of fundamental explanatory convergence. This is because they disagree about why nature (or its parts) has final value and therefore why it ought to be treated well. Broad

Anthropocentrism holds that the fundamental explanation of why nature ought to be treated well is its important relationship with humans. Whatever Non-Anthropocentrism holds as the fundamental explanation of why nature ought to be treated well, it cannot be nature's relationship with human beings. Thus, these two views will not exhibit any amount of fundamental explanatory convergence.

7.2 Degrees of Value

A related worry concerns whether Broad Anthropocentrism and Non-Anthropocentrism exhibit a non-trivial degree of convergence on the question of *how* (finally) valuable certain parts of nature are. One might worry that even if Broad Anthropocentrism and Non-Anthropocentrism agree that certain parts of nature are finally valuable they will disagree about *how* valuable they are and therefore issue in different recommendations on when to treat these parts of nature well. Take the example of a particular ecosystem that has played a crucial role in keeping some group of people alive, e.g., by providing rainwater, plants to eat, and animals to eat, but no longer plays this role. According to Broad Anthropocentrism, this ecosystem is very valuable because it was crucial in protecting the lives of many human beings. In fact, we might imagine that were it not for this ecosystem, many humans that are alive today would not have been alive. However, some version of ecocentric Non-Anthropocentrists might claim that all ecosystems are valuable, but not very valuable. If this is the case, then they will disagree about when to treat nature well when there are countervailing considerations. For example, if we have to choose between protecting this ecosystem and another ecosystem that didn't benefit humans much, the aforementioned Non-Anthropocentrist might think that we can pick which one to help at random (e.g., by flipping a coin). But the Broad Anthropocentrist will say that we ought to protect the one that had *previously* benefited humans to a great degree.

While I cannot provide a fully convincing argument that there will be a good amount of convergence between Broad Anthropocentrists and Non-Anthropocentrists on the value of nature or some part of nature, I think there is reason to be optimistic. In particular, I'm going to argue that both views converge on the claim that nature is "very" valuable. Of course, "very" refers to a wide range of values and so my claim is quite modest. To argue for something more precise would require making many controversial assumptions about how to precisely measure value and such a discussion is beyond the scope of this thesis.

In order to show that Broad Anthropocentrism and Non-Anthropocentrism will exhibit some non-trivial amount of convergence on the question of how valuable nature is we need to figure out what determines the amount of nature's value according to each view.

As we saw above, Broad Anthropocentrism claims that what makes nature finally valuable is the fact that it played and continues to play a unique and important role in the sustaining of some tremendously valuable such that all humans are benefited. That is, because of its special role in sustaining human lives. How valuable nature (or some part of it) is therefore depends on a few factors. First, how much nature contributed to sustaining human life, e.g., how many resources like oxygen and food it supplied. Second, how unique nature is as a source of these resources. Third, how valuable human life is. Fourth, how much all humans were benefited. All of these factors are a matter of degree and so how valuable nature is according to Broad Anthropocentrists will be a function of these factors.

I think Broad Anthropocentrists will hold that nature is very valuable. Here is why. It seems clear that nature contributed and continues to contribute many resources for sustaining human life (as enumerable above). Nature seems to be the unique source of many of these resources (e.g., sunlight, water, plants, medicine, building materials, etc.) Human life is obviously very valuable. All humans have benefited a lot from nature. So, it seems rather

clear that Broad Anthropocentrists will think that nature is very valuable. What about Non-Anthropocentrists?

While there are a variety of properties that Non-Anthropocentrists have argued make nature valuable, I will focus on Samuelsson (2008)'s view because it incorporates many of the features into a single view. Samuelsson argues that one thing that makes nature finally valuable is that is "complex".[161] Being complex in the relevant sense is to possess a cluster of interrelated features: intricacy, integrity, self-going or self-running, and interacting. All of these features can be hard in degrees and so, under normal circumstances, the higher the degree to which nature has these features and therefore the more "complex" it is, the more valuable it is. In order to see the plausibility of this view, let me outline what Samuelsson means by intricacy, integrity, self-going, and interacting.

First, nature is extremely intricate, i.e., extremely complicated. In fact, as Samuelsson notes, nature is so intricate that humans cannot come close to create things that accurately imitate it. In other words, nature is inimitably intricate.[162] Second, nature has integrity. That is, it is a self-managing system in the sense that it is self-sustaining and self-regulating. This is because it can adapt to change and cope with stress or bounce back from interference.[163] Third, it is self-going or self-running. That is, nature supplies itself with all the energy it needs to sustain itself and to adapt.[164] Finally, parts of nature interact with each other. For example, ecosystems interact with adjacent ecosystems and the organisms in these ecosystems.[165]

---

[161] Samuelsson (2008): 147. Samuelsson actually argues that nature's complexity is what gives us a reason to appreciate it and treat it well. But, for our purposes, we can reinterpret his view as being about value.

[162] Samuelsson (2008): 147. The following philosophers think that integrity matters to the value of nature: Leopold (1949: 224-5), Routley and Routley (1979), and Elliot (1992: 151).

[163] Samuelsson (2008): 147-148.

[164] Samuelsson (2008): 148.

[165] Samuelsson (2008): 149.

According to Samuelsson, these features are interrelated. First, nature must be very intricate if it is going to have a high level of integrity and be self-going. Second, nature must have a level of integrity in order for it to remain self-going, i.e., it must be sustaining.[166] Third, some part of nature must interact with its environment if it is to be self-going, because it needs to get its energy from somewhere. Fourth, integrity is needed precisely because parts of nature interact with their environment, which interferes with these parts of nature, resulting in nature needing to cope or adapt.[167]

Thus, on a view like Samuelsson's, the amount of value nature has is a function of how complex it is, which, in turn, is a function of the degree to which it is intricate, has integrity, is self-going, and interacts with its environment. According to Samuelsson, nature has all these features to a high degree, it is plausible that nature is very valuable according to this view and any view that takes any one of the features of "complexity" to determine the degree of nature's value.[168]

Related views also see diversity as adding to the value of nature.[169] For example, Westra (1994) writes “biodiversity contributes to integrity”.[170] While diversity can refer to a variety of things, these authors generally have in mind diversity in the species in a given environment and diversity in the ecosystems in a given environment. Nature in general and many environments in nature are quite diverse. This is more reason to think that Non-Anthropocentrists who think diversity matters to the value of nature will also agree that nature is very valuable.

---

[166] Proponents of such a view include, Woodley et al. (1993), Westra (1994): 24-5, Lemons & Westra (1995) and Lemons et al. (1998).

[167] Samuelsson (2008): 150.

[168] Samuelsson (2008): Chapter 5.

[169] Westra (1994): 24-5 and Russow (1981): 109.

[170] Westra (1994): 24-5.

Some authors think that the fact that nature has value in part because it has not been shaped by human intention.[171] In particular, some think it is the fact that nature's extremely intricate organization came about (and is sustained) without the help of any human intention.[172] Elliot (1997) writes:

> The fact that nature's organizational complexity arises in the absence of intention and design itself contributes crucially to nature's aesthetic value. Moreover, this fact transforms the aesthetic value in question into the kind of aesthetic value that gives rise to moral value.[173]

Samuelsson (2008) agrees that the non-intentional source of nature's intricate organization matters for its value. However, he adds that this organization is not only extremely intricate but also that we cannot imitate it. That is, we cannot by ourselves create objects with as intricate an organization.[174] Nature developed and sustains this intricate organization *completely* without the help of human intention. So, it has this property to a high degree as well. That is, it might have been the case that humans interfered with nature's organization and made it even more intricate so that some of its intricacy was non-intentional and some of it was intentional. But, that is not the case. Nature all by itself developed and continues to sustain its complicated structure. So, Non-Anthropocentrists who think that the non-intentional source of nature's intricacy adds to its value will also likely agree that nature is very valuable.

It looks like Non-Anthropocentrists with diverse views on what makes nature valuable will likely agree that nature is very valuable given that it has the value-relevant features to a high degree. Therefore, it looks like there is a kind of convergence between Broad Anthropocentrism and Non-Anthropocentrism on roughly how valuable nature is. That is,

---

[171] Brennan (1984), Elliot (1992, 1997), and Katz (1992).

[172] Elliot (1997): 61 and Samuelsson (2008): 163.

[173] Elliot (1997): 61.

[174] Samuelsson (2008): 163.

they seem to exhibit total or mass convergence on the answer to the question of "How valuable is nature?" Of course, as I noted earlier, this is convergence on a somewhat vague answer. To agree that nature is "very" valuable is not to agree on a precise value that nature has. However, this convergence at least shows that Broad Anthropocentrists and Non-Anthropocentrists are basically on the same side when it comes to the amount of nature's value. And, if this is true, then it's likely that they will also exhibit a non-trivial amount of practical convergence.

7.3 My View vs Norton's View

Recall that one problem with Norton's view is that Weak Anthropocentrism and Non-Anthropocentrism don't exhibit any degree of non-fundamental or fundamental explanatory convergence. One advantage of my view over Norton's is that, as I've just argued, is that my view shows that there is non-fundamental explanatory convergence between Broad Anthropocentrism and Non-Anthropocentrism. This is because they both hold that nature is *finally* valuable. Moreover, they also both hold that nature is *very* valuable. This means that there will be some non-trivial amount of practical convergence between these two views. So, even though Broad Anthropocentrism and Non-Anthropocentrism don't exhibit any fundamental explanatory convergence, because they don't agree on what gives nature final value, they at least argue that natural a particular kind of value and having this value explains why we have reason to treat nature well.

8. Conclusion

In this chapter, I argued against Norton's "Convergence Hypothesis," i.e., Weak Anthropocentrism and Non-Anthropocentrism exhibit total or mass practical convergence. I explained why Anthropocentrism and Non-Anthropocentrism are thought to be inconsistent. I then explained why Norton thought these two views could exhibit total or mass practical convergence. I objected to Norton's convergence hypothesis. I argued that Weak

Anthropocentrism cannot be made to converge with Non-Anthropocentrism because the former allows humans to have fully considered desires that are inconsistent with treating nature in the way that Non-Anthropocentrism requires. I also argued that, even if Norton was right about mass or total practical convergence, many Anthropocentrists and Non-Anthropocentrists will be unsatisfied with his argument, because what they really care about is what properties make actions right or wrong. I then suggested a weaker convergence argument according to which Broad Anthropocentrism exhibits a non-trivial amount of practical and non-fundamental explanatory convergence with Non-Anthropocentrism. I also argued that there is some reason to think that Broad Anthropocentrists and Non-Anthropocentrists converge on the question of roughly how valuable nature is.

Chapter 4: Converging Utilitarianism and the Moral Rights View on Animal Experimentation

1. Introduction

A prominent debate in animal ethics concerns the moral permissibility of animal experimentation.[175] In these, and related debates, animal ethicists often consider what the Utilitarianism and MRV entail about a particular issue.[176] As should be clear from previous chapters, Utilitarianism will hold animal experimentation is morally permissible just so long as it has the right consequences overall.[177] According to MRV, animal experimentation is either never or only rarely morally permissible.[178] This is because animals have moral rights which limit the ways in which they can be morally permissibly treated in the same way that humans have moral rights that limits the ways in which they can be morally permissibly treated. According to MRV, the source or foundation of the moral rights that these animals have is their inherent value.

Although they ground morality in different properties, both theories can exhibit some practical convergence. For example, both views tend to forbid using animal for entertainment (e.g., in circuses), animal hunting, and animal factories, etc. From the point of view of MRV, using animals for entertainment is a violation of the animal's moral rights and therefore it is morally wrong. From the point of view of Utilitarianism, the pain and suffering of animals

---

[175] For defences of animal experimentation, see Cohen (1986), McCloskey (1987), Brody (2003), and Frey (2005).

[176] For a review of the ethical issues involved in animal experimentation, see DeGrazia (1991), Frey (2003), and Norcross (2007).

[177] LaFollette and Shanks (1995) argue that Utilitarianism does not typically support animal experimentation.

[178] For defences of some version of the Moral Rights View, see Feinberg (1974), Francione (2003), and Regan (2004).

caused by animal abuse diminishes the pleasure and happiness of all sensitive beings and therefore it is morally wrong.

However, on the issue of animal experimentations, these two theories seem to be unable to reach an agreement.[179] Whether non-human animals should be used in scientific experiments is a point of serious contention between Utilitarianism and MRV. In short, unlike other cases, the issue of animal experimentation seems to be the starkest conflict between these two main theories of animal ethics.

Gary Varner thinks that it is important to try to show that there is convergence between Utilitarianism and MRV on the issue of animal experimentation. He thinks it's important because he thinks doing so could refocus the debate in a way that would at least lead to some agreement on animal experimentation policy and this would benefit proponents of both sides.[180]

He argues that Utilitarianism and MRV can converge on which cases of animal experimentation are morally permissible, required, or forbidden. In particular, he argues that a particular form of Utilitarianism (i.e., Preference Utilitarianism) and a particular form of MRV (i.e., Tom Regan's worse-off principle) exhibit non-trivial converge on the morality of of animal experimentation.[181] Preference Utilitarianism is the view that an action is morally required if and only if (and because) it maximises overall desire satisfaction.[182] The worse-off principle only concerns cases in which there is a conflict between two or more people's right not to be harmed. This principle says that, in such cases, when the harms of the rightsholders

---

[179] For empirical research on the use of animals in experimentation, see ABPI (2007), AFMA (2004), Animal Procedures Committee (2003), and Nuffield Council on Bioethics (2005).

[180] Varner (1994): 24.

[181] Varner (1994): 27.

[182] While preferences and desires are clearly distinct (e.g., preferences are always comparative but desires need not be), I will refer to this view as Preference Utilitarianism because that is what the authors I discuss call it. I thank Robert Cowan for suggesting this clarification.

are non-comparable, one ought to avoid harming those that would be worse-off.[183] A's harm is not comparable to B's if A's harm frustrates more desires or more important desires than B's harm. Or, if the harm faced by a few individuals makes them worse-off than a larger group of individuals, then one ought to avoid harming the few.[184] For example, if one must choose between torturing 2 people and giving 1,000 mild headaches, one is morally required to give the 1,000 people headaches because torture is a harm that is much worse than a mild headache.

The worse-off principle supposedly permits or requires animal experimentation in which the kind of harm done to non-human animals is less bad than the kind of harm that humans will suffer if the experiments are not done. In fact, Varner argues that the death of non-human animals, while a harm, is always less of a harm than the death of a human. So, the worse-off principle permits or requires lots of animal experimentation, even when the non-human animals die.

Varner thinks that Preference Utilitarianism and the worse-off principle will converge because both permit and require lots of animal experimentation. In fact, they can both permit or even require animal experimentation when the animals die. For the worse-off principle, all that is required for requiring some particular animal experimentation is that at least one human will be prevented from suffering a harm that is worse than the harm suffered by any single non-human animal in the experiment. That is, what is compared is not the *overall* harm of doing the experiment vs not doing the experiment. What is compared is the *kind* of harm suffered by any individual if the experiment is performed vs the kind of harm suffered by any individual if the experiment is not performed. Therefore, according to the worse-off principle,

---

[183] While Varner thinks that the worse-off principle is a version of MRV, I will argue in section 5.2 that it is not.

[184] Lafollette and Shanks (1996): 41.

the *aggregate amount* of harm is irrelevant, all that matters is that one single individual suffers a harm that is worse than the harm suffered by any other individual.

In this chapter, I argue that even appealing to the above two forms of Utilitarianism and MRV is not sufficient for showing that Utilitarianism and MRV exhibit some non-trivial degree of convergence on the issue of animal experimentation. First, the worse-off principle is not by itself plausible as a *complete* version of MRV, i.e., proponents of MRV will endorse further principles and these further principles will prevent convergence with Preference Utilitarianism. Second, even if Varner is correct that there is some convergence between Preference Utilitarianism and the worse-off principle, it is not the kind of convergence that would satisfy proponents of either view.

Why should we care about whether Utilitarianism and MRV convergence on the moral permissibility of at least some animal experimentation? There are a few reasons. First, it would be interesting to find out that there is more agreement than either side realised there was. Second, Utilitarianism and MRV are the two most prominent normative ethical theories endorsed by people in charge of animal policy and so if their two preferred views agreed on the morality of certain animal experiments, it will be easier to get animal policies enacted. Third, as Parfit (2011) argues, disagreement between different normative ethical theories calls into question whether there are moral truths. After all, if animal ethicists, who are equally intelligent and well-informed, disagree about what animal experimentations are permissible, required, or wrong, then we have some reason to think that there are no facts of the matter. Finally, such disagreement might suggest that, even if there are moral facts, we can't figure out what they are. So, if we can show that there is some important amount of convergence, we have less reason to worry about these kinds of scepticisms.

The plan of this chapter is this. In section 2, I will explain what Utilitarianism and MRV amount to in more detail. In section 3, I will explain the worse-off principle in more detail

and explain why Varner thinks that it exhibits some degree of convergence with Preference Utilitarianism. In section 4, I consider and reject an objection from Alan Clune that the worse-off principle only applies in exceptional situations and that animal experimentation is not an exceptional situation. In section 5, I argue that Varner's arguments fail to show that there is any meaningful convergence between Utilitarianism and MRV. One problem with Varner's argument is that he treats the worse-off principle as a complete version of MRV, but no plausible version of MRV would only have the worse-off principle and a full version of MRV would include principles to forbid much animal experimentation. Moreover, I argue that even if Varner is correct that there is some convergence between Utilitarianism and MRV, I argue that the kind of convergence he argues for will not satisfy proponents of either Utilitarianism or MRV.

Before explain the different normative ethical theories that Varner claims exhibit practical convergence, I need to briefly clarify what kind of non-human animals this debate concerns. In order to minimise controversy, both Utilitarians and proponents of MRV focus only on mammals and, in particular, mammals that are over a year old.[185] Whether younger mammals or other non-human animals count morally will, following the debate, be left to the side.

The reason to focus on mammals over a year old is that all the main authors involved in this debate agree that desire satisfaction and frustration as very morally important and it seems clear to these authors that mammals over one year old tend to have desires that can be frustrated.[186] In particular, these creatures have "forward-looking" desires, i.e., desires about the future. One of the most important of these desires is the desire to go on living.[187]

---

[185] Regan (2004): 78.

[186] Varner (1994): 25; Singer (1993); and Regan (2004).

[187] Singer (1993): 125 and Varner (1994): 25.

However, it is more controversial whether younger mammals and other non-human animals have such desires.

2. Utilitarianism and the Moral Rights View

Varner's thesis concerns two particular version of consequentialism (i.e., Preference Utilitarianism) and deontology (i.e., The Moral Rights View). Thus, in order to properly assess his argument, we need to have a grasp on what these particular theories hold.

In this section, I will explain both Preference Utilitarianism and MRV in greater detail and show how they differ.

2.1 Preference Utilitarianism

The most common version of Utilitarianism is known as hedonistic Utilitarianism because it concerns maximising *pleasure* and minimising pain.[188] However, Varner thinks that in order to get some degree of non-trivial convergence between Utilitarianism and MRV, he must appeal to a different version of Utilitarianism, i.e., Preference Utilitarianism.[189] According to Preference Utilitarianism, an action is morally required if and only if (and because) it maximises overall desire satisfaction.[190] So, Varner should be seen as endorsing, for the purpose of his convergence argument, the claim that maximising human happiness means maximising overall desire satisfaction. What does it mean to maximise *overall* desire satisfaction? While we needn't answer this question precisely for the purposes are arguing for convergence, here is one plausible story. All desires come in strengths, i.e., *how much* someone desires something. How much overall desire satisfaction is in the world is determined by adding together the strength of all the desires that are satisfied by some action. One way to maximise overall desire satisfaction is to satisfy many weak desires and another

---

[188] Sinnott-Armstrong (2019).

[189] Varner (1994): 25.

[190] For defences of preference Utilitarianism, see Hare (1981), Singer (1993), and Harsanyi (1996).

way is to satisfy a few very strong desires. Either way, what matters is the overall strength of desire satisfaction that results from actions.

Thus, in order to determine the moral status of an action, one must look at the total amount of desire satisfaction created by the action that one is considering whether to perform. That is, each individual being's desires count and the action which produces the greatest amount of desire satisfaction is the morally required action and any other action is morally wrong. For example, if I have to choose between saving 1 or 2 people from drowning, then, as long as everyone involved has the same number of desires satisfied by my action (and they all desire to not drown), Preference Utilitarianism requires that I save the 2 people. Thus, according to Preference Utilitarianism, the fundamental right-making property is desire satisfaction maximisation and the fundamental wrong-making property failure to maximise desire satisfaction.

Preference Utilitarianism is also committed to the claim that the satisfaction/frustration of different people's desires is comparable or commensurable, i.e., the satisfaction/frustration of different people's desires can be compared and weighed against each other. In other words, there is a standard for measuring and comparing the satisfaction of everyone's desires. As a result, the satisfaction of 100 of A's desires counts just as much as the satisfaction of 100 of B's desires (assuming that all of these desires are had with the same intensity).

Moreover, following Peter Singer, Varner assumes that the individual desire satisfaction/frustration of non-human animals count just as much as the individual desire satisfaction/frustration of humans. Varner doesn't make it clear, but he must be assuming the desires of the humans and the desires of the non-human animals that count equally are had with the same intensity.[191] So, the idea is that the satisfaction/frustration of desires with equal intensities count equally from the moral point of view.

---

[191] Singer (1993): 67.

Finally, Preference Utilitarianism is committed to a certain claim about value. In particular, it's committed to the claim that only the satisfaction of individual desires has "intrinsic" value.[192] As we saw in Chapter 2, for X to be intrinsically valuable is for X to be finally valuable (i.e., valuable in its own right or for own sake) in virtue of some set of X's internal or non-relational properties.

In order to move from facts about the value of something to claim about what is morally right or wrong, Preference Utilitarianism implicitly endorses a claim like the following: the only morally permissible way to treat something with intrinsic value is to maximise the amount of it that exists. This means that when someone has the ability and opportunity to bring about different amounts of something that has intrinsic value (e.g., happiness, desire satisfaction, pleasure, etc.), they are morally required to choose to bring about the most or the highest quantity of it that they are able to bring about given their circumstances. Preference Utilitarians think that the satisfaction of desires is intrinsically valuable and so they are required to bring about the greatest amount of desire satisfaction that they are able to bring about given their circumstances. As I noted earlier, the overall amount of desire satisfaction will be determined both by the number of desires that they satisfied and by the intensity or strength of those desires.

While Preference Utilitarianism does take the satisfaction/frustration of non-human animal desires into account, what matters for whether any particular animal experimentation is morally permissible or required is whether it will result in greater *overall* desire satisfaction. Thus, any experiment can be morally required so long as it produces the greatest overall desire satisfaction given an agent's options. Thus, according to Preference Utilitarianism, many animal experiments are morally justified because they result in the

---

[192] Rabinowicz and Österberg (1996).

maximisation of *human* desire satisfaction even if they involve frustrating non-human animal desires.[193]

Utilitarians do not ignore the feelings of non-human animals or the fact that animals can suffer, but they tend to consider the greatest interests for the largest number of individuals after balancing and aggregating those benefits and harms from those experiments. Therefore, for utilitarians, since the humans who will benefit from animal experimentation constitute a majority, the greatest benefit will be produced by sacrificing a small number of animals.

One might wonder why Varner chooses to argue that Preference Utilitarianism, as opposed to Hedonistic Utilitarianism, exhibits some convergence with MRV. There are two possible reasons. First, it is the version of Utilitarianism endorsed by the most prominent animal ethicist i.e., Peter Singer.[194] Second, the worse-off principle gives an important role to desire satisfaction and frustration (more on this below) and therefore Varner might think that Preference Utilitarianism is more likely to converge with the worse-off principle than other forms of Utilitarianism.

2.2 The Moral Rights View

According to the Moral Rights View (MRV), certain beings have rights that limit the way that others are morally permitted to treat them. According to Regan (2004), the source of a being's moral rights is that being's inherent value. And, whether someone has inherent value is independent of their pleasures, pains, benefits or ills, or any feelings. As we saw in Chapter 2, what matters for whether a creature has inherent value is whether that creature is the "subject of a life." Regan explains as follows:

> To be the subject-of-a-life… involves more than merely being alive and more than merely being conscious… [I]ndividuals are subjects-of-a-life if they have beliefs and desires; perception, memory, and a sense of the future, including their own future; an emotional life together with feelings of pleasure and pain; preference- and welfare-

[193] Francione (2003).

[194] Singer (1993). Varner (1994) explicitly notes this (25).

> interests; the ability to initiate action in pursuit of their desires and goals; a psychophysical identity over time; and an individual welfare in the sense that their experiential life fares well or ill for them.[195]

Regan, therefore, disagrees with Singer who argues that mere sentience (i.e., the ability to have experiences of pleasure and pain) is sufficient for inherent value.[196] What is essential for inherent value, according to Regan, is a set of psychological states and abilities (e.g., beliefs, desires, a sense of the future, experiences of pleasure and pain, the ability to initiate action to achieve goals, *psycho*physical identity over time) and interests (e.g., interests in having preferences and wellbeing protected).

Moreover, as we saw in Chapter 2, Regan argues that inherent value is not something that people can have to varying degrees. That is, everyone that has inherent value has it equally. One's level of inherent value doesn't vary with varying levels of intelligence, kindness, rationality, and so on.

One might wonder why all creatures with inherent worth have it equally. After all, if the basis of inherent worth are certain psychological states and abilities (beliefs, desires, an emotional life, the ability to initiate action in pursuit of a goal, and son) and these states and abilities are a matter of degree, why isn't inherent value a matter of degree. Regan's argument is that if different creatures had inherent value to differing degrees of inherent value then those creatures with less inherent value could be justly required to serve the interests of the creatures with more inherent value and the creatures with less inherent value would have no

---

[195] Regan (2004): 243.

[196] Singer (1990 and 1993).

ground of complaint. But, Regan argues, this view of what is required by justice is implausible and so it must be that all creatures with inherent worth have it equally.[197]

Recall that inherent value is importantly different from both intrinsic value and instrumental value. Recall that something has inherent value when it is finally valuable in virtue of being the subjective of a life, i.e., having the above psychological and interest properties discussed above. Whether being the subjective of a life is an intrinsic property of a being need not concern of here. The point of difference for my purposes is that only a certain set of properties makes something inherently valuable. So, for example, Picasso's paintings might be intrinsically (aesthetically) valuable, but they lack inherent value because they don't have the required psychological or interest properties.[198]

Inherent value is also importantly different from instrumental value. This is because things that are inherently valuable are finally valuable, i.e., valuable for their own sake, but instrumentally valuable things are non-finally valuable, i.e., they are valuable only as a means to something else that is valuable. Steel is instrumentally valuable because its value lies in its ability to be used to make tools. However, it has no inherent value.

Another difference between inherent value on the one hand and intrinsic and instrumental value on the other is that while the latter two come in degrees, the former does not. That is, as noted above, inherent value is not a matter of degree, something either has it or it doesn't. Both humans and non-human animals have inherent value and they have it equally. But both intrinsic and instrumental value are always possessed to some degree. We know that some steels are of better quality than others and can produce better tools. Such

---

[197] Regan (2004): 237. I find Regan's argument suspicious, because he appears to be offering a practical reason to believe that all creatures that have inherent value have it equally. His reason is that it would be morally bad if non-human animals had less inherent value than humans because then it would be morally justifiable to use animals in a way that was harmful to them in order to benefit humans. But it is controversial whether there are practical reasons to believe. I discuss this matter more below.

[198] They plausibly also have instrumental value because of the pleasure they cause or what they teach people about fine art.

steel tools are usually more instrumentally valuable. Similarly, Picasso's paintings have higher intrinsic aesthetic value than the paintings of a lesser, but still talented, artist.

How do we get from the fact that something has inherent or intrinsic value to the claim that they have rights? As we saw in the last chapter, Reagan argues for what he calls, "The Respect Principle," i.e., "We are to treat those individuals who have inherent value in ways that respect their inherent value." Thus, individuals who have inherent value have a fundamental moral right to respectful treatment. According to Regan, it follows from the Respect Principle that individuals with inherent value have a moral right to not be harmed intentionally—even "on the grounds that all those affected by the outcome will thereby secure "the best" aggregate balance of intrinsic values (e.g., pleasures) over intrinsic disvalues (e.g., pains)."[199]

Thus, according to MRV, the fundamental right-making property is respecting an inherently valuable being's moral rights and the fundamental wrong-making property is violating an inherently valuable being's moral rights. One primary way of violating an inherently valuable being's moral rights is to intentionally or knowingly cause them harm.

According to MRV, what matters most fundamentally for determining what one is morally required to do are the rights of *individual* inherently valuable beings. That is, the thing that people ought to do is try to protect the rights of individuals. This is in stark contrast to Utilitarianism, which holds that what matters most fundamentally to what one is morally required to do is the *overall* or the aggregate pleasure or happiness of all beings related to an action. It is not any individual's pleasure or happiness that matters, but the *sum total* of pleasure or happiness. MRV, therefore, holds that violations of an individual's moral rights are always wrong—regardless of any good consequences that might follow from such a violation.

---

[199] Regan (2004): 286.

What about cases of conflicts between rights? The rights of individuals cannot be compared or weighed. That is, one is not morally required to minimise the number of rights violations. It is not clear, according to MRV by itself, what one is morally required to do when one must choose between violating either of two people's moral rights. For example, it's not clear based on MRV alone what one ought to do if one must choose between killing one innocent person and killing another, equally innocent, person. Moreover, because all beings that have inherent value have it equally and one's moral rights are based solely on one's inherent value, all beings with moral rights have equal weight or stringent rights. Thus, a non-human animal's moral right to not be killed is not weaker than a human's moral right to not be killed.

When it comes to the moral permissibility of animal experimentation, all that matters for MRV is whether a particular experiment violates an animal's moral rights, e.g., the right to not be intentionally harmed. Thus, any experiment that involves knowingly harming a non-human animal involves violating a moral right and is therefore morally wrong. Because many animal experiments clearly involve knowingly inflicting harm on animals, a large number of animal experimentations are morally wrong according to MRV. Moreover, these experiments are morally wrong, according to MRV, even if we know that they will prevent human or other non-human animals from experiencing suffering and even if it minimises the number of future rights violations.

3. Converging Utilitarianism and the Moral Rights View

In this section, I will explain Varner's argument that there can be a non-trivial amount of practical convergence between Utilitarianism and MRV concerning animal experimentation. To preview: Varner argues that if we treat Tom Regan's worse-off principle as a version of MRV, then there is some non-trivial convergence between Utilitarianism and MRV. The worse-off principle permits the overriding of rights in certain cases that involve what Regan

calls “non-comparable” harm. In what follows, I explain the worse-off principle in greater detail as well as the difference between comparable and non-comparable harm.

3.1 The Worse-Off Principle

Given what we’ve seen, Preference Utilitarianism and MRV are fundamentally opposed on the issue of moral experimentation. This is because MRV forbids any animal experimentation in which animals are knowingly harmed—regardless of the good consequences that may come from it. However, Preference Utilitarianism is happy to permit any animal experimentation as long as it results in the maximisation of overall desire satisfaction. Given that these views conflict so much, how should we proceed in order to figure out what we are morally permitted to do during animal experimentation? One strategy is to argue that one of these normative ethical principles is correct and therefore the other is false. This is not the strategy I’m interested in. A second strategy is to explore whether there is a way of getting Preference Utilitarianism and MRV to exhibit total practical converge. This is the strategy I’m interested in.

According to Varner, we can avoid the debate between Utilitarianism and MRV concerning animal experimentation if we appeal to what Regan calls the worse-off principle as well as to Regan’s preferred view of harm.[200] I will explain each view in turn.

Regan formulates the worse-off principle as follows:

> *The worse-off principle:* The rights of the many may be overridden in favour of the rights of the few if the harm faced by the few would make them worse off than any of the many.[201]

This principle has two elements that need to be clarified. First, recall that MRV didn’t allow for a weighing of rights, i.e., treating some rights as being weightier than others. In fact, it was not clear, on that view, how one should act when more than one being's rights are in

---

[200] Varner (1994): 27.

[201] Regan (2004): 308.

conflict. The worse-off principle differs from MRV by allowing for there to be cases where some being's rights outweigh the rights of others even though all the beings involved have equal inherent value.

Second, the worse-off principle only applies when at least one of the rights holders would be harmed in a way that makes them worse off than other rights holders would be harmed. In other words, the worse-off principle only applies when the conflict between rights involves what Regan calls "non-comparable" harms. To see what he has in mind, consider what he says about comparable harms:

> …Two harms are comparable when they detract equally from an individual's welfare, or from the welfare of two or more individuals. For example, separate episodes of suffering of a certain kind and intensity are comparable harms if they cause an equal diminution in the welfare of the same individual at different times, or in two different individuals at the same or different times…. [202]
>
> …Other things being equal, that is, it is reasonable to assume that like harms have like effects--that is, detract equally from individual welfare and so are to be counted as comparable.[203]

This means that when the welfare of individuals or the welfares of two or more individuals, is equally infringed upon then the harm is comparable. For example, death can be treated as a comparable harm if the loss of opportunities it marks are equal in any two cases.

Two harms to two different people are non-comparable when they detract *differently* from each individuals' welfare. Varner further explains that different kinds of harms can have different losses for the same individual and the same kind of harm can result in different losses and injuries for different individuals. For example, losing an arm is more harmful than hurting one's toe because it can frustrate more of an individual's desires. Likewise, the same harm can frustrate different desires in different individuals. For example, losing eyesight will

---

[202] Regan (2004): 304.

[203] Regan (2004): 304.

result in different losses for a visual artist than it will for a musician. While the former needs vision to create their art, the latter does not (under normal circumstances).

One might wonder how a proponent of MRV like Regan could ever support a view like the worse-off principle given that MRV seems to forbid the overriding of rights and the worse-off principle allows them. While Regan supports a version of MRV, he thinks that moral rights are *prima facie*, which he defines as follows:[204]

> To say this right is a prima facie right is to say that (1) consideration of this right is always a morally relevant consideration, and (2) anyone who would harm another, or allow others to do so, must be able to justify doing so by (a) appealing to other valid moral principles and by (b) showing that these principles morally outweigh the right not to be harmed in a given case.[205]

Therefore, Regan rejects an absolute moral right not to be harmed, because he thinks this right can be overridden. However, he thinks that a person's right not to be harmed always have weight when one is determining what one is morally required to do. Thus, Regan seems to view moral rights as being *pro tanto* in the same sense that Ross thinks we have duties that are *pro tanto*. Moreover, harming someone and therefore violating this right always required that the harmer be able to justify what they are doing. Thus, going from Regan's version of MRV to the worse-off principle is not too puzzling, because he thinks rights can be overridden.[206]

To see how the worse-off principle would work, consider the following case. Imagine that there are two wells with miners trapped in them. The first well has only one miner in it while the second has fifty miners in it. The current rescue tools are limited and can only save miners in one well. The miner who is alone in that well has relapsed due to a heart attack and

---

[204] Regan (2004): 287. It seems to me that Regan should have used the term "pro tanto" instead of "prima facie" because a person's moral right not to be harmed has nothing to do with what is *prima facie*, i.e., it is not a matter of how things *appear on the surface*. Rather, as Regan uses the term, it means that the right is always in force or matters for determining what one ought morally to do, but that it can be overridden or outweighed sometimes.

[205] Regan (2004): 287.

[206] I thank Robert Cowan for pressing me to say more on this matter.

could die if not rescued immediately, while the other fifty miners in the other well were very well at that moment and they had no problem waiting for the next rescue.

In this case, according to the worse-off principle, it is entirely correct to consider rescuing the single miner first. This is because even if every individual has equal moral right not to be harmed, this does not mean every harm itself is equally harmful. In the above case, other things being equal, the miner with a heart attack's possible death is a greater harm than other fifty miners' thirst or nervousness, and the greater harm faced by the miner with a heart attack would make him worse-off than any of the individual fifty miners would be. Therefore, we should consider the greater harm of one miner instead of the lesser harms of the fifty miners. This is because the potential harm to the first miner is not comparable with the potential harms to the other miners.

Thus, according to the worse-off principle, an action is morally required if and only if (and because) it prevents one individual from suffering a harm that is worse than the harm that any other single individual involved would suffer.[207] An action is morally wrong if and only if (and because) it allows one individual to suffer a harm that is worse than a harm suffered by any other single individual involved.

How is the worse-off principle related to MRV? The core of MRV is The Respect Principle mentioned above, i.e., "We are to treat those individuals who have inherent value in ways that respect their inherent value." Regan thinks that it follows from this that we ought to see every individual as having an *equal* right to not be harmed. This is because they have

---

[207] A separate issue that I cannot hope to address here is whether the worse-off principle is plausible. I have my doubts. One might find it intuitively implausible that one can be required to maim 50 people instead of killing one person or letting them die. But this is what the principle holds. After all, we are only supposed to compare the harms suffered by individuals and not groups. So, in order to determine what to do, we must compare a single case of maiming with a single case of death and because death is usually a worse harm than maiming, we morally ought to kill or let die 50 people instead of killing one. While I'm sympathetic to such objections, I think a related objection is stronger. Imagine having to choose between inflicting 100 units of pain on one person and inflicting 99.999 . . . units of pain on one million people. Even if we compare harms on a one-to-one basis, i.e., compare the worse harm from the first person with the worst harm of any individual from the group of one million, it seems strange to require that we cause the 99.999 . . . units of pain to the one million.

equal inherent worth. But, one might wonder, how can the view that every individual with inherent worth has equal inherent worth be compatible with overriding one individual's rights to protect another's?

Regan argues that these two claims are compatible in the following way:

> To say that two individuals, M and N, have an equal right not to be harmed, based on the equal respect each is owed, does not imply that each and every harm either may suffer is equally harmful. Other things being equal, M's death is a greater harm than N's migraine. If we are to show equal respect for the value and rights of individuals, therefore, we cannot count a lesser harm to N as equal to or greater than a greater harm to M. To show equal respect for the equal rights of the two, one must count their equal harms equally, not their unequal harms equally, a requirement that entails, other things being equal in prevention cases, that M's right override N's when the harm done to M would be greater if one choice were made than the harm done to N would be if another option were chosen.[208]

So, Regan thinks that the fact that two individuals have an equal right to not be harmed doesn't entail that they are owed equal treatment *if* the harms involved are *unequal*. In fact, it would be unfair to treat N's lesser harm as being *as serious as* M's harm. So, Regan thinks, it is compatible with the equal inherent worth and equal right of each individual that one individual's right to not be harmed is overridden by another individual's right to not be harmed.

What makes one harm worse than another? Varner interprets Regan as holding a desire satisfaction view of welfare and thus as holding that harms consist in the reduction in one's desire satisfaction or a reduction in one's capacity to form and satisfy desires.[209] Thus, harm A is worse than harm B to a person just in case harm A frustrates more of a person's desires than B does. And death is the worst harm because it destroys the capacity to form or satisfy any desires.[210]

---

[208] Regan (2004): 309.

[209] Varner (1994): 26.

[210] Varner (1994): 26.

However, Varner misinterprets Regan's view. Regan explicitly argues that, "Individuals are harmed when their welfare is seriously diminished", but Regan thinks that more than desire satisfaction is part of welfare.[211] He writes:

> To live well, relative to one's capacities, involves more than having benefits. One must also take satisfaction in having, or in using, one's benefits, not only sporadically, but on balance, over time. To live well is to have a life that is characterized by the harmonious satisfaction of one's desires, purposes, and the like, taking account of one's biological, social, and psychological interests. More generally, animals (and humans) live well relative to the degree to which (1) they pursue and obtain what they prefer, (2) they take satisfaction in pursuing and obtaining what they prefer, and (3) what they prefer and obtain is in their interests.[212]

Regan does think that an animal's pursuing and obtaining what they *prefer* is part of their welfare (or living well). Of course, however, preferences and desires are distinct (e.g., preferences are always comparative). But, even granting that they are the same thing, Regan still thinks more is required for well-being or welfare.

First, Regan thinks that animals must *take satisfaction* in pursuing and obtaining what they prefer or desire. But taking satisfaction in acquiring what one desires is not the same thing as having that desire satisfied. After all, taking satisfaction in the fact that one's desire is satisfied require *knowing* that the desire is satisfied. But, one's desire can be satisfied without one knowing it. For example, one might desire to be respected by one's colleagues, but because one's colleagues are guarded, one never knows if they respect one

Second, Regan thinks that what animals prefer or desire and obtain must *actually* be in their interest. As Regan notes, one can prefer or desire something that is not in one's interest, e.g., doing a drug that harms one's health.[213] And, one can lack a preference or desire for something that *is* in one's interest, e.g., to exercise regularly.[214]

---

[211] Regan (2004): 94.

[212] Regan (2004): 117.

[213] Regan (2004): 87.

[214] Regan (2004): 87.

Thus, it seems clear that Varner is incorrect that Regan holds a simple desire-satisfaction account of welfare. However, as we will see shortly, despite what Varner argues, precisely what account of welfare Regan holds doesn't matter for Varner's argument.

3.2 The Worse-Off Principle and Animal Experimentation

Varner argues that from the combination of the worse-off principle and Regan's conception of harm, it follows that at least some animal experimentation is permissible and even required.[215] However, Regan's precise conception of harm is not relevant for Varner to make his argument. Rather, all Varner needs to show is that Regan's view of harm (whatever that view is) and the worse-off principle entail that the death of an average individual human being constitutes a greater harm than the death of an average non-human animal used in an experiment. This is because, if experiments result in the saving of at least one human life, then, even if the experiments involve the death of many non-human animals, the experiment will be morally required. This is because the worse-off principle says that in cases in which we must harm certain inherently valuable individuals and all these individuals have an equal right to not be harmed, then we should prevent the worst harm from occurring. This is because the worse-off principle is based on the idea that no inherently valuable individual ought to suffer a greater harm than any other inherently valuable individual.

Moreover, Varner is correct that Regan thinks that, under normal circumstances, death is a greater harm for an individual human than it is for an individual non-human animal. He argues for this by first proposing the following case:

> *Lifeboat*: There are five survivors on a lifeboat; four normal adults and one ordinary dog. However, there are only four positions on board, one of the passengers must be thrown overboard.[216]

He notes that both the humans and the dog have equal inherent value. However, he notes:

---

[215] Varner (1994): 27.

[216] Regan (2004): 324.

> [T]he harm that death is, is a function of the opportunities for satisfaction it forecloses, and no reasonable person would deny that the death of any of the four humans would be a greater prima facie loss, and thus a greater prima facie harm, than would be true in the case of the dog. Death for the dog, in short, though a harm, is not comparable to the harm that death would be for any of the humans. To throw any one of the humans overboard, to face certain death, would be to make that individual worse-off (i.e., would cause that individual a greater harm) than the harm that would be done to the dog if the animal was thrown overboard.[217]

Why does the death of an average individual human constitute a greater harm to that individual than the death of an average non-human animal constitutes for that animal? This is because Regan holds that "the magnitude of the harm that death is, is a function of the number and variety of opportunities for satisfaction it forecloses" and the average human life has more and a greater variety of opportunities for satisfying their preference, taking satisfaction in those preferences being satisfying and actually obtaining what it is in their interest.[218]

Moreover, Regan makes clear that "the numbers don't matter" and so any number of non-human animal deaths could be justified if their deaths were necessary for saving a single human life. Regan writes:

> Let the number of dogs be as large as one likes; suppose they number a million; and suppose the lifeboat will support only four survivors. Then the rights view still implies that, special considerations apart, the million dogs should be thrown overboard and the four humans saved. To attempt to reach a contrary judgment will inevitably involve one in aggregative [i.e., utilitarian] considerations.[219]

Thus, it looks like the worse-off principle combined with Regan's view death is a greater harm to humans than to non-human animals entails that some animal experimentation is morally required—at least when we know that it will lead to the saving of human lives.

To be clear, however, the fact that death is a greater harm to humans than to non-human animals doesn't entail that *any* harmful experiment is required or permissible. This is

---

[217] Regan (2004): 324.

[218] Regan (2004): 314.

[219] Regan (2004): 325.

because, as Regan notes, there are harms *greater than death*. For example, he gives the case of "A life of protracted, intense, untreatable suffering, one that promises no relief" and we can add cases of long protracted imprisonment and excruciatingly painful torture.[220] Thus, there are cases of non-human animal harm that are worse than human death.

Finally, Varner could also argue that it follows from the worse-off principle and Regan's view of harm that many other kinds of animal experimentation are morally required. All that needs to be true is that the experiment will prevent the death of a single human or prevent a single human from suffering a harm that is worse than the worst harm any individual non-human animal will suffer during the experiment. So, experiments that involve causing animals pain or that involve imprisoning them can be morally required as long as they can help prevent a human (or non-human animal) from suffering an even worse harm.

3.3 The Worse-Off Principle and Convergence on Animal Experimentation

Varner argues as follows: because the worse-off principle and the view that human death is worse than non-human animal death together entail that some cases of animal experimentation are morally required, it follows that Utilitarianism and MRV exhibit a non-trivial among of practical convergence. However, there is a problem with this reasoning. Just because the worse-off principle and the desire frustration view of harm entails that some cases of animal experimentation are morally required, it doesn't follow that these cases are the same as the cases that Preference Utilitarianism entails are morally required.

We can easily see that there won't be total practical convergence because Preference Utilitarianism allows for the aggregation of harm while the worse-off principle doesn't. Preference Utilitarianism entails that cases of animal experimentation that result in a greater sum total of frustrated animal desires over satisfied human desires, then that experiment is morally wrong. For example, return to the case of having to choose between killing some

---

[220] Regan (2004): 100.

large number of dogs (e.g., 1 million) vs killing a single human life. The worse-off principle (and the view that human death is usually worse than non-human animal death) entails that in this case one is morally required to kill the dogs and save the human. This is because the human would suffer a greater harm than any individual dog would. However, Preference Utilitarianism would entail that one is morally required to kill the human or let them die. This is because the sheer number of frustrated desires that would come about by killing 1 million dogs would outweigh the number of satisfied human desires—if one doesn't think that the death of a million dogs is enough to outweigh the harm of a single human death, then make it a billion or a trillion.

There will be two broad kinds of cases on which these two views would agree. First, there are cases in which a large number of human lives could be saved by doing experimentation. For Utilitarians, the large amount of desire satisfaction that would result from saving a large number of human lives would likely outweigh the amount of desire frustration resulting from harming many non-human animals. For proponents of the worse-off principle, the fact that a single human life could be saved would make morally required to harm any number of non-human animals. Second, there are cases in which the number of animals that needed to be harmed is not too much greater than the number of humans that would indirectly benefit from the harm. For Utilitarians, as long as enough humans benefited from the animals' harm, the experiment would be morally required. For proponents of the worse-off principle, as long as one human life would be saved, harming a few non-human animals would be required. These two views mainly seem to disagree when the number of animals that are harmed are killed is far greater than the number of humans that benefit.

Given this, Varner concludes:

> [T]he foregoing discussion illustrates how the implications of a true animal rights view can converge with those of researchers' animal welfare philosophy. Even someone who attributes moral rights in the philosophical sense to animals, and whose

> ethical theory thus differs dramatically from most animal researchers' [i.e., Utilitarianism], could think that some medical research is justified.[221]

The fact that Varner is claiming that there is there will be agree that "*some* medical research is justified [my emphasis]," means that he thinks the convergence is non-trivial, but not mass or total. After all, if he had just trivial convergence in mind, it wouldn't be worse arguing for, so he must have in mind more than trivial convergence. However, he is clear that there is only *some* convergence and so he can't have mass or total practical convergence in mind. However, as I will argue in the next section, even if he is correct that there is some non-trivial practical convergence between Preference Utilitarianism and the worse-off principle plus the desire frustration view of harm, it doesn't follow that "the implications of a true animal rights view can converge [on the question of what to do] with those of researchers' animal welfare philosophy [i.e., Utilitarianism]."[222]

Varner's strategy for convergence between Preference Utilitarianism and the MRV involves a few strategies that we saw in Chapter 1. First, as we will see in Section 6, a complete version of MRV consists of numerous principles, not just the worse-off principle. However, Varner argues only for convergence between Preference Utilitarianism and the worse-off principle. This is similar to the view discussed in Chapter 1 in which a person might argue for convergence with Rossian Pluralism by showing that their view exhibits non-trivial convergence with *one* of Ross's *prima facie* duties.

Second, because Varner treats the worse-off principle as a complete version of MRV, he seems to change the focus of MRV from protect people's rights against being harmed or used to benefit other people into a view that is primarily concerning with minimising the worse kinds of harm. In particular, the view argues that what matters most is preventing anyone from suffering a non-comparable harm. This makes MRV similar to Preference Utilitarianism

---

[221] Varner (1994): 27.

[222] Varner (1994): 27.

which is at least indirectly concerned with minimising overall harm (i.e., desire frustration). Minimising overall harm will often involve preventing the most serious kinds of harm (e.g., death) and so there will be some non-trivial amount of convergence between the two views.

Thus, Varner uses a strategy similar to another strategy we discussed in Chapter 1 in which one view is altered so that it has a morally-relevant property that is similar to the other view's morally-relevant property. In Chapter 1, we considered the case of a consequentialist view being altered so that it was concerned with minimising harm to innocent people. I argued that this would lead to at least some practical convergence with standard deontological views that are concerned with prevent harm to innocent people. In this case, the deontological view (i.e., MRV) is altered to become concerned with minimising a particular kind of harm and the consequentialist view (i.e., Preference Utilitarianism) is at least partially concerned with minimising over all harm (i.e., desire frustration).

4. Clune's Objection

The key to Varner's practical convergence argument is to view the Moral Rights Views as essentially consisting solely in Regan's worse-off principle. However, as Alan C. Clune argues, Regan thought that the worse-off principle only applied to *exceptional* cases like Lifeboat from above. However, Clune argues, testing on nonhuman animals are not exceptional cases like Lifeboat.[223]

Generally, in routine biomedical research, testing on nonhuman animals happens on an institutional scale, across numerous disciplines and in numerous ways more or less consistently, and Clune, therefore, considers animal experimentation as not being an exceptional case.[224] So the worse-off principle cannot be applied. Clune agrees with Regan's point of view, they both oppose the application of the worse-off principle to the normal

---

[223] Clune (1996) writes "What [MRV] implies should be done in exceptional cases including lifeboat cases, are exceptional cases — cannot fairly be generalised to unexceptional cases (325)."

[224]Clune (1996): 237.

institutional field and insist that the worse-off principle only applies to special circumstances.[225]

But what reason do we have for thinking that the worse-off principle only applies to exceptional circumstances? The principle itself does not specify anything about what circumstances it applies in. Rather, it is a principle that seems to apply to all cases in which there is a conflict between at least two rightsholders where at least one of their rights must be overridden. Recall that the principle roughly says, "The rights of the many may be overridden in favour of the rights of the few if the harm faced by the few would make them worse off than any of the many."[226] Moreover, conflicts between different people's rights not to be harmed is quite common, so the content of the worse-off principle doesn't seem to entail that it only applies to rare or exceptional circumstances. So, at first glance, it is unclear why we should think that the worse-off principle only applies in exceptional circumstances. As an analogy, imagine that I claimed that some version of Utilitarianism is true, e.g., one is morally required to maximise happiness, but then also claimed that this principle only applies in exceptional circumstances like Lifeboat. This claim seems completely unmotivated.

Clune argues that one reason why the worse-off principle doesn't apply to everyday circumstances is that it would be morally bad if it did, i.e., it would have bad moral consequences. In particular, it would lead to a kind of moral perfectionism and both Clune and Regan find to be unacceptable. Clune writes,

> And there is a good reason not to allow application of the worse-off principle at the institutional level. Allowing such an application would be equivalent to embracing perfectionism: this is the position that all burdens ought to be shifted to the less perfect in society so that the most perfect will encounter the least possible resistance to realizing their potential. And this would clearly be disrespectful to the inherent value of vast numbers of subjects-of-a-life, many of whom could in principle be humans.

---

[225] Animal testing in the scientific domain is institutional and common, also it is not possible for scientists to use only one animal for an experiment. In fact, they have to use a certain amount of animals to take research data from these certain animals. Thus, animal experimentations are not isolated cases like the lifeboat case.

[226] Kemmerer (2006): 67.

> Perfectionism is disrespectful precisely because on the rights view every individual subject-of-a-life deserves respect that is "logically independent of [that individual's] utility for others and logically independent of [that individual's] being the object of anyone else's interests.[227]

Clune's idea seems to be that if we applied the worse-off principle to everyday circumstances, it would lead to "less perfect" beings having to suffer more burdens so that "more perfect" beings could make progress.

There are numerous problems with Clune's argument. First, it is not clear why "less perfect" beings would end up bearing most of the burden. By "less perfect", Clune must mean non-human animals who don't have the same level of psychological ability and potential that humans have. So, he must be thinking that most of the time when an animal can be used to help a human, we should use that animal even if it violates the animal's right. But the worse-off principle doesn't say anything like this. In fact, in cases in which non-human animals are at risk of more severe harm than humans, it is human's moral rights that will be overridden according to the worse-off principle. This is because the worse-off principle entails that no creature with a moral right not to be harmed should suffer more than any other individual creature with a moral right not to be harmed.

Second, while Clune cites Regan as making this same argument, Clune is actually misrepresenting Regan's view. What Regan thinks will lead to an unacceptable form of perfectionism is that claiming that *inherent value is a matter of degree* and not that the worse-off principle is true. The worry is that humans possess the basis of inherent value (e.g., certain psychological states and abilities) to a greater degree than non-human animals and so one might think that they have a greater amount of inherent value than non-human animals. If humans have a greater amount of inherent value than non-human animals, then one might also think that humans' right not to be harmed is *always* stronger than non-human animals'

---

[227] Clune (1996): 237.

right not to be harmed. If this were true, then *any* conflict between a non-human animal's right not to be harmed and a human's right not to be harmed will result in the non-human animal's right being overridden. And that would seem to lead to an unacceptable form of perfectionism.

Third, even if Clune were correct that the worse-off principle applies to everyday circumstances would lead to a morally unacceptable form of moral perfectionism, it wouldn't follow that it is not *true* that the worse-off principle applies in everyday circumstances. This is because the fact that something, if true, would have very bad moral consequences doesn't mean that it's not true. The fact that it would be morally bad if something is true is not a reason to think it's not true. Clune seems to be offering the wrong kind of reason to think that something is true.

Finally, it is not clear to me that animal experimentation or testing is a common or everyday occurrence. That is, it seems to me that animal experimentation is an exceptional circumstance. And, this is true, even if it occurs all the time. This is because what counts as common or exceptional is relative to a certain context. If you're looking at the domain of biomedical research, then of course animal experimentation is not exceptional. However, the correct perspective for assessing whether a situation is exceptional or every day is from the perspective of most people in most jobs. For most people in most jobs, there is nothing like choosing where to cause harm to a non-human animal in order to benefit humans or even other non-human animals. So, from most people's perspectives, animal experimentation is quite exceptional. Some hold for saving human lives. Doctors save human lives every day, but that this is still an exceptional thing to do because the vast majority of people never save lives, never mind doing it every day. Therefore, it seems quite clear that, even if the worse-off principle only applies to exceptional circumstances, it still applies to animal experimentation. This is because animal experimentation is actually an exceptional circumstance.

5. Why Varner's Convergence Argument Fails

In this section, I argue that Varner's argument fails for two reasons. First, most proponents of versions of MRV (including Regan) think that there are additional principles (i.e., other than the worse-off principle) that apply to the case of animal experimentation. And, it is plausible that these principles will forbid animal experimentation and thus prevent convergence between some version of MRV and Preference Utilitarianism. Second, even if Varner is correct that there is some convergence between Preference Utilitarianism and the worse-off principle, it is not the kind of convergence that would satisfy proponents of either view.

5.1 Voluntarily Accepting Risk

The first problem with Varner's argument is that he treats the worse-off principle as the only principle that proponents of MRV would appeal to in discussing animal experimentation. However, as Regan himself argues, one reason that most animal experiments are morally wrong according to MRV is that they include putting the non-human animals at risk of harm without those animals voluntarily accepting the risk.[228] Of course, animal experimentation is meant to reduce the risk that humans undertake in using certain products, but Regan argues, one is not morally justified in *transferring* the risks from one being with rights to another being with rights unless the latter voluntarily accepts these risks.[229] But, it's quite obvious that animals do not and cannot voluntarily accept these risks.

Moreover, it will not do to argue that because they *could* not voluntarily accept these risks given their cognitive capacities, it doesn't matter if risks are imposed on them. This is because Varner and Regan agree that the non-human animals we're discussing have preferences and desires and being involved in animal experimentations conflicts with those

---

[228] Regan (2004): 377.

[229] Regan (2004) writes, "risks are not morally transferrable to those who do not voluntarily choose to take them (377)".

desires and preferences.[230] For example, Regan notes that when animals are stuck in laboratories and are being caused pain or discomfort, they are being forced to do things they desire not to do. Furthermore, Regan argues that exposing non-human animals to risks so that we don't have to be exposed to risk (or as much risk) is to treat them as "resources for others" and this is incompatible with treating them with respect.[231]

Varner acknowledges that Regan would reject his argument that MRV permits or requires some animal experimentation by appealing to the above anti-risk principle, but he argues that this principle is false. That is, Varner argues that it is sometimes morally permissible to transfer risks from one group to another without the latter's voluntary acceptance.[232] He gives two examples:

> [M]odifying price supports can redistribute the financial risks involved in farming, and changing draft board policies in a time of war can redistribute the risk of being killed in defence of one's country. Yet most people believe such transfers are justifiable even if involuntary.[233]

Thus, he argues that it is morally permissible for a government to modify price supports (e.g., subsidies, price control, or producing quotas) even though they redistribute certain financial risks involved in being a farmer and it is morally permissible for a government to have a military draft even though it redistributes the risk of being killed in war.[234]

I don't think that either of these cases shows that it is false that one cannot morally transfer risks from one group (or individual) to another without the latter's voluntary acceptance. Let's start with the case of price supports. It seems clear that people voluntarily accept the risks associated with price supports when they voluntarily choose to become

---

[230] Regan, (2004): 382.

[231] Regan (2004): 378.

[232] Varner (1994): 27.

[233] Varner (1994): 27.

[234] Varner (1994): 27.

farmers. They might not intentionally choose to put themselves at financial risk due to changes in price support, but they do voluntarily choose a profession that is known to have certain risks attached to it. So, this is not a counterexample to Regan's anti-risk principle.

Now let's consider the military draft case. First, it's not at all clear to me that this is a case of a *morally permissible* practice. It seems morally wrong to force people to risk their lives even in defence of a justified war. Take an analogous case in which you know that a stranger has begun attacking an innocent person and you have the ability to at least slow that person down or maybe prevent them from continuing to harm the innocent victim. However, it is also true that if you try to interfere with the attack, you bear a substantial risk of being killed (just as soldiers do in war). Would it be morally permissible for another bystander to push you into the attacker or otherwise force you to interfere with the attack? If everyone involved knows that forcing you to interfere with the attack would put your life at substantial risk, it seems clearly morally impermissible to force you to get involved. That being said, you might be praiseworthy if you tried to help the innocent victim on your own, but that doesn't mean it's morally permissible for someone else to force you to interfere. Moreover, the morality of military drafts is controversial among moral philosophers and so I don't think this case provides a counterexample to Regan's anti-risk principle. [235]

Second, even if it were morally permissible, the risk that is transferred to citizens is not actually the risk of being killed in war, but a disjunction of risks. That is, they are exposed to a risk of (1) being killed in war, (2) having to perform an alternative service (e.g., in healthcare or community services), or (3) being imprisoned in their own country if they refuse to go to war or perform an alternative service. Here the risks involved don't seem nearly as bad. But the animals chosen for experimentation cannot choose other options, e.g., non-harmful experimentation or working for a government agency (e.g., as a police dog).

---

[235] See the debate between: Pattison (2012) and Sagdahl (2018).

These animals only face the risk of death and harm. Nor for that matter could they choose these alternatives, because they wouldn't understand that they had the options. Or, if someone were making their choices for them and had their best interest in mind, they wouldn't choose they have the animals put at risk of death or harm.

Third, many think that at least some governments have a kind of practical authority over their citizens. That is, the government (or its officers) can force citizens to do certain things and citizens have a duty to not interfere with government actions. Of course, it's controversial *how* a government gets this authority, but that need not concern us here.[236] However, it is far less clear that individual human beings or researchers have any practical authority over non-human animals.

So, there are important differences between non-human animals being put at risk of being killed in order to prevent something worse happening to human beings and people being put at risk of being killed at war because of a military draft.

Thus, I don't think Varner has shown that Regan or other proponents of versions of MRV are committed to a large amount of animal experimentation and so he hasn't shown that Utilitarianism and MRV exhibit some non-trivial degree of convergence.

5.2 Kinds of Practical Convergence

There are two broad ways in which two normative ethical principles can come to have some level of practical converge. First, there can be some practical converge because there is partial explanatory convergence (e.g., theories appeal to the same values or they agree on one or more fundamental right- or wrong-making properties). Second, there can be some practical convergence without any agreement at the explanatory level.

In this section, I argue that even if there was some non-trivial amount of practical convergence between Utilitarianism and MRV, it would only be the second kind of

---

[236] For more on this issue, see Christiano (2020).

convergence. If this is true, then the practical convergence will not satisfy proponents of either Utilitarianism or MRV that there is any *meaningful* convergence because there will still be disagreement about what matters most, i.e., the fundamental right- and wrong-making properties.

Recall that the fundamental explanatory right- and wrong-making properties of Preference Utilitarianism have to do with maximising (or failing to maximise) overall or aggregate desire satisfaction. On the other hand, the fundamental explanatory right- and wrong-making properties of the worse-off principle have to do with preventing or allowing a harm to befall an inherently valuable individual that is worse than the harm suffered by any other single inherently valuable individual involved in one's action. It might seem that there is some degree of explanatory convergence once we understand harm as desire frustration. However, as we saw above, this is not how Regan understands harm. For him, desire/ preference satisfaction is only one part of what harm consists in for Regan. There is also *taking satisfaction* in getting one what one wants and actually getting things in one's interest (e.g., health). Moreover, it is only harm done to individuals with *inherent value* and therefore a *moral right* against being harmed. So, it's not even harm *per se* that fundamentally matters, but only harm done to those with inherent value. Thus, the worse-off principle is silent about the harm done to individual creatures that lack inherent value. On this view, the fundamental explanatory right- and wrong-making properties of the worse-off principle have to do with preventing or allowing a single individual to be harmed to a greater degree than any other single individual involved in the action. So, while any creatures' desires matter for Preference Utilitarianism, the worse-off principle only cares about the desire of inherently valuable individuals.

Moreover, Preference Utilitarianism's fundamental right- and wrong-making properties have to do with *overall* or *aggregate* desire satisfaction. Therefore, the desire satisfaction of

any single individual doesn't matter in their own right. They only matter as a part of the sum of the desire satisfaction. Any action to any person can be morally required as long as it results in enough desire satisfaction. However, according to the worse-off principle, what matters is the desires of *each individual* person. The amount of overall or aggregate desires is not morally important. Thus, one cannot do just anything to a person because enough desires from other people are satisfied.

Second, Preference Utilitarianism holds that the only thing that has *intrinsic* value is desire satisfaction while the worse-off principle holds that the only thing that has *inherent* value are beings that are subjects of a life. While both intrinsic and inherent value are ways of being finally valuable, they disagree agree about the source of the final value. Intrinsically valuable things are finally valuable in virtue of internal properties and inherently valuable things are finally valuable in virtue of being the subjective of a life. Therefore, these views don't even agree on what kind of value matters for determining the rightness or wrongness of actions. Thus, we can see that even though both principles involve desire satisfaction (to some degree), they are quite different in what think is fundamentally morally important.

It is important to see that Preference Utilitarianism and the worse-off principle disagree about what count as fundamental right- and wrong-making properties, because it is the instantiation of these properties that proponents of each view really cares about. That is, what Preference Utilitarians really care about is the maximisation of overall desire satisfaction and what proponents of the worse-off principle really care about making sure that no single individual suffers harms that are worse than those suffered by other individuals in the same situation. Thus, it is misleading to say that these two views "converge" or "agree" on the moral status of some action. What matters is not calling an action "right" or "wrong" but rather whether the action maximises desires satisfaction (if one is a Preference Utilitarian) or

prevents a single individual from being harmed to a greater extent than any other single individual involved in the action (if one is a proponent of the worse-off principle).

As mentioned in the previous chapter, we can get a better sense of why practical convergence without explanatory convergence will not satisfy proponents of either view by considering a disagreement in aesthetics. Imagine that two art critics both think that a painting is beautiful. The first critic thinks that the painting is beautiful because of the gracefulness of its lines and brush strokes and while the second critic thinks that the painting is beautiful because of its evocative colour palette. Imagine further that the first critic doesn't think that the colour palette at all contributes to making the painting beautiful while the second critic doesn't think the lines or brush strokes at all contribute to making the painting beautiful. So, while there is a kind of superficial convergence on the question of whether the painting is beautiful, there is a deep disagreement remaining between the critics. It is this kind of deep disagreement that also remains between Preference Utilitarianism and the worse-off principle.

Notice that proponents of MRV or the worse-off principle will reject the fundamental right- and wrong-making features of Preference Utilitarianism. First, they will reject the idea that *only* desire satisfaction matters fundamentally. Second, they will reject the idea that the *aggregate* amount of desire satisfaction matters fundamentally. Only that matters fundamentally for the proponent of the worse-off principle are individuals with inherent worth. Likewise, Preference Utilitarianism will reject the fundamental right- and wrong-making features of the worse-off principle. First, they will reject the claim that what matters fundamentally has anything to do with "inherent worth." Second, they will reject the claim that we should only compare harms on a one-to-one basis and thus never aggregate harm. Third, they will reject the claim that what matters fundamentally is taking (or being able to take) satisfaction in one's desires/preference being fulfilled or obtaining what is objectively in

one's interest. Thus, not only do these two theories disagree about what fundamentally matters, they think the other theory focuses on morally-irrelevant features.

It is true that practical convergence without explanatory convergence might satisfy some who are most concerned with promoting animal experimentation policies. This is because they might be able to persuade critiques by pointing out that even the critic's preferred view agrees on what policy is best. Of course, as I argued above, I don't think that proponents of a version of MRV will be convinced by Varner's argument because the worse-off principle is not the only MRV-friendly principle that applies to animal experimentation.

Moreover, the convergence might not even be particularly helpful to those concerned with promoting certain policies. This is because the kind of convergence exhibited between Preference Utilitarianism and the worse-off principle is accidental. That is, the theories could easily cease agreeing on certain ethical matters. For example, imagine that both theories initially agree that some experiment is morally required because it will save 10 human lives at the cost of 5 non-human animal lives. Then imagine that once the experiment begins and people learn about the death of the 5 non-human animals, there is public outrage and continuing the experiment will now result in an overall amount of desire *frustration* and not satisfaction. This would entail that Preference Utilitarians would cease to support the experiments, but proponents of the worse-off principle would continue supporting it—after all, they don't care about *aggregate* desire satisfaction or *only* desire satisfaction. Likewise, imagine that after the experiment started, the experimenters learned that the experiment could no longer save humans lives, but could save 6 non-human lives. In this case, Preference Utilitarians would agree that the experiment would be required, but proponents of the worse-off principle wouldn't have anything to say because that principle only applies to cases in which unequal harms will befall the animals to be experimented on and the animals

benefitting from the experiment. And, if they endorsed other principles related to MRV, they would likely claim that the experiment is no longer permissible.

However, policies are meant to be long-term and stable. So, simply showing that Preference Utilitarians and proponents of the worse-off principle agree *now* about the morality of some animal experimentation wouldn't be enough to justify enacting that policy. What policymakers would need to see is that the two ethical views would *continue* to agree on the policy as time went on. However, in many cases, this will not be true because whether an action is required or wrong according to these two views is contingent (e.g., on whether they will save human lives and how many).

Finally, because the convergence is not based on any meaningful agreement about what the right- and wrong-making properties are, the convergence is merely accidental. The fact that convergence is merely accidental is unlikely to satisfy anyone who wants to argue for convergence in order to curb concerns about scepticism about moral facts or moral knowledge. The idea is that agreement that is merely accidental doesn't show that ethicists on both sides of the debate are agreeing with each other, it just means that sometimes the properties that they respectively care about overlap. But showing that accidents like this happen will do little to assuage any worries caused by peer disagreement. After all, the kind of agreement that this kind of convergence constitutes is like the agreement between the two art critics: they're willing to say that the painting is beautiful, but they each think the other is wrong about what matters for making something beautiful and thus what beauty consists in.

6. Conclusion

In this chapter, I examined Gary Varner's argument that there can be a non-trivial amount of practical convergence between Utilitarianism and the Moral Rights View (MRV) on the issue of animal experimentation by looking at Preference Utilitarianism and the worse-off principle. I argued that even appealing to the above these forms of Utilitarianism and MRV

are not sufficient for showing that Utilitarianism and MRV exhibit some non-trivial degree of convergence on the issue of animal experimentation. First, the worse-off principle is not by itself plausible as a *complete* version of MRV, i.e., proponents of MRV will endorse further principles and these further principles will prevent convergence with Preference Utilitarianism. Second, I argued that even if Varner is correct that there is some convergence between Preference Utilitarianism and the worse-off principle, it is not the kind of convergence that would satisfy proponents of either view. This is because proponents of each view really care about whether an action instantiates their preferred right- or wrong-making properties.

Chapter 5: Caring and Convergence in Environmental and Animals Ethics

1. Introduction

In this thesis, I have argued that despite the fact that there is at least non-trivial convergence between Broad Anthropocentrism and Non-Anthropocentrism, philosophers have failed to show that there is *total* or *mass* practical or explanatory convergence between normative ethical principles in environmental ethics. Moreover, I argued in Chapter 4, that philosophers have not shown that there is even *non-trivial* convergence in animal ethics (concerning animal experimentation).

In Chapter 1, I distinguished practical convergence from explanatory convergence. Practical convergence occurs when two or more normative ethical principles agree on the answer to the question, "What (morally) should I do?" Explanatory convergence occurs when two or more normative ethical principles agree on the answer to the question, "What explains why I (morally) should do what I should do?" I argued that it was implausible that well-known normative ethical theories (e.g., Utilitarianism, Kantianism, and Contractualism) exhibit mass or total explanatory convergence by considering the available options for how these theories might exhibit explanatory convergence and also considering Derek Parfit's argument that these three views (Utilitarianism, Kantianism, and Contractualism) exhibit some kind of explanatory convergence. Two or more normative ethical theories exhibit total explanatory convergence when they always agree about why some action is right or wrong. Two or more normative ethical theories exhibit mass explanatory convergence when they often agree about why some action is right or wrong. I concluded that it was more likely that different normative ethical theories would exhibit mass or total *practical* convergence.

In Chapters 3 and 4, I considered arguments for practical convergence in environmental ethics and animal ethics and argued that these arguments are unsuccessful. In Chapter 3, I argued against Bryan Norton's claim that there is mass or total practical convergence between Weak Anthropocentrism and Non-Anthropocentrism. Such convergence counts as total when these normative ethical principles *always* agree on the answer to the question, "What (morally) should I do?" and convergence counts as mass when they often agree on the answer to this question. Weak Anthropocentrism, roughly, is the view that what matters for figuring out how one ought to treat nature is determined solely by the desires that a person would have if she were rational and well-informed. Non-Anthropocentrism is the view that nature is intrinsically valuable and that both human interests and nature matter in their own right. I argued that one can be rational and well-informed and still fail to desire that nature to be treated well. So, it is implausible that Weak Anthropocentrism and Non-Anthropocentrism will exhibit mass or total practical convergence. However, I argued that what I called Broad Anthropocentrism and Non-Anthropocentrism exhibited a non-trivial amount of practical and explanatory convergence as well as a non-trivial amount of convergence on the question of how valuable nature is.

In Chapter 4, I argued against Gary Varner's view that Preference Utilitarianism and the Moral Right View exhibit non-trivial practical convergence on the question of whether people should engage in animal experimentation. Two or more normative ethical theories exhibit trivial convergence when they agree on the answer to the question, "What (morally) should I do?" for some small number of cases. Two or more normative ethical theories exhibit non-trivial practical convergence when they agree on the answer to the practical question in more than a trivial amount of cases, but still less than many cases. Preference Utilitarianism claims that what one (morally) should do is to maximise overall desire satisfaction in all cases. The Moral Rights View holds, roughly, that certain non-human animals are inherently valuable

and therefore have a moral right not to be harmed, even if doing so maximising goodness (e.g., desire satisfaction). Varner argues that if we understand the Moral Rights View as consisting of a certain principle (i.e., the worse-off principle), then the Moral Rights View and Preference Utilitarianism will agree on the answer to the question "Morally speaking, should I (or other people) conduct harmful animal experimentation?" I argued that the worse-off principle is not a plausible version of the Moral Rights View and so Varner fails to show that Preference Utilitarianism and the Moral Rights View exhibit even non-trivial practical convergence.

Given the fact that philosophers have failed to show that there is mass or total practical convergence between normative ethical theories on issues in environmental ethics and animal ethics, I think we have good reason to look for alternative ways that these normative ethical theories might exhibit mass or total convergence. Inspired by McShane (2007), I think we should look at what normative ethical theories in environmental ethics and animal ethics have to say about what we should *feel* and not just what we should *do*. Even though these normative ethical theories were not designed to answer questions about one should feel, all of them make claims about: (1) what kinds of things have value and (2) what kind of value those things have. With these value claims in mind, we can get a better sense of what these normative ethical theories think we should care about.

In this chapter, I will argue for two convergence claims. First, I will argue that Anthropocentrism and Non-Anthropocentrism will exhibit total convergence on the answer to the question, "Should I *care* about nature?" In particular, I will argue that they both entail that one should care about nature. And, I argue that they exhibit some form of explanatory convergence, i.e., they converge (to some degree) on the question of why one should care about nature. Second, I will argue that Preference Utilitarianism and MRV will exhibit total convergence on the answer to the question, "Should I care about animals (or animal

welfare)?" In particular, I will argue that they both entail that one should care about animals (or animal welfare). And, I argue that they exhibit some form of explanatory convergence, i.e., they convergence (to some degree) on the question of why one should care about animals. Thus, while these theories might not exhibit mass or total convergence on what to do with nature and animals, they exhibit mass or total convergence on how to *feel* about nature and, to some degree, on *why* we should feel a certain way toward nature.

Before making my argument, it is important to address two potential worries. The first worry is this: one might doubt that how one should feel is morally important. However, McShane points to several reasons why feelings can be morally important that have nothing to do with action.[237] First, we can morally assess people based on their feelings. For example, if someone hates people of a certain race just because they belong to this race, then we can negatively assess this person for having such feelings. And, we can assess them negatively even if their feelings do not motivate immoral action. Second, being a virtuous person requires having the right feelings. For example, a virtuous person will be concerned about other people's welfare and be motivated to help others while a vicious person might be motivated to harm others.[238] Third, when we are trying to figure out how to live good lives, we often ask ourselves how we should feel, "Should I feel indignant about the mistreatment of my friend?", "Should I feel remorse for lying?", or "Should I empathise with people in distant countries?"[239] Therefore, it seems like how we feel is morally important because they can be good/bad or criticisable or praiseworthy (e.g., love and hate), they are part of being virtuous or vicious, and they matter for living well or living a good life.

---

[237] McShane (2007): 174-175.

[238] Ibid.

[239] Ibid.

The second worry is this: why should we care about showing that there is convergence between different normative ethical theories on the question of whether we should care about nature or animals? This is an important question. As we saw above, we can morally assess someone for how they feel. For example, we can think that someone is a bad person if they have certain feelings (e.g., hatred toward certain races) or that someone is a good person if they have certain other feelings (e.g., they care deeply about other people's welfare). So, if one cares about being a good person or living well, then one should care about how one should feel.

Moreover, if there is disagreement between philosophers about how we should feel, e.g., whether we should care about nature, we might have reason to doubt that there are certain moral facts, e.g., how a good person should feel about other people, nature, or non-human animals. What makes this potential disagreement concerning is that proponents of different normative ethical theories likely have all the same morally-relevant information and so if they were to disagree about when we should care about things, this might cast doubt on certain moral facts about what we should care about. This motivation for exploring the question of how to feel is analogous to Derek Parfit (2011)'s motivation for exploring whether normative ethical theories converge on what to do.[240] Parfit was concerned by the fact that equally smart and well-informed moral philosophers disagreed about what actions were morally right and wrong. Analogously, I think it would be concerning if these philosophers disagreed about whether to care about certain things, e.g., nature and animals. Moreover, if one could show that even seemingly opposing normative ethical theories agree that we should care about animals or nature, then this might inspire others to care as well or at least consider whether to care. And, even those who are sceptical about whether we should care about animals or nature, would have some reason to doubt their view.

---

[240] Parfit (2011): 418-419.

Here is the structure of this chapter. First, I provide important background information. I explain what is involved in caring about something, what Anthropocentrism, Non-Anthropocentrism, Utilitarianism, and MRV hold, what I mean by saying that one "should" care about nature, and I explain the different ways that normative ethical theories can convergence on the answer to the question, "How should I feel about something?" Second, I argue that Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I finally care about nature?" I also argue that they exhibit total convergence on one answer to the question, "Why should I care about nature?" Third, I argue that Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animals?" I also argue that they exhibit total convergence on one answer to the question, "Why should I care about animals?" Fourth, I answer an important objection about the reason for which people should care about nature and animals.

2. Background

Broadly speaking, my main thesis in this chapter is that certain normative ethical theories converge on questions about what one should care about. In order to make my thesis clear, I will need to explain a few things. First, I need to explain what is involved in caring about something. Second, my argument will rely on the claim that if something is valuable, then it is appropriate to care about it. Therefore, I need to show that all the normative ethical theories I discuss make claims about has value. I will then argue that these views entail that nature and non-human animals have value. Finally, I need to explain what it means for two theories to exhibit total convergence on the answer to the questions, "Should I care about nature?" and "Should I care about animals (or animal welfare)? " With these issues explained, I will then present my main thesis.

2.1 Caring

Caring about something consists in having a set of attitudes and dispositions to have certain attitudes about what one cares about. While there are numerous theories of what caring consists in, I will rely on claims that many theories of caring agree on.[241] First, if one cares about X, then one thinks that X is important and worthy of care.[242] For example, if you care about your pet, you will see him or her as being important and worthy of care. Second, if one cares about X, then one is disposed to have certain emotions depending upon how well X is doing. For example, if X is harmed or damaged, one will be disposed to experience negative emotions (e.g., sadness or anger). If X flourishes or does well, then one will be disposed to experience positive emotions (e.g., happiness or joy). For example, people care a lot about their children and when their children fail or are injured, people tend to be upset. Third, if one cares about X, then one is disposed to desire that X flourishes or does well and to desire that X is not damaged or destroyed. For example, parents strongly desire that their children are happy and successful. Fourth, one is disposed to be motivated to promote X's flourishing and to protect X from damage or destruction. For example, parents are motivated to help their children succeed and avoid injury. Likewise, if someone cares about their expensive cast iron pan, then they will be motivated to do things to keep it in good condition, e.g., drying it after washing it to prevent rusting, using the correct method of cleaning it, and coating it with carbonised oil and so on.

There are two broad ways in which one can care about X. One can *instrumentally* care about X and one can *finally* (or non-instrumentally) care about X. When one instrumentally cares about X, then one cares about X only because X has the potential to be used to give something else that one cares about. For example, this is normally the correct way to care about money. This is because money, under normal circumstances, is only instrumentally

---

[241] What I call "caring" is sometimes called "valuing." For different accounts of what is involved in caring about something or valuing it, see Scheffler (2011), Seidman (2009, 2016), and Jaworska (2007).

[242]Jaworska (2007): 563, n. 97, 564 and 562, n. 94, Seidman (2009): 285, Scheffler (2011): 32.

valuable. If X is instrumentally valuable, then it is appropriate to instrumentally care about X. If you only care X instrumentally, then you could easily case to care about X if something else, e.g., Y, could equally be used to bring something else about that one cares about. For example, one instrumentally cares about the particular money notes in one's wallet, because, if they were to be replaced with different money notes of equal monetary value, one would not care. One would instantly cease caring about the original money notes and start caring about the new ones.

When one finally (i.e., non-instrumentally) cares about X, then one cares about X for its own sake and not just because of what it can be used to get. Moreover, one would not cease caring about X even if it could no longer be used to bring about other things that one cares about getting. It is only appropriate or fitting to finally care about X when X is finally valuable. (I will explain exactly what I mean by "appropriate" or "fitting" shortly).

A case will help bring this out.[243] Imagine Lars' first guitar. He got the guitar as a child and spend many hours getting pleasure from playing the guitar. In fact, the guitar was a means of his developing his musical abilities. The guitar also helped him find friends and contributed to him having the opportunity to play music on stage in front of people. Moreover, imagine that he played this guitar whenever he was having a hard time in life and it always made him feel better. This his experience with this guitar and the important role it played throughout his life, it's no surprise that he cares about the guitar for its own sake and just because of what it has the potential to bring about now.

Samuelsson (2008) argues that three things are involved when you finally value X. First, even if you don't use X anymore or take advantage of what X can be used to bring about, you would still be sad or disappointed if X were stolen or destroyed. Second, one hopes that X is taken care of even after one dies. Third, one would not trade X for something

---

[243] This case comes from Samuelsson (2008): 188.

of the same kind, e.g., Y, even if Y had a higher monetary value and even if Y were more instrumentally valuable.

We can easily imagine that Lars would be sad if his guitar was stolen, even if he stopped playing it regularly. We also imagine that he'd want someone to take care of his guitar even if when he passes away. At the very least, he would not want someone to just throw it in the trash. Finally, he wouldn't sell his guitar even if he were offered more money than it was worth (at least under normal circumstances), nor would he trade his guitar for a guitar that sounded better or was easier to play.

The difference between instrumentally caring about X and finally caring about X is important because I'm going to argue that Anthropocentrists and Non-Anthropocentrists can agree that nature is finally valuable and therefore, one has reason to finally care about nature. However, I will argue that even though Utilitarians and proponents of MRV will agree that non-human animals are valuable (in fact very valuable), they will disagree about whether animals are finally valuable and thus they will disagree about how to care about them.

2.2 Normative Ethical Principles and Value

In Chapters 1, 3, and 4, I explained what normative ethical theories are concerned with and what Anthropocentrism, Non-Anthropocentrism, Utilitarianism, and MRV hold. However, I want to focus on the fact that normative ethical theories also often make claims about what kinds of things are valuable and what kind of value they have.

Anthropocentrism holds that nature is valuable only insofar as it can be used as means for serving the interests of humans.[244] In other words, nature is only *instrumentally* valuable. However, as I argued in Chapter 3, I think a plausible version of Anthropocentrism (i.e., Broad Anthropocentrism) holds that nature is finally valuable. Non-Anthropocentrism holds

---

[244] Norton (1984) and McShane (2007). This is the standard account of Anthropocentrism. As we saw in Chapter 3, Norton (1984) gives an alternative version. I focus on the standard account because it is less controversial.

that humans and their interests have intrinsic value but that nature is also instrumentally valuable. Hedonistic Utilitarianism holds that the only things that have intrinsic value are experiences of pleasure.[245] Moreover, it does not matter whose pleasure it is and so both human and non-human animal pleasure count as intrinsically value. Preference Utilitarianism holds that the only things that have intrinsic value are the satisfaction of desires.[246] However, this view is neutral about whose desires count morally. That is, it allows that the desires of any creatures to be intrinsically valuable. So, the satisfaction of animal desires is intrinsically valuable as is the satisfaction of human desires. MRV holds all creatures that are a subject of a life have inherent value.

Figuring what each theories thinks has value, because, as I will soon argue, if something is valuable, then it is worthy of some degree of care. And, depending on what kinds of value it has (i.e., final vs instrumental), it is worthy of being caring about in a certain way.

2.3 The Fittingness or Appropriateness of Caring

I will argue that certain normative ethical theories will convergence on the answer to the questions, "Should I care about nature?" and "Should I care about non-human animals?" What do I mean by "should" in this context? What I have in mind is not the claim that one is *required* (morally or otherwise) to care. Rather, I have a weaker claim in mind. What I mean is that it is fitting or appropriate to care about nature or non-human animals.[247] This is the sense of fitting or appropriate in the same sense that it is: (a) fitting to enjoy delicious food, beautiful artwork, entertaining films or (b) appropriate to blame wrongdoers, and praise good people. It's not as if one is *required* to enjoy these things or to blame/praise people. Rather,

---

[245] Sinnott-Armstrong (2021).

[246] Rabinowicz and Österberg (1996).

[247] For more on fittingness, see Howard (2018).

one has a reason to enjoy them. In particular, one has a reason to enjoy these things such that when one enjoys them one is *justified* in enjoying them. So, it's not just that one has *a* reason to do these things, one a reason that is sufficient to *justify* one's doing them.[248]

2.4 Convergence About How to Feel

There are numerous ways in which normative ethical theories can exhibit convergence about how to feel. First, they can exhibit convergence on the answer to the question, "How should I feel about X?" For example, they can agree that one should care about X, desire X, love X, etc. Second, they can exhibit convergence on the answer to the question, "Why should I feel that way about X?" That is, they can agree on why one should care about, love, desire, or admire X. For example, two views might agree that I ought to admire someone because they are hardworking or talented.

Like practical convergence, convergence about how and why to feel a certain way is a matter of degree. Two or more normative ethical theories exhibit *total* convergence on the answer to the question, "How should I feel about X?" when they agree on all of the kinds of feelings that one should have concerning X. For example, two normative ethical theories might agree that one should only love and admire X. Two or more normative ethical theories can also exhibit total convergence about whether to feel a certain way about X. Two or more normative ethical theories exhibit total convergence on the answer to the question, "Should I feel way W toward X?" if they agree on all the circumstances in which one should feel way W toward X. For example, two theories might agree on all the circumstances in which one should admire another person.

Two or more normative ethical theories exhibit *mass* convergence on the answer to the question, "How should I feel about X?" when they agree about many or most of the kinds of feelings that one should have concerning X. For example, two normative ethical theories

---

[248] I follow Gert (2007) in claiming that a reason can justify an action or reaction without requiring it.

might agree that one should love X, but disagree about whether one should admire X. Two or more normative ethical theories can also exhibit mass convergence about whether to feel a certain way about X. Two or more normative ethical theories exhibit mass convergence on the answer to the question, “Should I feel way W toward X?” if they agree on many or most of the circumstances in which one should feel way W toward X. For example, two theories might agree on many or most of the circumstances in which one should admire another person.

There is also a kind of convergence that exists for feelings, but not for action. This is convergence on how intense one's feeling should be. That is, two or more normative ethical theories can convergence on the answer to the question, "How much should I feel way W toward X?" For example, two normative ethical theories might convergence or diverge on how angry one should be about a betrayal or how happy one should be about some success. Important for our purposes is that two normative ethical theories can convergence on *whether* one should feel a certain way toward an object but diverge on the intensity of the feeling one should have. For example, both theories might agree that I should feel angry with a person, but one theory might hold that I should only be a little angry while other theories hold that I should be very angry. Yet, they still converge on the answer to the question, “Should I be angry with this person?”

Just as there is explanatory convergence concerning practical convergence, there is also explanatory convergence concerning how one should feel. First, two or more normative ethical theories can exhibit total or mass convergence on the answer to the question of “Why should I have this particular set of feelings about X?” For example, they might hold that the fact that someone is kind explains why one should have all the feelings one should have about that person, e.g., why one should admire, praise, respect, and love that person.

Second, two or more normative ethical theories can exhibit total or mass convergence on the answer to the question of "Why should I feel way W toward X?" This occurs when these theories agree on either all or many/most of the reasons why one should feel a certain way about a certain object. For example, two or more normative ethical theories might hold that one ought to be angry with a person only in circumstances when they mistreat you. This is the kind of explanatory convergence I will focus on in this chapter.

Third, there are two kinds of explanations of why I should feel way W toward X: a derivative or non-fundamental explanation and a non-derivative or fundamental explanation. Two or more normative ethical theories exhibit derivative explanatory convergence when they agree on why one should feel a certain way, but the only reason that this explanation is an explanation of why one should feel a certain way is because of some further, more fundamental properties. For example, take the case of admiring two beautiful paintings, e.g., the Mona Lisa and Starry Night. I think it is clear that one should admire these paintings and one explanation of why one should admire them is because they are beautiful. However, this explanatory is only derivative because there is a more fundamental explanation, i.e., the explanation of why the paintings are beautiful. Whatever explains why these paintings are beautiful is the fundamental explanation of why one should admire them.

Two or more normative ethical theories exhibit fundamental explanatory convergence when they agree on why one should feel a certain way and there is no further explanation of why we should feel a certain way. Let us return to the case of admiring the Mona Lisa and Starry Night. Imagine that I explain why you should admire the Mona Lisa by listing the properties of the painting that make it beautiful, e.g., the particular brushwork, the balance of light and shadow, the arrangement of certain shapes and colours, and so on. In this case, it doesn't seem like there is a further explanation to be given of why these properties make it so that I should admire the painting.

2.5 Clarifying My Argument

With this terminology on the table, I can now more clearly state my main claims in this chapter. First, I will argue that Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I *finally* care about nature?" That is, they will agree about all the circumstances in which one should care about nature. Second, I will argue that Preference Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animal welfare?" That is, they will agree about all the circumstances in which one should care about animal welfare. Third, I will argue that Anthropocentrism and Non-Anthropocentrism exhibit total non-fundamental explanatory convergence between they agree that nature is very, finally valuable. Fourth, I will argue that Preference Utilitarianism and MRV exhibit total non-fundamental explanatory convergence because they argue that nature is very valuable. However, I will note that they don't converge on the question of what kind of value animals have and thus they don't converge on the question of how to value animals (i.e., either instrumentally or finally).

3. Convergence Between Anthropocentrism and Non-Anthropocentrism

3.1 Convergence on Whether to Finally Care

In this section, I will argue that Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I finally care about nature?" That is, they agree on all the circumstances in which I should finally care about nature.

It is not surprising to claim that Non-Anthropocentrism entails that one should care about nature. This is because, as we saw above, Non-Anthropocentrism holds that nature is intrinsically valuable and it seems quite plausible that one should care about things that are intrinsically valuable. For example, it seems plausible that one should care about masterpieces of art, music, and literature because they are intrinsically (or finally) valuable.

And, therefore, if Non-Anthropocentrism is correct that nature is intrinsically valuable, one should care about nature.

McShane (2007) claims that there is a close relationship between X being valuable and it being true that one should care about X. She writes:

> Claims about why something has value are claims about why we, as moral agents, have reason to care about the thing. More precisely, they are claims about why the thing is worth caring about.[249]

There are two ways of reading McShane. On one reading, she is making the stronger claim that claims about why something has value *are the same thing* as claims about why we have reasons to care about things. In other words, claiming something about something value is just to make a claim about what reasons people have to care about it. On a second, weaker, reading, McShane is claiming that the following conditional holds: if something has value, then one should (i.e., has justifying reason to) care about it. I will only rely on this second, weaker, claim.

So, it seems pretty straightforward that Non-Anthropocentrism entails that one should care about nature. But why think that Anthropocentrism entails that one should care about nature? After all, Anthropocentrism claims that nature is only instrumentally valuable. In particular, it claims that nature is valuable only insofar as it is a means of serving human interests.

In fact, McShane (2007) provides some reason to think that believing in Anthropocentrism is incompatible with caring about nature in the right way. First, she says that if Anthropocentrism is true, then one should believe that nature is valuable only insofar as it can benefit human interests. Second, she notes that certain kinds of feelings require that one think that the object of those feelings is finally valuable and not just valuable only because it can benefit one. She focuses on feelings of love, respect, and awe. She writes:

---

[249] McShane (2007): 172.

> But to love a friend is in part to deny that her value is just a matter of her serving your interests . . . I think that there are other attitudes besides love of which this is also true. Respect certainly seems to work this way; awe (at least in some manifestations) might do so as well. To respect something is in part to see it as making a claim on your moral attention in its own right. It is to attribute to the thing a kind of independent standing in your scheme of 'things that matter'. To be in awe of something is in part to see it as having a kind of greatness that goes beyond you - beyond your needs, interests, or attitudes.[250]

While she doesn't say that one cannot properly care about something unless one thinks that that thing is finally valuable, one can imagine that something similar can be said about caring. For example, perhaps I don't properly care about my friend is I believe she is valuable only as a means to benefitting me.

I think McShane's argument is implausible because I think she confuses the following two claims:

1. I should care about nature because it benefits *me*.
2. I should care about nature because it benefits human interests (including my own).

It might be true that one cannot believe that nature is only valuable because it serves one's interests and at the same time really care about nature. However, it does not follow that one cannot believe that nature is valuable only because it serves the interests of many or most humans. This is the difference between caring about something in a fully selfish way and caring about it in a partly selfless or other-directed way.

In addition, I think McShane fails to consider all the kinds of benefits that nature might provide to humans and many of the kinds of benefits it provides seem to make it nature worthy of care. It might be that there is something bad about believing that nature is only valuable because it serve human interests in simple pleasures (e.g., feeling a cool breeze, seeing a pretty mountain range, or hearing a pleasant bird song) and really caring about nature. However, I think that Anthropocentrism need not restrict the human interests it is

---

[250] McShane (2007): 176.

concerned with to interests in experiencing simple pleasures. Rather, it also considers human beings' interest in having meaningful experiences, feeling like part of something greater than oneself, experiencing beauty, preserving part of one's cultural heritage, and so on.[251]

I think it is clear that nature can serve these higher human interests. For example, many people think that their experiences in nature contribute to the meaningfulness of their lives or give them the sense that they belong to something greater than themselves. Nature provides many opportunities to experience beauty and it's plausible that experiencing beauty is finally valuable. For example, thinking about experiencing the beauty of masterpieces of art. Many pieces of nature (e.g., parts of land or trees) belong to a person's cultural heritage and so parts of nature are valuable as parts of one's culture or heritage.

Moreover, Anthropocentrism can also hold that nature is very valuable because nature is necessary for many finally valuable things to exist. For example, if nature were in a state of decay, e.g., because of global warming or air pollution, this negatively affects the health and well being of millions of people. The health of nature can be tremendously valuable according to Anthropocentrism because it is required for so many finally valuable things to exist.

Therefore, it seems to me that Anthropocentrists can claim that nature is tremendously valuable, even if this value is only instrumental value, and that being tremendously instrumentally valuable makes it so that people should care about nature. If this is true, then Anthropocentrism will claim that one should care about nature.

However, as we saw in Chapter 3, I think more can be said about Anthropocentrism. I think that we need to recall that the core value claim of Anthropocentrism need not be that nature's value is determined solely by how it can be used to serve human needs. Rather, I

---

[251] For examples of how people think nature or experiences with nature contribute to the meaningfulness of their lives, see Chan et al. (2016).

think the core value claim is just that whatever nature has, it must be explained by some relationship it has with human beings. Standard Anthropocentrism seems to assume that the only relationship with human beings that matters for determining nature's value is an instrumental relationship. But, as I argued in Chapter 3 and as I will briefly recap here, I think this is a mistake.

Recall that there are two ways of being instrumentally valuable. The first is having the *potential* to bring about something else of value. The second is having *actually* already brought about something else with value. As I argued in Chapter 3, things with the second kind of instrumental value can sometimes have final value. For example, Kagan (1998) gives the case of the pen that Abraham Lincoln used to sign the Emancipation Proclamation, which lead to millions of American slaves being freed. I noted that the reason that Lincoln's pen has final value is that the following related facts obtain: (1) the finally valuable object played an important role in actually bringing about something else with final value, (2) the finally valuable object *uniquely* played an important role in actually bringing about something else with final value, (3) the thing it brought about was tremendously valuable, and (4) the thing it brought about benefited all human beings.

In the case of Lincoln's pen, it played an important role in freeing millions of American slaves (which is finally valuable). It uniquely played this role because it's not as if other pens were used to sign the proclamation. The freeing of millions of American slaves was not only finally valuable, it was also tremendously valuable. Finally, it benefitted all human beings.

Finally, I argued that nature seems to share these properties with Lincoln's pen. First, nature played an important role in sustaining something of final value (i.e., all human life). Second, it played a unique role in doing this. This is because nothing else is responsible for making the same kind of contribution to the sustaining of human life that nature has. After all only nature provided oxygen, drinkable water, edible plants and animals, comfortable

climates, materials for building shelter and tools, and so on. Third, the finally valuable thing it sustained was tremendously valuable. This is because nature helped to sustain innumerable humans being—including all currently-living humans—and human lives are tremendously valuable. And, fourth, the finally valuable thing it sustained benefited all humans. This is because all of us have benefited from the climate, food, water, shelter, medicine, and materials that nature has provided that has helped us remain alive.

Thus, I claimed that a version of Anthropocentrism (i.e., Broad Anthropocentrism) entails that nature is finally valuable. In fact, it entails that nature is very, finally valuable. But, because the source of nature's final value is solely based on its previous or current relationship with humans, this view is completely Anthropocentric.

Anthropocentrism entails that nature is finally valuable because being intrinsically valuable is just a way of being finally valuable. Moreover, as we saw in Chapter 3, Anthropocentrists will claim that nature is also very valuable. In Chapter 3 and as recapped here, Broad Anthropocentrism also holds that nature is finally valuable and indeed very valuable. So, it looks like Broad Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the question of whether nature is very, finally valuable.

Because these two views exhibit total convergence on the question of whether nature is very, finally valuable, they also exhibit total convergence on the answer to the question, "Should I finally care about nature?" The final value of nature and finally caring about nature are linked by a principle like the following:

> *Value-Caring Link*: If X is finally valuable, then one should (i.e., it is fitting to) finally care about X.

Thus, we have total convergence on the question of finally caring about nature.

3.2 Convergence on Why to Finally Care

In this section, I will argue that there is total *derivative* explanatory convergence on Anthropocentrism and Non-Anthropocentrism on the answer to the question, "Why should I finally care about nature?" I will then explain why these two views do not exhibit any *fundamental* explanatory convergence.

As we saw above, Broad Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I finally care about nature?" Now the question of whether they exhibit any kind of derivative explanatory convergence depends upon whether they agree to any degree on any explanation of *why* we should care about nature. I think it is clear that they do agree on one explanation. In particular, they both agree that we should care about nature because it is *finally valuable*. And, we saw above, it is plausible that: if X is finally valuable, then one should (i.e., it is fitting to) finally care about X.

However, there is no fundamental explanatory convergence between them. This is because they disagree about what fundamentally explains why nature is finally valuable and therefore why we should finally care about it. Recall that Broad Anthropocentrism thinks that nature is valuable because of its unique and important role in sustaining human life and benefitting all human beings. However, Non-Anthropocentrists think that nature is finally valuable because of features having nothing to do with its relationship with humans.[252]

3.3 Convergence on How Much to Care

Recall that in Chapter 3, it was important to figure out *how* valuable nature is according to both Broad Anthropocentrism and Non-Anthropocentrism. The reason for this was that in order for these two views to exhibit non-trivial convergence on the question of whether one should treat nature well, we needed to know how strong of a reason one has to treat nature

[252] I remain neutral on exactly what properties Non-Anthropocentrists think nature has that gives it intrinsic or final value.

well. This is because there might be countervailing reasons that count against treating nature well and exactly how valuable nature is will determine how often our reason to treat nature well will be defeated by these other reasons. After all, in these cases, we can either treat nature well or fail to treat it well. Those were our only two options.

However, the same doesn't hold for caring. That is, we can finally care about two things even if they are in conflict with each other. I can finally care about my best friend and my sister, even if they are fighting with each other and in fact, dislike each other. While my finally caring about each of them might annoy them, I don't seem to be making any kind of mistake in caring as I do. So, in order to show that Broad Anthropocentrism and Non-Anthropocentrism can exhibit *total* convergence on the question of whether one should finally care about nature, we don't need to figure out how valuable nature is according to each of these views.

Moreover, recall that in order to make any precise claims about just how valuable nature is according to Broad Anthropocentrism and Non-Anthropocentrism, I will have to take on board controversial claims about how to precisely measure value. Such an enterprise is surely worth pursuing, but it would take us too far afield to be worth investigating. This is because we can already see that Broad Anthropocentrism and Non-Anthroponcetrism can exhibit *total* convergence on the answer to the question of whether to finally care about nature regardless of *how* valuable nature is.

What we can say is that it seems clear that both Broad Anthropocentrism and Non-Anthropocentrism will convergence on the somewhat vague claim that nature is "very" valuable. Recall that this is because nature has the properties that make it finally valuable to a high degree according to both Broad Anthropocentrism and Non-Anthropocentrism. For example, Broad Anthropocentrists think that nature played an especially important and unique role in sustaining a tremendous amount of final value that benefitted all humans. Non-

Anthropocentrists think that nature is very complex, very intricately organised without the help of any intention, and very diverse in terms of species and ecosystems. So, both views agree that nature is not merely finally valuable, but very finally valuable.

Given that Broad Anthropocentrism and Non-Anthropocentrism agree that nature is not only finally valuable, but also "very" valuable, they might also exhibit total or mass convergence on the answer to the question, "Should I finally care *a lot* about nature?" Again, it is vague exactly what is it to care "a lot" about something, but there it looks like there is some hope that there is even more convergence between these two views on *how much* to finally care about nature.

4. Convergence Between Utilitarianism and MRV on Whether to Care

4.1 Convergence on Whether to Care

In this section, I will argue that Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animals?" That is, they agree on all the circumstances in which I should care about animals. However, I will argue that they disagree about how I should care (i.e., instrumentally vs non-instrumentally) and about the fundamental explanation of why I should care.

It is not surprising to claim that MRV entails that one should care about animals. This is because, as we saw above, this holds that animals are inherently valuable and it seems quite plausible that one should care about things that are inherently valuable. In particular, it seems plausible that one should care about things that are finally valuable and when something is inherently valuable, it is finally valuable. We saw above that it is plausible that one should care about masterpieces of art, music, and literature because they are finally valuable. We can add to the list of finally valuable things knowledge, pleasure, and living a meaningful life. We can also see that it's plausible that we should care about other entities with not just final value, but inherent value. The most obvious case is human beings. Humans are subjects of a

life according to Regan's definition and so also have inherent value. And, it seems very plausible that we should care about other humans and humankind in general. So, it looks like one should care about animals.

Moreover, there are only a few kinds of beings that can be the subject of life and so only a few kinds of beings that can have inherent value. It is common to think that the rarity of an otherwise valuable object might increase that object's value. For example, Dancy (2003) argues that the value of a book that provides knowledge increases if all the other copies of it are destroyed.[253] Something similar can be said about the inherent value of animals. Given that only two kinds of beings can have inherent value (i.e., humans and mammals over a year old), these beings are more valuable than they would be were to be many more beings that could have inherent value.[254]

Proponents of MRV also think that non-human animals are very valuable. One reason to think this is that, according to Regan (2004), non-human animals and humans have the same level of inherent value.[255] I think we can safely assume that human beings are very valuable. So, we can conclude that proponents of MRV think that non-human animals are very valuable.

So, it seems pretty straightforward that MRV entails that one should care about animals. But why think that Utilitarianism entails that one should care about animals? This is an important challenge because as Regan (2004) argues, Utilitarianism doesn't think that

---

[253] Dancy (2003), 633.

[254] This last point is somewhat controversial because Regan (2004) claims that all entities with inherent value have inherent value equally. First, even if I'm wrong that rarity would increase inherent value, my main point is that being inherently valuable is sufficient for making it the case that we should care about animals. Second, Regan might just be wrong about this. It might be that different beings have different amounts of inherent value. In fact, this is plausible because different beings have the psychological states and abilities to a greater degree than others. For example, humans have more complex beliefs, desires, and emotional lives and they are better at achieving their goals, and so forth. I thank Robert Cowan for bringing this last thought to my attention.

[255] Regan (2004): 241-245.

individual animals are valuable themselves or that animal kind is valuable itself. Rather, as Singer (1993) argues, animals are "mere receptacles" of what is intrinsically valuable. In the case of Hedonistic Utilitarianism, they are mere receptacles of pleasure. In the case of Preference Utilitarianism, they are mere receptacles of desire satisfaction. Either way, it is not them that are valuable. Regan (2004) gives the following helpful analogy:

> Suppose we think of moral agents and patients as cups into which may be poured either sweet liquids (pleasures) or bitter brews (pains). At any given time, each cup will have a certain hedonic flavour: the liquid it contains will be more or less sweet or bitter. Now, what we are to aim to bring about, according to hedonistic Utilitarianism, is not the best-tasting liquid for this or that particular individual; rather, what we must aim to achieve is the best aggregated balance of the sweet and the bitter among all those individuals affected by what we do; it is the best total balance of the sweet over the bitter that we aim to realise. That being so, there is no reason why it may not be necessary to redistribute the contents of any given cup among the others or, indeed, why it may not be necessary to destroy a given cup ("receptacle") quite completely.[256]

So, we see that the real value of animals, according to Utilitarianism, is as receptacles of what is intrinsically valuable.

If this is correct, then animals, according to Utilitarianism, are not intrinsically or finally valuable. Then the worry is: if animals are not intrinsically or finally valuable, why should we care about them? However, I don't think that this means that we shouldn't care about them. Even if non-human animals are mere receptacles of intrinsic value, they are still themselves valuable, just not *finally* valuable. What kind of value do they have? We can think of the value that animals (and even humans) have as a kind the value something gets by providing a place for value to come into existence. That is, they are valuable because they are well-suited to provide a space or container for something that is finally valuable to exist or it provides the opportunity for something finally valuable to come into existence. Following Raz (2011), we can call this *facilitative* value.[257] As Raz notes, this is the kind of value that

---

[256] Regan (2004): 205-206.

[257] Raz (2011): 220.

something has when it can be used to create or bring about something else of value or when it creates opportunities for something of value to be achieved.

Some cases will help to illustrate what I have in mind. First, consider a case based on one from Zimmerman and Bradley.[258] Imagine that you allow a struggling, talented artist, to use your home as their art studio while you are on vacation. Imagine further that, given their talent, the art that they will create will be finally valuable. However, if they didn't have this place to work, they would not have been able to create their masterworks of art. It's not clear that the workspace you provide to them is valuable in its own right. And, if it is, it has nothing to do with the artist and their work. Moreover, your house doesn't cause the artist to create masterpieces of art. Whether they painted was up to them. The value that your house seems to have is as a space that provided the artist with the opportunity to create things with final value. The same holds for any workspace that provides the opportunity for final value to be created, e.g., the workspaces of fiction and non-fiction writers, scientists, medical doctors, and so on. Thus, these spaces seem to be extrinsically valuable, i.e., valuable in virtue of some set of their relational properties.

Or, consider museums. Art museums provide visitors with the opportunity to experience beauty and wonder, which are plausibly finally valuable experiences. Natural history and science museums provide visitors with the opportunity to gain knowledge about important facts (e.g., human history, life on earth, and scientific theories). Notice that museums do not cause visitors to have this knowledge. After all, one could go to the museums and ignore everything around as a small child might. Nor it is clear that museums are finally valuable.

I think it is plausible that the value provided by the workspaces of talented artists and museums means that we should care about these spaces. For example, it seems like we should

---

[258] Zimmerman and Bradley (2019).

be disappointed or sad if a brilliant artist's workspace is destroyed and she is therefore prevented from creating finally valuable work for some period of time. Likewise, I think we should desire that art, history, and science museums should stay open so that people can gain the final value that is available at these places. So, it looks like it's plausible that we should care about things that provide the space or opportunity for intrinsically or finally valuable things to come about.

One reason we should care about animals is that they have the kind of facilitative value that artist studios and museums have. Any particular non-human animal is capable of having what utilitarians think is intrinsically (or finally) valuable, e.g., either experiences of pleasure or satisfied desires. The same holds for non-human animals in general (i.e., mammals over one year old). As a group, they provide many opportunities for intrinsic value to exist. Therefore, just as we should care about the workspaces of artists, scientists, and medical doctors, and museums, we should care about animals.

However, there is a way in which non-human animals are even more valuable than these workspaces and museums. This is because while there are many kinds of places that are capable of providing the opportunity for people to create good art or to find knowledge, there are relatively view kinds of creatures that can feel pleasure or have desires satisfied. Of course, humans can experience pleasure and have desires satisfied and therefore, even according to utilitarians, we should care about humans. But other than humans, only the non-human animals I'm talking about (mammals older than a year old) can experience pleasure and have desires satisfied.

This means that even if they are valuable only because of the opportunity to provide for intrinsic (or final) value to come into existence, the fact that they are one of only a few kinds of creatures that can provide the opportunity for this kind of value to come into existence should increase their value. The claim that the value of animals increases because they are

one of two kinds of beings that can provide opportunities for intrinsic (or final) value is supported by the common thesis that the rarity of an object can increase the value of it. For example, recall Dancy's case of a book increasing in value if all the other copies of it are destroyed.[259] The fact only one particular book provides the opportunity to gain some bit of knowledge makes it more valuable than if there were more copies of it. Likewise, if there were many kinds of beings that provided the opportunity for intrinsic (or final) value to come into existence, then the value of animals would decrease.

So, I think that even utilitarians should agree that animals are quite valuable because they are one of two kinds of beings that provide the opportunity for intrinsic (or final) value to come into existence. Given that they have this kind of opportunity value and this kind of opportunity value is rare, even utilitarians should think that we should care about animals.

One might wonder if facilitative value is really different from the instrumental value that utilitarians normally talk about. I think it is important to distinguish facilitative value from instrumental value for two reasons. First, there is a difference between A causing/ producing B and A providing the opportunity for B to be caused (e.g., by C). If facilitative value can be explained in purely causal terms, then this just means that there is more than one way to be causally related to something of value. Second, when utilitarians are talking about "instrumental" value, they often have in mind things that can directly *cause* or *produce* the best state of affairs to occur.[260] So, it is important to distinguish facilitative value from "instrumental" value even if they end up being two kinds of causal relation and so both are broadly instrumental.

Given that Utilitarianism should think that animals are very valuable and proponents of MRV think that animals are very valuable, they seem to exhibit total convergence on the

---

[259] Dancy (2003): 633.

[260] Fletcher (2013), Driver (2014), and Nathanson (n.d.).

answer to the question, "Should I care about non-human animals?" The answer they both give is: Yes. Moreover, because they both think that animals are very valuable, they should also exhibit total convergence on the question of roughly how much one should care about animals. Of course, as I noted above, to say that something is very valuable is vague and so total convergence on the answer to a vague question is not the most impressive kind of convergence. However, it is still a kind of convergence and a more important kind of convergence than simply agreeing that animals are valuable. After all, one view might think that animals are valuable, but only a little bit while the other view might think that they are the most valuable thing in the world. So, the fact that there is convergence that animals are very valuable is somewhat important.

What utilitarians and proponents of MRV don't exhibit any convergence about is *how* one should care about animals (i.e., either instrumentally or finally). Because utilitarians think that animals only have value as a kind of home or container for intrinsic value, it seems like they think that one should only instrumentally care about animals. However, proponents of MRV think that animals are finally valuable and so one should finally care about them.

One might wonder why utilitarians cannot make a move similar to the move made by Broad Anthropocentrists and claim that what has intrinsic value must be related to happiness, pleasure, or desire satisfaction, but it not be only be related to these things by producing them. Maybe something can be instrumentally valuable because it *previously* made someone happy, gave them pleasure, or satisfied a desire. And, perhaps we can claim that these things can come to have final value as just Lincoln's pen or nature came to have final value.

However, Utilitarianism makes a specific claim about what has intrinsic (or final) value that excludes any previous means of happiness, pleasure, or desire satisfaction from being finally valuable. In particular, Utilitarianism endorses value monism, i.e., they claim that only

one thing has intrinsic (or final) value.[261] For Hedonistic Utilitarians, only pleasure or happiness is intrinsically (or finally) valuable. For Preference Utilitarians, only desire satisfaction is intrinsically (or finally) valuable. So, these views already exclude anything else from being finally valuable. But notice that Broad Anthropocentrism only claims that the value of nature must be based on its relationship with human beings. It need not be specific which relationship. Moreover, utilitarians tend to think that what has intrinsic (final) value is tied to the mental states of human beings (e.g., pleasure or desire) and the actual individuals are not what matters.

One final worry is that there might be circumstances in which utilitarians will actually claim that we shouldn't care about animals. In particular, there might be situations in which it would maximise overall utility (e.g., happiness, pleasure, or desire satisfaction) if one didn't care about animals. For example, imagine a billionaire who hates animals will donate billions of dollars to charities around the world if one ceases to care about animals. Let us grant that her donating this money will maximise utility. However, the fact remains that animals are valuable as containers of what is intrinsically valuable. Shouldn't the utilitarian hold that one *shouldn't* care about animals in this case?

First, even if were true that one should cease caring about animals in this case, all that would follow is that there is at least mass convergence between Utilitarianism and MRV on the answer to the question, "Should I care about animals?" This would still be a far greater amount of convergence than we've seen before between these two views.

Second, I think that utilitarians can simultaneously maintain that you "should" care about nature even when it would not maximise utility. To see why recall that I argued earlier that by "should" I meant that it is fitting or appropriate or that one has a justifying reason. However, there is no conflict with a utilitarian maintaining that it would be fitting or

---

[261] Skorupski (2003): 209.

appropriate to care about nature and at the same time also maintaining is that it would be all things considered morally wrong to do so. What the utilitarian is claiming, in this case, is that one's reason to care about animals is just outweighed by a reason to not care about them.

Third, one might even argue that there is no conflict because one does not have direct control over what one cares about, but one can only be required to do something if one has direct control over doing it. So even if it would maximise utility to stop caring, one cannot be required to do so. Rather one might be required to do something else that might affect whether one cares, but that is not a reason itself to stop caring. So, even in these cases, Utilitarianism doesn't tell one that it is fitting to care and also not that one should not care, only that it is fitting to care and that one should do something else that might make one cease caring.

4.2 Convergence on Why to Care

In this section, I will argue that there is total *derivative* explanatory convergence between Utilitarianism and MRV on the answer to the question, "Why should I care about animals?" I will then explain where these two views do not exhibit any *fundamental* explanatory convergence.

As we saw above, Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animals?" Now the question of whether they exhibit any kind of derivative explanatory convergence depends upon whether they agree to any degree on any explanation of *why* we should care about animals. I think it is clear that they do agree about one explanation. In particular, they both agree that we should care about animals because they are valuable.

Perhaps one will object that the fact that X is simply valuable is not enough to make it so that one should care about X. Perhaps everything has *some* degree of value, but it certainly wouldn't follow that one should care about everything (e.g., rock, leaves, dirt, bottlecaps,

loose thread, and so on). Fortunately, both Utilitarianism and MRV can hold that non-human animals are very valuable. Of course, as we noted above, this is vague. But it at least allows us to separate non-human animals from things like rocks and lint. Thus, utilitarians and proponents of MRV exhibit total convergence on the question, "Why should I care about animals?" The answer they will both give is that animals are "very valuable".

This also means that they might exhibit mass or total convergence on the answer to the question, "Should I care a lot about nature?" However, because figuring out precisely how valuable each theory thinks animals are is beyond the scope of this chapter, I can only offer a hedged answer to this question.

However, there is no fundamental explanatory convergence between them. This is because they disagree about what fundamentally explains why animals are valuable and therefore why we should care about them. Recall that Utilitarianism thinks that animals are valuable because provide the opportunity for intrinsically valuable things to come into existence. However, MRV holds animals are valuable in their own right in virtue of being subjects of a life. Therefore, Utilitarianism and MRV do not exhibit any fundamental explanatory convergence.

5. An Objection

One might worry that the fact that nature is very, finally valuable and that the fact that animals are very valuable are not *themselves* the right kind of reason for caring about them. For example, if you asked me why I care about nature and I say, "Because it's (very) valuable" or "Because it's finally valuable," you might be confused. This might sound like an overly intellectual or cold reason to care about nature or animals. You would probably expect me to explain my caring about nature by saying things like, "It's beautiful," or "It gives me a sense of peace" or "I used to go camping here as a child."[262]

---

[262] I thank Michael Brady for this objection.

I agree that claiming that you care about something because of its value or final value sounds odd. However, I don't think this is a problem for my argument in this chapter. To see why we need to distinguish a motivating reason from a normative or justifying reason. A motivating reason is a reason for which someone performs an action or forms some attitude. In the case of caring, one's motivating reason is the reason that they care about X. A normative or justifying reason to perform some action is a fact that makes it the case that one is *required* or *justified* in performing that action or forming that attitude. These two kinds of reasons can come apart. For example, someone might drink a liquid because they think it's poison. That is, their motivating reason for drinking the liquid is "it is poison." But that cannot be a normative reason to drink the liquid, because the fact that the liquid is poison does not justify their drinking the liquid at all. In fact, it justifies them in *not* drinking the liquid.

What I have argued in this chapter is that: (1) Broad Anthropocentrists and Non-Anthropocentrists agree that one has a *normative/justifying* reason to care about nature and they agree that one has a normative reason to care about nature because they agree that nature is very, finally valuable; (2) Utilitarians and proponents of MRV agree that one has a normative/justifying reason to care about animals and they agree that one has a normative reason to care about animals because they agree that animals are very valuable. So, I was not suggesting that people have the motivating reason "it's (very) valuable)" or "it's finally valuable"

I also think that the oddness of my response that I value some part of nature "Because it's finally valuable" can be explained away. Notice that it's not as if it sounds like I'm making a *false* claim or lying. So, my answer is not odd because it's false. Rather, one way of understanding why my answer seems odd is because is it not informative enough. In particular, it does not answer the question of why I care about *this* part of nature as opposed

to *that* part of nature. Nor does it help distinguish why I care about this part of nature as opposed to why I care about my pet. I likely find both this part of nature and my pet to be finally valuable. But that doesn't tell my interlocutor what I find *particularly* or *especially* valuable about this part of nature.

This explanation is supported by the observation that if it would be natural for my interlocutor to respond to my answer by asking, "But why do you find it valuable?" Notice that the interlocutor is not doubting the veracity of my response, but rather seeking more specificity. But the fact that it would be uninformative to tell someone that I care about something because it is finally valuable, does not make it false that I care about it for that reason. So, it might be that my reason for finally caring is that I find that part of nature to be finally valuable.

Finally, one might find my response that I value some part of nature because it's finally valuable to be odd because my response is not subjective enough. That is, I have explained a psychological fact, i.e., that I finally care about some part of nature, by appealing to a mind-independent fact, i.e., that it is finally valuable. I can avoid this by rephrasing my response to appeal to my own psychological states concerning this part of nature. So, for example, if asked why I care about some part of nature, I can answer: "Because it seems valuable to me" or "Because I think it's quite valuable" or "Because I know it's valuable." This sounds less odd to me and the reason might be that I'm indicating why *I* care about this part of nature as opposed to why one should care about nature.

6. Conclusion

In this chapter, I argued for two claims about the convergence of normative ethical theories in environmental and animal ethics. First, I argued that Broad Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I finally care about nature?" In particular, I argued that they both entail that one should finally care

about nature. I argued that they exhibit total derivative explanatory convergence, i.e., they converge on the answer to the question of why one should finally care about nature. I also argued that they might exhibit mass or total convergence on the answer to the question, "Should I finally care a lot about nature?"

Second, I argued that Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animals?" I argued that they both entail that one should care about animals. I argued that they exhibit total derivative explanatory convergence, i.e., they convergence on the one answer to the question of why one should care about animals. In particular, they agree that animals are very valuable. I indicated that they might also exhibit mass or total convergence on the answer to the question, "Should I care a lot about animals?" However, they don't exhibit any convergence concerning *how* to care about animals (instrumentally or finally) and they exhibit no amount of fundamental explanatory convergence.

Conclusion

In this thesis, I have explored the limits of convergence between prominent normative ethical theories in environmental ethics and animal ethics. In this concluding chapter, I will outline the major conclusions of this investigation.

Chapter 1

In Chapter 1, I clarified what normative ethical theories are and explained the core features of the most prominent normative ethical theories. I then explained the kinds of convergence that they can exhibit. I argued that they could exhibit practical, non-fundamental, and fundamental convergence. Moreover, I argued that all of these kinds of convergence come in degrees: total, mass, non-trivial, and trivial. I argued that all the plausible strategies for arguing for total or mass convergence at the practical and explanatory level have important limitations. For example, I argued that Parfit's attempt to show that there is fundamental explanatory convergence between Kantian Contractualism, Scanlonian Contractualism, and Rule Consequentialism fails. First, it requires making the fundamental right-/wrong-making property of Rule Consequentialism a *non*-fundamental right-/wrong-making property. Second, it requires making all three fundamental explanatory properties *conditional* in the sense that they can only make an action right or wrong *on the condition* that the other properties are present in the action as well. But normative ethical theories are usually meant to posit unconditional, fundamental explanatory properties. I concluded by arguing that the most plausible kind of convergence to argue for, when it comes to total or mass convergence, is non-fundamental explanatory convergence. This is because normative ethical theories are supposed to not only entail that certain kinds of actions are right or wrong (i.e., that certain kinds of properties make actions right or wrong) but also explain why these properties tend to make actions right or wrong.

Chapter 2

In Chapter 2, I distinguished between the kinds of value that played crucial roles in the debates between Anthropocentrism/Non-Anthropocentrism and the Utilitarianism/MRV and show how they are related (i.e., objective value, subjective value, final value, instrumental value, extrinsic value, intrinsic value, and inherent value). I argued that once we have a clear picture of these kinds of values and how they are related, we can see that while much of the literature claims to concern the importance of intrinsic value, what philosophers really care about is objective *final* value. Recognising this fact matters not only as a means of clarifying the debates but also because it shows that the convergence between Broad Anthropocentrism and Non-Anthropocentrism on the final value of nature is an important kind of convergence.

Chapter 3

In Chapter 3, I argued that Norton failed to show that there is mass or total convergence between what he called Weak Anthropocentrism and Non-Anthropocentrism. Weak Anthropocentrism involves the claim that humans who are rational and fully-informed will prefer that nature be treated well in just the same way that Non-Anthropocentrists would prefer it to be treated, i.e., as if it had intrinsic value. I argued that Norton's argument failed for two reasons. First, I showed that rational and fully-informed people could rationally either fail to prefer that nature be treated well or that prefer that nature not be treated well. Second, I argued that Weak Anthropocentrism doesn't exhibit any explanatory convergence with Non-Anthropocentrism and so even if his claim about practical convergence were correct, it would not satisfy many Anthropocentrists and Non-Anthropocentrists. I then argued that we should interpret Anthropocentrism more broadly as being the claim that the value of nature must come from its relationship with humans, but this relationship need not be merely instrumental. Instead, I argued that nature has and continues to play a uniquely important role in the sustaining of human life. I argued that this role makes nature finally valuable. This

means that Broad Anthropocentrism and Non-Anthropocentrism exhibit total non-fundamental explanatory convergence because they agree that nature is finally valuable. After all, Non-Anthropocentrism holds that nature is intrinsically valuable and this entails being finally valuable. I also argued that they will agree that nature is *very* valuable. Given this convergence, it is plausible that Broad Anthropocentrism and Non-Anthropocentrism exhibit a non-trivial amount of practical convergence as well.

Chapter 4

In Chapter 4, I argued that Gary Varner fails to show that there can be a non-trivial amount of practical convergence between Utilitarianism and the Moral Rights View (MRV) on the issue of animal experimentation. Varner argues that by looking at Preference Utilitarianism and Regan's worse-off principle, we can see that there is a non-trivial amount of convergence between Utilitarianism and MRV. I argued that even appealing to the above these forms of Utilitarianism and MRV are not sufficient for showing that Utilitarianism and MRV exhibit *any* non-trivial convergence on the issue of animal experimentation. First, I argued that proponents of MRV will hold principles in addition to the worse-off principle that will prohibit animal experimentation. Second, I argued that even if Varner is correct that there is some convergence between Preference Utilitarianism and the worse-off principle, it is not the kind of convergence that would satisfy proponents of either view. This is because proponents of each view really care about whether an action instantiates their preferred right- or wrong-making properties. In fact, even those who only care about animal experimentation policy will not be satisfied with the kind of practical convergence Varner argues for. This is because such policies are meant to be long-term and stable. But it doesn't follow from the fact that Utilitarianism and the worse-off principle agree with a policy *before* it's enacted that they will agree with that policy after it's been enacted. This is because certain contingent facts could change that entail that experiments *become* wrong according to one view, but not

another. For example, if the conducting of an experiment causes many people to feel unhappy or to have their desires frustrated, then that experiment would cease being permissible for Utilitarians, but not proponents of the worse-off principle.

Chapter 5

In Chapter 5, I argued that there is a fair amount of convergence on the question of whether and why to care about nature and animals between Broad Anthropocentrism and Non-Anthropocentrism and Utilitarianism and MRV, respectively. First, I argued that Broad Anthropocentrism and Non-Anthropocentrism exhibit total convergence on the answer to the question, "Should I finally care about nature?" I argued that they exhibit total derivative explanatory convergence, i.e., they exhibit total convergence on one answer to the question of why one should finally care about nature. This is because they agree that nature is finally valuable. I also argued that they might exhibit mass or total convergence on the answer to the question, "Should I finally care a lot about nature?" because they exhibit total convergence on the question of how valuable nature is. I argued that there is reason to think that they will both hold that nature is "very" valuable. However, because it is beyond the scope of this thesis to offer an account of how to precisely measure value, I concluded that "very" indicates a vague range of value. Moreover, I concluded that they don't exhibit any amount of explanatory convergence concerning why to finally care about nature.

Second, I argued that Utilitarianism and MRV exhibit total convergence on the answer to the question, "Should I care about animals?" Moreover, I argued that they exhibit total derivative explanatory convergence on one answer to the question of why one should care about animals. In particular, they agree that animals are very valuable. I indicated that they might also exhibit mass or total convergence on the answer to the question, "Should I care a lot about animals?" But, as with the debate concerning Anthropocentrism and Non-Anthropocentrism, it is beyond the scope of this thesis to offer an account of how to precisely

measure value. Therefore, I concluded merely that Utilitarianism and MRV would exhibit mass or total convergence on the vague claim that animals are "very" valuable. Finally, I argued that these two views don't exhibit any convergence concerning *how* to care about animals (i.e., either instrumentally or finally) and they exhibit no amount of fundamental explanatory convergence.

Final Thoughts

I argued that we should often be sceptical about the prospects of practical and fundamental explanatory convergence when it comes to actions. For example, I argued that Varner fails to show that Preference Utilitarianism and MRV exhibit even non-trivial practical or explanatory convergence concerning the morality of animal experimentation. I also argued that Parfit (2011) fails to show that there is explanatory convergence between Scanlonian Contractualism, Kantian Contractualism, and Rule Consequentialism. So, I don't think Parfit's or Varner's arguments should assuage concerns about the threat of peer disagreement to moral truth or moral knowledge. Although, there are certainly alternative arguments against such sceptical views.

I also argued that there is practical and non-fundamental explanatory convergence between Broad Anthropocentrism and Non-Anthropocentrism. This means that Anthropocentrism, once properly understood, has much more in common with Non-Anthropocentrism than many thought. I extended this thought in Chapter 5 when I argued that there is also total convergence between these two views on whether and why to finally care about nature. I think that these arguments should assuage, to some extent, sceptical concerns about the moral importance of nature or of moral knowledge concerning nature. Moreover, I think that this much convergence should help environmental policymakers see that the two most prominent views on the moral importance of nature agree that nature is very, finally valuable and should therefore be protected.

This conclusion is practically important because, as I noted in the Introduction, climate change poses an imminent threat to both living creatures and to the non-living parts of nature (e.g., ecosystems). Moreover, in order to minimise the extent of the damage that climate change will lead to, policymakers need to agree on and try to enact policies to protect nature as soon as possible. My arguments in Chapters 3 show that there is a good amount of convergence between the most prominent views in environmental ethics. More precisely, they agree that nature is finally valuable and therefore worth protecting.

Moreover, my development of Broad Anthropocentrism also has the potential to help in the defence of environmentally-friend policies. At least some environmental ethicists are generally sceptical of the claim that nature has intrinsic value or they think that talk of intrinsic value is unnecessary for supporting environmentally-friendly policies.[263] Broad Anthropocentrism provides a way of grounding the final value of nature solely in terms of its relationship with human beings. This allows one to defend environmentally-friendly policies solely on anthropocentric grounds.

Finally, it is more important than ever to consider the morality of animal experimentation. As the world tries to develop new vaccines to combat new variants of the coronavirus, more and more non-human animals will be used to test these vaccines.[264] One the first step in vaccine development is testing the vaccine on non-human animals. The current vaccines for the coronavirus were developed by experimenting on mammals such as mice, rats, pigs, ferrets, hamsters, rabbits and non-human primates.[265] Given the urgency of developing new vaccines, it is essential that we have a clear view of how much and what kind of animal experimentations are morally permissible. If Varner had been correct that the

---

[263] Light (2002), Morito (2003), Norton (1984, 1991, and 1995), Weston (1996).

[264] NIAID (2021).

[265] EFPIA (2021).

worse-off principle is a plausible and complete account of MRV, then he would have been right that even deontologists think that animal experiments that kill or harm animals are morally permissible as long as they save a single human life. He would also have been able to argue that MRV and Utilitarianism exhibit at least a non-trivial amount of practical convergence. However, as I argued in Chapter 4, the worse-off principle is not a plausible or complete account of MRV and proponents of MRV would certainly appeal to additional principles to argue that harmful or lethal animal experimentation is often morally impermissible. Thus, while I didn't offer a view on exactly which and how much animal experimentation is morally permissible, my arguments indicate that there isn't much agreement between Utilitarians and proponents of MRV.

Appendix: Treating Nature Well Without Convergence

In this thesis, I have argued that one version of Anthropocentrism (i.e., Broad Anthropocentrism) and Non-Anthropocentrism exhibit a good amount of convergence on the question of whether and why to treat nature well and whether and why to finally care about nature. One motivation for exploring these questions was to show that environmental policymakers could appeal to this convergence between seemingly opposing views to promote policies that protect nature. Recently, however, some environmental scientists and philosophers (especially Chan et al. and Knippenberg et al.) have argued that appealing to the value of nature will not help promote policies that protect nature. They argue that few people make decisions based solely on the kind of value something has (e.g., inherent, intrinsic, or instrumental).[266] Rather, they often base their decisions on what they think is appropriate to do based on the kinds of relationships they have.[267] That is, what is required or appropriate given the norms of their particular relationships. For example, people make decisions based on what they think a good friend or a virtuous spouse would do. They call the norms and virtues associated with certain relationships "relational values." They argue that appealing to norms of relationships is more conducive than appealing to instrumental or intrinsic value in promoting policies that protect nature.[268] In this Appendix, I will critically examine these recent views. I argue that different authors mean different things by relational value but that none of them offer a plausible source of moral reasons to treat nature well. Moreover, I argue that their views are really just disguised versions of Anthropocentrism concerned with the instrumental value of nature.

---

[266] Chan et al. (2016): 1462.

[267] Chan et al. (2016): 1463.

[268] Chan et al. (2016): 1462.

Chan et al. and Knippenberg et al. argue that humans already have a certain kind of valuable relationship with nature that is similar to the interpersonal relationships that we have with our friends and family. Chan et al.'s view has been widely endorsed by environment scientists and scholars working on environmental policy.[269]

Both Chan et al. and Knippenberg et al. argue that this relationship with nature is valuable and therefore that we have reasons to treat nature well because of the value present in this relationship, which they call "eudaimonic value" or value associated with a living a good life.[270] For example, Knippenberg et al write:

> [R]elational values are 'values relative to the meaningfulness of relationships'. It appears we have a freedom for the term to mean either or both of first, the intrinsic value of relational triads or second, the constitutive value of the relationship per se.[271]

On this view, the moral reason we have to treat nature well is based on the value that is either intrinsic or constitutive of the relationship with have with nature.

Chan et al. make a similar suggestion by writing that one source of moral reasons to treat nature well is "relational values," which they define as:

> preferences, principles, and virtues associated with relationships, both interpersonal and as articulated by policies and social norms.[272]

Unfortunately, this definition is rather vague and unclear and Chan et al. provide no other explanation of relational value in general.

However, they do say more about interpersonal relationships and the preferences, principles, and virtues that are associated with them. They write, "Relational values are not present in things but derivative of relationships and responsibilities to them [i.e., those

---

[269] Schulz and Martin-Ortega (2018), Muradian and Pascual (2018), Klain et al. (2017). Grubert (2018), and Stenseke (2018).

[270] Chan et al. (2016): 1462.

[271] Knippenberg et al. (2018): 41.

[272] Chan et al. (2016): 1462.

relationships]."[273] As I interpret Chan et al, they are using term "values" in the sense of "ideals" or "norms" that are specific to a certain domain. For example, there are certain norms or ideals specific to certain workplaces, e.g., arrive on time, keep a tidy desk, do not make phone calls using company phones, etc. Likewise, there are values, norms, or ideals of good sportsmanship, e.g., one should not "show off" or cheat, one should be polite and cheer for one's teammates. Likewise, there are certain values, ideals or norms of interpersonal relationships. For example, friends ought to trust each other, help each other out with expecting anything in return, care about each other, and so on, and spouses are supposed to be supportive, understanding, and honest. On this view, humans have a reason to treat nature well because they have a certain kind of valuable relationship with nature and relationships of that kind come with certain values, ideals and norms that are specific to that kind of relationship.

Another possible interpretation of Chan et al. is that by "relational values" they are talking about what people call "special obligations." Special obligations are moral obligations that we have only to people with whom we have certain intimate, personal relationships, e.g., friends and family and perhaps colleagues or fellow citizens.[274] For example, you might have a special moral obligation to help your brother move if he asks, but you do not have a moral obligation to help a stranger move if they ask. Likewise, you might have an obligation to pay for your child's education, but the not education of a stranger's child.

There is a deep problem with Chan et al.'s and Knippenberg et al.'s suggestion that we have reasons to treat nature well that are based on our relationship with nature. The problem is that this view is only plausible if the relationship we have with nature is relevantly *similar* to the relationship we have with other people, e.g., friends and family. However, there are

---

[273] Chan et al. (2016): 1462.

[274] Jeske (2019).

crucial ways in which our relationship with nature, whatever it might be, must be different than our relationships with our friends and family.

First, interpersonal relationships are most often reciprocal, i.e., not only do we care about our friends and voluntarily and intentionally help them out, but our friends care about us and voluntarily and intentionally help us out. Moreover, we help our friends and family *for their own sake* and not because of some value we will get because we help them. That is, we treat them and their interests as being the source of our reasons to help and not some benefit we will get by helping them. This is because we finally care about them in the sense discussed in Chapter 5. In fact, one might think that the source of the special value of such relationships is the voluntary and intentional actions are both performed for the loved one's own sake and these actions are reciprocated. But nature is not capable of voluntarily and intentionally helping us for our own sake or voluntarily and intentionally helping us at all. So, there is also no reciprocity of this kind in any relationship with nature.

Second, even when those relationships are not fully reciprocal (e.g., the relationship between a mother and a newborn baby), the person who cannot reciprocate in action will often reciprocate in love or care. For example, while an infant cannot do their parents a favour, they do respond to what their parents do for them by loving or caring about the parents. In these cases, because the parent loves their child and the child loves them back, there is at least reciprocal love. Moreover, the parent and child love each other for the sake of their loved ones. That is, parents love their children for their children's own sake and the children love their parents for the parents' own sake. In other words, they love each other *finally*. And, the special value of these kinds of relationships seems to be at least partly based on the reciprocity of this non-instrumental love.[275] But nature is not capable of caring about or loving humans at all (never mind of loving them non-instrumentally).

---

[275] Also, children will, most often, become capable of reciprocating in action once they grow up.

Third, interpersonal relationships only involve two people and this creates a sense of intimacy between them. But when it comes to whatever relationship I have with nature, many, if not most, people have exactly the same relationship with nature. So, there is nothing special about my relationship with nature. And therefore, the value that relationships can have in virtue of their intimacy or rarity cannot be had in whatever relationship I (and most other people) have with nature.

Because the reasons that certain interpersonal relationships are valuable don't apply to human-nature relationships, I think Chan et al. and Knippenberg et al. are wrong to argue that we have the kind of relationship with nature that gives us a reason to treat nature well.

Chan et al. also argue that people care more about living a good life than the instrumental or intrinsic value of things when they're deliberating about what to do. They write:

> Few people make personal choices based only on how things possess inherent worth or satisfy their preferences (intrinsic and instrumental values, respectively). People also consider the appropriateness of how they relate with nature and with others, including the actions and habits conducive to a good life, both meaningful and satisfying.[276]

Chan et al. seem to be talking about what reasons actually motivate people to act, but we can reinterpret what they say as being about what moral reasons people have to nature well. More precisely, it seems like they are arguing that we should treat nature well because it is "conducive" to living a good life. That is, we have a moral reason to treat nature well because doing this tends to lead to having a good life or because it makes a good life more likely.

This interpretation of Chan et al. is strengthened by what they call statements of "rational values",

- Being in nature provides a vehicle to connect with people (Social cohesion).

---

[276] Chan et al. (2016): 1462.

- Care for some piece of land fulfils people, helps them lead a good life (Stewardship eudaemonic).[277]
- Some people's identities are rooted in long-term care and stewardship, such as volunteer stream-keepers and urban or rural farmers.
- Some people and social organisations hold worldviews that encompass kinship between people and nature.[278]

I do not think that these reasons are reasons that everyone has to treat nature well. To see why not, let us look at some of these reasons more closely.

One reason listed by Chan et al is that taking care of some land "fulfils" people. If we understand this as the claim that taking care of some piece of land makes them feel fulfilled or makes them have some sense of satisfaction, then this cannot be a reason for everyone to take care of nature. Not everyone will have that feeling of fulfilment and so this cannot be a reason for everyone to take care of nature. Likewise, another reason they list to treat nature well is that being in nature provides a vehicle for people to connect with people. But being in nature will not be a vehicle for everyone to connect with other people. There are plenty of other vehicles for connecting with people. The limits of the reasons they provide can be seen more clearly in the last two reasons they give. Both of these reasons explicitly note that they apply only to "some people." This means that all of these reasons are reasons that only some people and it is possible that at some point no one will have these reasons because they will not be affected by nature in the ways Chan et al. list.

Also, some of the good things that nature provides to people are not conducive to living a good life. For example, Chan et al. imply that connecting with people and feeling fulfilled are important for living a good life. But that is not true. What if the people you connect with are bad people (e.g., Nazis)? Is connecting with Nazis conducive to living a good life? I do not think so. Likewise, whether feeling fulfilled is conducive to living a good

[277] Chan et al. (2016): 1462.

[278] Chan et al. (2016): 1463.

life depends on why one feels fulfilled. For example, if one feels fulfilled by taking care of nature because one likes controlling or having dominion over things, then taking care of nature does not seem conducive to living a good life.

Therefore, I think that Chan et al. and Knippenberg et al. fail to provide a reason to treat nature well that is different from Anthropocentrists and Non-Anthropocentrists. This is because the reasons they offer do not apply to everyone or they assume that people have a certain relationship with nature that we cannot have.

Finally, it looks like their view is just a "disguised" version of Anthropocentrism. After all, what seems to explain the importance of our personal relationships is that they are part of living a good life. For example, Chan et al. write:

> Few people make personal choices based only on how things possess inherent worth or satisfy their preferences (intrinsic and instrumental values, respectively). People also consider the appropriateness of how they relate with nature and with others, including the actions and habits conducive to a good life, both meaningful and satisfying.[279]

> Whereas instrumentalism considers value as derived from the satisfaction of preferences whatever they are, the relational notion of eudaimonia ("flourishing") entails reflection on the appropriateness of preferences, emphasizing that value is derived from a thing's or act's contribution to a good life, including adhering to one's moral principles and maintaining the roots of collective flourishing.[280]

This sounds as if the reason to have a relationship with nature or to follow the norms of such relationships is that doing so contributes to one flourishing or living a good life.

Likewise, Knippenberg et al. argue that we should have a relationship because it is instrumental to our flourishing or living a good life. They write:

> First, on a theoretical note, we may posit nature — wild nature especially — as our most fundamental Other, and therewith, following Aristotle's principle that the other teaches us who we are [54], as our deepest source of identity and context for meaning [55]. Second, the experience of relationship with nature has capacities for growth, health and healing that no other relationship can provide, as shown in countless studies on short medical interventions [57] and long wilderness experience [58].

---

[279] Chan et al. (2016): 1462.

[280] Chan et al. (2016): 1464.

> Third, virtually all committed actors for nature in the already cited BIOMOT research shared lively stories of how encounters with nature co-constituted their lives.[281]

In other words, we should have a certain kind relationship with nature because: (1) it will help teach us who we are, (2) it has the ability to help us grow, become healthier, and heal, and (3) encounters with nature are often listed as an important part of the lives of people who committed to protecting nature.

What we can see in both Chan et al. and Knippenberg et al. is that being in a certain kind of relationship with nature is valuable because is instrumental to our living a good life or flourishing. But this just sounds like Anthropocentrism. Of course, Norton (1984) interpreted Anthropocentrism as claiming that nature is valuable only as a means for bringing about our unreflective and consumeristic desires or preferences. However, as we saw in Chapter 3, the core value of Anthropocentrism is just that nature's value is derived from its having *some kind* of relationship or relationships with humans. Nothing about this core value claim entails that nature's value is dependent on how much it can help with getting what we desire or prefer before reflection. It is thus perfectly consistent with Anthropocentrism to claim that nature's value is partly derived from its ability to help us live well.

If what I've argued is true, then Chan et al. and Knippenberg et al. are both just appealing to the instrumental value of nature in order to argue that we should treat nature well. However, they talk as if they are appealing to a "third class" of value, i.e., relational value.[282] So, not only do Chan et al. and Knippenberg et al. fail to show that our supposed relationship with nature gives us reasons to treat nature well, they also fail to show that there is a justification for treating nature well that doesn't appeal to intrinsic or instrumental value.

---

[281] Knippenberg (2018): 43.

[282] Chan et al. (2016): 4062.

References

ABPI. 2007. Facts & Statistics from the Pharmaceutical Industry. London, UK: Association of the British Pharmaceutical Industry. http://www.abpi.org.uk/statistics/ section.asp?sect=3

Adams, Robert Merrihew. 1976. Motive Utilitarianism. *Journal of Philosophy* 73(14): 467-481.

AFMA. 2004. New Survey Among Doctors Suggests Shift in Attitude Regarding Scientific Worth of Animal Testing. Listed as EFMA Survey of 500 General Practitioners. Los Angeles, CA, USA: Americans For Medical Advancement. www.curedisease.com

Agar, Nicholas. 2001. *Life's Intrinsic Value: Science, Ethics and Nature*. New York, NY: Columbia University Press.

Ashford, Elizabeth and Tim Mulgan. Contractualism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). <https://plato.stanford.edu/archives/sum2018/entries/contractualism/>.

Animal Procedures Committee. 2003. Review of Cost–Benefit Assessment in the Use of Animals in Research. London, UK: Animal Procedures Committee. www.apc.gov.uk

Anscombe, G.E.M. 2008. Twenty Opinions Common among Modern Anglo-American Philosophers. In Luke Gormally & Mary Geach (Eds.), *Faith in a Hard Ground: Essays on Religion, Philosophy and Ethics*, 66-68. Charlottesville, VA.: Imprint Academic.

Aquinas, Thomas. *Summa Theologica* II-II, Question 110, Article, 3

Baron, Marcia W., Philip Pettit, and Michael Slote. 1997. *Three Methods of Ethics*. Oxford: Blackwell Publishing.

Baumann, Marius. 2021. In Search of the Trinity: A Dilemma for Parfit's Conciliatory Project. *Ethical Theory and Moral Practice*: 1-19.

Beardsley Monroe C. 1965. Intrinsic Value. *Philosophy and Phenomenological Research* 26(1): 1-17.

Boyle, Joseph. 1991. Who is Entitled to Double Effect? *The journal of medicine and philosophy* 16(5): 475-494.

Bradley, Ben, 1998, Extrinsic Value, *Philosophical Studies* 91(2): 109-126.

Brady, Michael. 2015. Ethics. In Duncan Pritchard (ed.), *What is This Thing Called Philosophy?*, 3-32. London: Routledge.

Brennan, Andrew, 1984. “The Moral Standing of Natural Objects”, *Environmental Ethics* 6(1): 35-56.

Brody, Baruch. 2003. Defending Animal Research: An International Perspective. In Ellen Frankel Paul and Jeffrey Paul (eds.), *Why Animal Experimentation Matters: The Use of Animals in Medical Research*, 131-147. New Brunswick, NJ: Transaction Publishers and the Social Philosophy and Policy Foundation.

Callicott, J. Baird. 1984. Non-Anthropocentric Value Theory and Environmental Ethics. *American Philosophical Quarterly* 21(4): 299-309.

Callicott, J. Baird. 1989. *Defense of the Land Ethic*. Albany, NY: State University of New York Press.

Callicott, J. Baird. 2009 . The Convergence Hypothesis Falsified: Implicit Intrinsic Value, Operational Rights, and De Facto Standing in the Endangered Species Act. In Ben A. Minteer (ed.), *Nature in Common? Environmental Ethics and the Contested Foundations of Environmental Policy*. Temple University Press, 142-166.

Chan et al. 2016. Why protect nature? Rethinking values and the environment. *Proceedings of the National Academy of Sciences* 113(6): 1462-1465.

Christiano, Tom. Authority. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2020 Edition). <https://plato.stanford.edu/archives/sum2020/entries/authority/>.

Clune, Alan C. 1996. Biomedical Testing on Nonhuman Animals: An Attempt at a "Rapprochement" Between Utilitarianism and Theories of Inherent Value. *The Monist* 79(2): 230-246 .

Cohen, Carl. 1986. The case for Biomedical Experimentation. *New England Journal of Medicine* 315: 865-870.

Cudd, Ann and Seena Eftekhari. Contractarianism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). <https://plato.stanford.edu/archives/sum2018/entries/contractarianism/>.

Dancy, Jonathan. 2003. Are There Organic Unities? *Ethics* 113 (3) : 629-50.

Dancy, Jonathan. 2004. *Ethics without Principles*. Oxford: Claredon Press.

DeGrazia David. 1991. The moral status of animals and their use in research: a philosophical review. *Kennedy Institute of Ethics Journal* 1(1):48-70.

DeVries, R.B.M. 2008. Intrinsic Value and the Genetic Engineering of Animals. *Environmental Values* 17(3): 375-392.

Dickens, Charles. 2003. *A Christmas Carol*. Peterborough, Ontario: Broadview Press.

Driver, Julia. 2005. Normative Ethics. In Frank Jackson and Michael Smith (eds.), *The Oxford Handbook of Contemporary Philosophy*, 31-62. Oxford: Oxford University Press.

Driver, Julia. The History of Utilitarianism. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2014 Edition). <https://plato.stanford.edu/archives/win2014/entries/Utilitarianism-history/>.

Edwards, Rem B. 1993. Tom Regan's Seafaring Dog and (Un)Equal Inherent Worth. *Between the Species* 9(4): 231-235.

Elliot, Robert. 1992. Intrinsic Value, Environmental Obligation and Naturalness. *The Monist* 75(2): 138-160.

Elliot, Robert, 1997. *Faking Nature: The Ethics of Environmental Restoration*. London: Routledge.

European Federation of Pharmaceutical Industries and Associations (EFPIA). 2021. Fitting vaccine research into one year – Were animals used? 27/10/2021. <https://www.efpia.eu/news-events/the-efpia-view/blog-articles/fitting-vaccine-research-into-one-year-were-animals-used/#_ftn4>

Feinberg Joel. 1974. The Rights of Animals and Unborn Generations. In William T. Blackstone (ed.), *Philosophy & Environmental Crisis*, 43-68. Athens, GA: The University of Georgia Press.

Finnis, John. 1991. *Moral Absolutes: Tradition, Revision, and Truth*. Washington, D.C.: Catholic University of America Press.

Fletcher, Guy. 2013. Act Utilitarianism. In James E. Crimmins (ed.), *The Bloombury Encyclopedia of Utilitarianism*, 1-5. London: Bloomsbury Academic.

Forde, Kathy Roberts and Bryan Bowman. 2017. Exploiting black labor after the abolition of slavery. *The Conversation*, February 7, 2017. Accessed September 20, 2021: < https://theconversation.com/exploiting-black-labor-after-the-abolition-of-slavery-72482>.

Francione, Gary L. 2003. Animal Rights Theory and Utilitarianism: Relative Normative Guidance. *Between The Species* 13(3): 1-30.

Francione, Gary L. 2008. *Animals as Persons: Essays on the Abolition of Animal Exploitation*. New York: Columbia University Press.

Frey, R. G. 2002. Moral standing, the value of lives, and speciesism. In: Hugh LaFollette (ed.), *Ethics in practice: an anthology*, 117-127. Oxford: Blackwell Publishing.

Frey, R. G. 2003. Animals. In Hugh LaFollette (ed.), *The Oxford Handbook of Practical Ethics*, 161-187. Oxford: Oxford University Press.

Frey, R. G. 2005. Animals and Their Medical Use. In Andrew I. Cohen and Christopher Heath Wellman (eds.), *Contemporary Debates in Applied Ethics*, 91-103. Malden, MA: Blackwell Publishers.

Gert, Joshua. 2007. Normative Strength and the Balance of Reasons. *Philosophical Review* 116: 533-562.

Gewirth, Alan. 1981. Are there any Absolute Rights? *Philosophical Quarterly* 31 (122): 1-16.

Gregorich, E. G., L.W. Turchenek, M.R. Carter, and D.A. Angers (eds.). 2001. *Soil and Environmental Science Dictionary*. Boca Raton, FL.: CRC Press.

Grenholm, Carl-Henric. 1997. Etik och djurförsök ["Ethics and Animal Experimentation"]. *Årsbok 1997 för Föreningen lärare i Religionskunskap*: 81-99.

Grubert, Emily. 2018. Relational values in environmental assessment: the social context of environmental impact. *Current Opinion in Environmental Sustainability*35: 100-107.

Hare, R.M. 1981. *Moral Thinking: Its Level, Method and Point*. Oxford: Clarendon Press.

Hargrove, Eugene C. 1992. Weak Anthropocentric Intrinsic Value. *The Monist* 75(2): 183-207.

Harris, Amy. 2018. 10 Examples of a Natural Ecosystem. *Sciencing.com*. Retrieved from https://sciencing.com/10-examples-natural-ecosystem-7836.html

Harsanyi, John C. 1996. Utilities, preferences, and substantive goods. *Social Choice and Welfare* 14(1): 129-145.

Hobbes, Thomas. 1668/1994. Leviathan. Edited, with introduction, by Edwin Curley. Indianapolis, IN: Hackett Publishing Company, Inc.

Howard, Christopher. 2018. Fittingness. *Philosophy Compass* 13:e12542.

Hurka, Thomas. 2001. *Virtue, Vice, and Value*. Oxford: Oxford University Press.

Hursthouse, Rosalind. 1999. *On Virtue Ethics*. Oxford: Oxford University Press.

Jaworska, Agnieszka. 2007. Caring and Internality. Philosophy and Phenomenological Research 74(3): 529-568.

Jeske, Diane. Special Obligations. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition. <https://plato.stanford.edu/archives/fall2019/entries/special-obligations/>.

Jørgensen, Sven E., Bernard C. Patten, and Milan Straškraba. 1992. Ecosystems emerging: toward an ecology of complex systems in a complex future. *Ecological Modelling* 62(1-3): 1-27.

Kagan, Shelly. 1998. Rethinking intrinsic value. *The Journal of Ethics* 2(4): 277-97.

Kant, Immanuel. 1996. On a supposed right to lie from philanthropy. In Mary J. Gregor (trans. and ed.), *Practical Philosophy*, 607-615. Cambridge: Cambridge University Press.

Kant, Immanuel. 1785/1997. *Groundwork of the Metaphysics of Morals*. Mary Gregor (translator and editor). Cambridge: Cambridge University Press.

Katz, Eric. 1985. Organicism, Community, and the "Substitution Problem'. *Environmental Ethics* 7: 241–56.

Katz, Eric, 1992. "The Big Lie: Human Restoration of Nature", in Frederick Ferré (ed.) *Technology and the Environment* (Research in Philosophy & Technology, vol. 12). Greenwich, Connecticut: JAI Press Inc.

Kemmerer, Lisa. 2006. *In Search of Consistency: Ethics and Animals*. Boston: Brill.

Klain et al. 2017. Relational values resonate broadly and differently than intrinsic or instrumental values, or the New Ecological Paradigm. *PLoS ONE* 12(8): 1-21.

Klötzli, Frank. 1993. Ökosystem. In Kuttler W (ed.) *Handbuch zur Ökologie*, 288-295. Berlin : Analytica.

Knippenberg et al. 2018. Relational value, partnership, eudaimonia: a review. *Current Opinion in Environmental Sustainability* 35: 39-45.

Korsgaard, Christine. 1983. Two Distinctions in Goodness. *Philosophical Review* 92 (2): 169–95.

Korsgaard, Christine M. 1997. Introduction. In Mary Gregor (translator and editor), *Groundwork of the Metaphysics of Morals*, vii-xxx. Cambridge: Cambridge University Press.

Kripke, Saul. 1980. *Naming and Necessity*. Cambridge, MA: Harvard University Press.

LaFollette, Hugh & Niall Shanks. 1996. The Origin of Specisiesm. *Philosophy* 71(275): 41-61.

Lee, Wendy Lynne. 2008. Environmental Pragmatism Revisited: Human-Centeredness, Language, and the Future of Aesthetic Experience. *Environmental Philosophy* 5(1): 9-22.

Lemons, John & Laura Westra (eds.). 1995. *Perspectives on Ecological Integrity*. Dordrecht: Kluwer Academic Publishers.

Lemons, John, Laura Westra & Robert Goodland (eds.). 1998. *Ecological Sustainability and Integrity: Concepts and Approaches*. Dordrecht: Kluwer Academic Publishers.

Lemos, Noah. 1994. *Intrinsic Value: Concept and Warrant*. Cambridge: Cambridge University Press.

Leopold, Aldo. 1949. *A Sand County Almanac*. Oxford: Oxford University Press.

Lewis, David. 1983. Extrinsic Properties. *Philosophical Studies* 44: 197-200.

Light, Andrew. 2002. Contemporary Environmental Ethics: From Metaethics to Public Philosophy. *Metaphilosophy* 33: 426-49.

Manela, Tony. 2018. Gratitude to Nature. *Environmental Values* 27(6): 623-644.

Marshall, Dan and Brian Weatherson, "Intrinsic vs. Extrinsic Properties", *The Stanford Encyclopedia of Philosophy* (Spring 2018 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2018/entries/intrinsic-extrinsic/>.

McCloskey H.J. 1987. The Moral Case for Experimentation on Animals. *The Monist* 70(1): 64 82.

McLaughlin P.B, 1995. Varieties of Supervenience. In E. Savellos, and U. Yalcin (eds.), *Supervenience: New Essays*, 16-59. Cambridge: Cambridge University Press.

McLaughlin, Brian and Karen Bennett. Supervenience. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition). <https://plato.stanford.edu/archives/sum2021/entries/supervenience/>.

McNaughton, David and Piers Rawling. 2007. Deontology. In David Copp (ed.), *The Oxford Handbook of Ethical Theory*, 424-458. Oxford: Oxford University Press.

McNaughton, S.J., and Michael B. Coughenour. 1981. The cybernetic nature of ecosystems. *The American Naturalist* 117(6): 985-990.

McShane, Katie. 2007. Anthropocentrism vs. Nonanthropocentrism: Why Should We Care? *Environmental Values* 16(2): 169-185.

McShane, Katie. 2009. Environmental Ethics: An Overview. *Philosophy Compass* 4/3: 407-420.

Mill, John Stuart. 2003. *Utilitarianism and On Liberty*. Edited with an Introduction by Mary Warnock. Oxford: Blackwell Publishing.

Morito, Bruce. 2003. Intrinsic Value: A Modern Albatross for the Ecological Approach. *Environmental Values* 12: 317-36.

Muradian, Roldan and Unai Pascual. 2018. A typology of elementary forms of human-nature relations: a contribution to the valuation debate. *Current Opinion in Environmental Sustainability* 35: 8-14.

Naess, Arne. 1973. The Shallow and the Deep, Long-Range Ecology Movement: A Summary. Inquiry 16: 95-100.

Naess, Arne and George Sessions. 1995. Platform Principles of the Deep Ecology Movement. In Alan Drengson and Yuichi Inoue (eds.), *The Deep Ecology Movement: An Introductory Anthology*, 49-53. Berkeley, CA: North Atlantic Books.

Nathanson, Stephen. n.d. Act and Rule Utilitarianism. *The Internet Encyclopedia of Philosophy*. < https://iep.utm.edu/util-a-r/ >.

National Institute of Allergy and Infectious Diseases (NIAID). 2021. The Important Role of Animal Research in mRNA COVID-19 Vaccine Development 18/08/2021. < https://www.niaid.nih.gov/news-events/role-animal-research-mrna-covid-19-vaccine-development>

NOAA. 2021. How do coral reefs protect lives and property? National Ocean Service website, https://oceanservice.noaa.gov/facts/eutrophication.html, 26/2/21.

Norcross, Alastair. 2007. Animal Experimentation. In Bonnie Steinbock (ed.), *The Oxford Handbook of Bioethics*, 648-667. Oxford: Oxford University Press.

Norton, Bryan G. 1984. Environmental Ethics and Weak Anthropocentrism. *Environmental Ethics* 6(2): 131-148.

Norton, Bryan G. 1987. Why Preserve Natural Variety? Princeton: Princeton University Press.

Norton, Bryan G. 1991. *Toward Unity Among Environmentalists*. New York: Oxford University Press.

Norton, Bryan G. 1995. Why I Am Not a Nonanthropocentrist: Callicott and the Failure of Monistic Inherentism. *Environmental Ethics* 17: 341-58.

Norton, Bryan G. 2002. *Searching for Sustainability*. Cambridge: Cambridge University Press.

Norton, Bryan G. 2009. Convergence and contextualism: some clarifications and a reply to Steverson. In Ben A. Minteer (ed.), *Nature in Common? Environmental Ethics and the Contested Foundations of Environmental Policy*, 87-100. Philadelphia, PA: Temple University Press.

Nuffield Council on Bioethics. 2005. *The Ethics of Research Involving Animals*. London.

O'Neill, John. 1992. The Varieties of Intrinsic Value. *The Monist* 75 (2): 119-137.

O'Neil, Rick. 2000. Animal Liberation versus Environmentalism: The Care Solution. *Environmental Ethics* 22: 183-190.

Parfit, Derek. 2011. *On What Matters* Vol. 1. Oxford: Oxford University Press.

Passmore, John. 1980. *Man's Responsibility for Nature: Ecological Problems and Western Traditions* 2nd ed. London: Duckworth.

Passmore, John. 2010. Attitudes to Nature. In David R. Keller (ed.), *Environmental Ethics: The Big Questions*, 103-110. London: Wiley-Blackwell.

Pattison, James. 2012. The Legitimacy of the Military, Private Military and Security Companies, and Just War Theory. *European Journal of Political Theory* 11(2): 131-154.

Peacock, Kent A. 2008. Ecosystems. In Sahotra Sarkar and Anya Plutynski (eds.), *A Companion to the Philosophy of Biology*, 351-367. Oxford: Blackwell.

Rabinowicz, Wlodek and Jan Österberg. 1996. Value Based on Preferences: On Two Interpretations of Preference Utilitarianism. *Economics and Philosophy* 12(1): 1-27.

Rawls, John. 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.

Raz, Joseph. 2011. *From Normativity to Responsibility*. Oxford: Oxford University Press.

Rea, Anne W and Wayne R Munns, Jr. 2017. The value of nature: Economic, intrinsic, or both?. *Integrated environmental assessment and management* 13(5): 953-955.

Regan, Tom. 1981. The Nature and Possibility of an Environmental Ethic. *Environmental Ethics* 3: 19-34.

Regan, Tom. 2004. *The Case for Animal Rights*. Berkeley & Los Angeles: University of California Press.

Roberts, M.A. 2011. An Asymmetry in the Ethics of Procreation. *Philosophy Compass* 6: 765-776.

Rogers, Chris Dinesen. 2019. How Does Sediment Affect Ecosystems? *Sciencing*. Accessed 7/9/2021 : <https://sciencing.com/soil-erosion-due-rainforest-deforestation-23042.html>

Rolston III, Holmes. 1988. *Environmental Ethics: Duties to and Values in the Natural World*. Philadelphia, PA: Temple University Press.

Rolston III, Holmes. 2012. *A New Environmental Ethics: The Next Millenium for Life on Earth*. London: Routledge

Ronnow-Ramussen Toni and Zimmerman Michael J. 2006. Introduction. In *Recent Work on Intrinsic Value*, Toni Ronnow-Ramussen and Michael J Zimmerman, xiii- xxxv. Dordrecht, The Netherlands: Springer.

Ronnow-Ramussen Toni. 2015. Intrinsic and Extrinsic Value. In Iwao Hirose and Jonas Olson (eds.), The Oxford Handbook of Value Theory, 29-43. Oxford: Oxford University Press.

Rosenberg, Alexander. 1992. Contractarianism and the "trolley" problem. J*ournal of Social Philosophy* 23(3): 88-104.

Ross, W. D. 1930/2002. The Right and the Good. Oxford: Oxford University Press.

Routley, Richard and and Val Routley. 1979. Against the Inevitability of Human Chauvinism. In K. E. Goodpaster and K. M. Sayre (eds.), *Ethics and Problems of the 21st Century*, 36-59. Notre Dame, IN: University of Notre Dame Pres.

Russow, Lilly-Marlene, 1981. Why do Species Matter? *Environmental Ethics* 3(2): 101-112.

Russow, Lilly-Marlene. 1988. Regan on Inherent Value. *Between the Species* 4(1), 46-54.

Sagdahl, Mathias S. 2018. Conscription as a Morally Preferable Form of Military Recruitment. Journal of Military Ethics 17(4): 224-239.

Samuelsson, Lars. 2009. *The Moral Status of Nature*. Saarbrücken: VDM verlag.

Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: The Belknap Press of Harvard University Press.

Scheffler, Samuel. 2011. Valuing. In R. Jay Wallace , Rahul Kumar, and Samuel Freeman (eds.), *Reasons and Recognition*, 23-42. Oxford: Oxford University Press.

Schulz, Christopher and Julia Martin-Ortega. 2018. Quantifying relational values — why not?, *Current Opinion in Environmental Sustainability* 35: 15-21.

Seidman, Jeffrey. 2009. Valuing and Caring. *Theoria* 75: 272-303.

Seidman, Jeffrey. 2016. The unity of caring and the rationality of emotion. *Philosophical Studies* 173: 2785-2801.

Singer, Peter. 1990. *Animal Liberation*. New York, NY: Avon Books.

Singer Peter. 1993. *Practical Ethics*, 2nd ed. New York: Cambridge University Press.

Singer, Peter. 2005. Ethics and Intuitions. *The Journal of Ethics* 9 (3/4): 331-352.

Sinnott-Armstrong, Walter, "Consequentialism", The Stanford Encyclopedia of Philosophy (Fall 2021 Edition), Edward N. Zalta (ed.), forthcoming URL = <https://plato.stanford.edu/archives/fall2021/entries/consequentialism/>.

Skorupski, John. 2003. Ethics. In Nicholas Bunnin and Eric Tsui-James (eds.), *The Blackwell Companion to Philosophy*, 202-230. Oxford: Blackwell Publishing.

Slote, Michael. 2001. *Morals from Motives*. Oxford: Oxford University Press.

Southwood, Nicholas. 2010. *Contractualism and the Foundations of Morality*. Oxford: Oxford University Press.

Stenmark, Mikael. 2009. The Relevance of Environmental Ethical Theories for Policy Making. In Ben A. Minteer (ed.), *Nature in Common? Environmental Ethics and the Contested Foundations of Environmental Policy*, 81-96. Philadelphia, PA: Temple University Press.

Stenseke, Marie. 2018. Connecting 'relational values' and relational landscape approaches. *Current Opinion in Environmental Sustainability* 35: 82-88.

Sterba, James. 2011. Biocentrism Defended. *Ethics, Policy and Environment* 14: 167-169.

Stratton-Lake, Philip. 2002. Introduction. In Philip Stratton-Lake (ed.), *The Right and the Good*, ix-l. Oxford: Oxford University Press.

Svoboda, Toby. 2011. Why there is no Evidence for the Intrinsic Value of Non-Humans. *Ethics and the Environment* 16(2): 25-36.

Taylor, Paul. 1981. The Ethics of Respect for Nature. *Environmental Ethics* 3(3): 197-218.

Taylor, Paul. 1986. *Respect for Nature: A Theory of Environmental Ethics*. Princeton, NJ: Princeton University Press.

Tenen, Levi. 2020. No Intrinsic Value? No Problem. Why Nature Can Still Be Valuable for Its Own Sake. *Environmental Ethics* 42(2): 119-133.

Timmons, Mark. 2012. *Moral Theory*. London: Rowman & Littlefield Publishers, Inc.

Thomson, Judith Jarvis. 1976. Killing, Letting Die, and the Trolley Problem. *The Monist* 59(2): 204-217.

United Stated Department of Agriculture. No date. Insects & Pollinators. USDA website. Accessed on 9/2/21: <https://www.nrcs.usda.gov/wps/portal/nrcs/main/national/plantsanimals/pollinate/>.

Varner, Gary. 1994. The Prospects for Consensus and Convergence in the Animal Rights Debate. *The Hastings Centre Report* 24(1): 24-28.

Varner, Gary E. 1998. *Nature's Interests? Interests, Animal Rights, and Environmental Ethics*. Oxford: Oxford University Press.

Verhoog, H. 1992. The Concept of Intrinsic Value and Transgenic Animals. *Journal of Agricultural and Environmental Ethics* 5(2): 147-60.

Warren, Mary Anne. 1987. Difficulties with the Strong Rights Position. *Between the Species* 2(4): 433-441.

Weston, Anthony. 1996. Beyond Intrinsic Value: Pragmatism in Environmental Ethics. In Andrew Light and Eric Katz (eds.) *Environmental Pragmatism*, 285-306. London: Routledge.

Westra, Laura. 1994. *An Environmental Proposal for Ethics: The Principle of Integrity*. Lanham, MD: Rowman and Littlefield Publishers, Inc.

Wildlife Conservation Society. 2020. 10 Reasons to Thank Nature. WCS.org. Access 2/9/2021: https://www.wcs.org/get-involved/updates/10-reasons-to-say-thanks-to-nature

Woodley, Stephen, James Kay & George Francis (eds.). 1993. *Ecological Integrity and the Management of Ecosystems*. Delray Beach: St. Lucie Press.

Worster, Donald. 1985. *Nature's Economy*. Cambridge: Cambridge University Press.

Zimmerman Michael. 1999. In Defense of the Concept of Intrinsic Value. *Canadian Journal of Philosophy* 29: 389-409.

Zimmerman, Michael J. and Ben Bradley. Intrinsic vs. Extrinsic Value. In Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2019 Edition). <https://plato.stanford.edu/archives/spr2019/entries/value-intrinsic-extrinsic/>