

Five Misguided Attempts to Defend Speciesism

Bryan Ross

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

School of Philosophy, Religion and History of Science

September 2020

The candidate confirms that the work submitted is his own and that appropriate credit has been given where reference has been made to the work of others.

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

The right of Bryan Ross to be identified as Author of this work has been asserted by Bryan Ross in accordance with the Copyright, Designs and Patents Act 1988.

Acknowledgments

I want to thank my supervisors, Gerald Lang and Rob Lawlor, for their priceless philosophical guidance. I am grateful for the patience, promptness, and diligence with which they have read and commented on so many drafts of each chapter of this thesis. This thesis would not have been possible without them.

Thank you very much to Jessica Isserow and Patrick Tomlin for giving up their time to be my examiners.

I am also very thankful for the financial support received from the University of Leeds, for the first 3 years of my PhD.

I am incredibly lucky to have met Adina Covaci, Alison Toop, Emily Paul, Gary Mullen, Marc Wilcock, and Will Gamester during my time at Leeds. I honestly do not think I could have done this without all of you.

Finally, I am indebted to my family - my parents, John and Belinda; my sisters, Cathy and Johanna; my brother, Johnny; my wife's family, John, Sean, and Chloe; and my wife, Lisa-Marie, whom the thesis is dedicated to. They have all kept me going and this thesis would not have been completed without them.

Table of Contents

Introduction	7
Chapter 1: Regan and the Problem of the Innocent Attacker: Falling Prey to the Predation Argument.....	14
1. The Predation Argument.....	15
2. A More Plausible Defence.....	20
3. A Deeper Worry.....	21
4. Regaining Consistency.....	24
4.1. Disposing of the Agency Requirement: Falling Prey to the Predation Argument.....	25
4.2. Disposing of the Permissibility Claim: Embracing a Duty of Martyrdom.....	30
5. Conclusion.....	33
Chapter 2: Regan and the Problem of the Innocent Attacker: Having One's Prey and Eating It Too.....	37
1. Justification Defeats Liability.....	38
2. Closing the Gap.....	41
3. Exploiting the Gap.....	54
4. Bear Attack Revisited.....	59
5. Conclusion.....	62
Chapter 3: A Series of Unfortunate Accounts of Misfortune.....	65
1. The Species Norm Account.....	67
1.1. Superchimp #1.....	68
1.2. Superchimp #2.....	69

1.3. Superchimp #1 Revisited.....	70
1.4. Superchimp #2 Revisited.....	73
2. The Intrinsic Potential Account.....	75
2.1. Problems with the Intrinsic Potential Account.....	79
2.1a. Intrinsic/Extrinsic Potential.....	79
2.1b. Connor & Abigail.....	84
3. The Role of Misfortune Verdicts.....	87
4. Conclusion.....	90
Chapter 4: How to Get Away with Speciesism.....	92
1. Is Speciesism a Mere Prejudice.....	93
2. Embracing Speciesism: Potentiality.....	103
3. Avoiding Speciesism: Personism.....	111
3.1. 'Typical' in what sense?.....	112
3.2. The relevant sense of 'kind'.....	115
4. Conclusion.....	119
Chapter 5: Humans are Persons or: How I Learned to Stop Worrying and Love the Generic.....	122
1. Essence and Generics.....	125
2. Generics and their Truth Conditions.....	130
3. Characteristic Dimensions of a Kind.....	132
4. Nature of Counterinstances.....	135
4.1. The Generics Debate.....	138
5. Essence and Personhood.....	139
5.1. Yet Another Plague.....	142

6. As For Those Pesky Intuitions.....	144
7. The Generics Debate Revisited.....	145
8. Conclusion.....	146
Conclusion.....	150

Introduction

Two well-known and related arguments within the animal ethics literature have interested me ever since I read Peter Singer's book *Practical Ethics* as part of an undergraduate 'Introduction to ethics' module. The first is the unfortunately named argument from marginal cases, and the second is the argument that discrimination based on grounds of species membership is objectionable for the same reasons that discrimination based on sex or race is. Both arguments are frequently appealed to by those who wish to defend some form of moral status for animals.

The argument from marginal cases, as described by Peter Singer, holds that human beings are, for the most part, not equal. If we are to seek some morally relevant property that all humans possess, it would have to be

a kind of lowest common denominator pitched so low that no human being lacks it. The catch is that any such characteristic that is possessed by all human beings will not be possessed only by human beings. For example, all humans, but not only humans, are capable of feeling pain; and while only humans are capable of solving complex mathematical problems, not all humans can do this. So it turns out that in the only sense in which we can truly say, as an assertion of fact, that all humans are equal, at least some members of other species are also 'equal' - equal, that is, to some humans (1975, p. 265).

The difficulty, then, is that for any morally relevant property that we can appeal to in order to justify giving more weight to the interests of humans, it will either be the case that some non-human animals also possess it, or it will be the case that not all humans possess it.

The second argument holds that those who attempt to defend the intuition that we should give more weight to the interests of humans than to the interests of non-human animals - in virtue of the fact that they are *human* interests - without pointing to a morally relevant property, are guilty of speciesism, a prejudice objectionable for the same reasons that racism and sexism are. Here is Singer again:

Racists violate the principle of equality by giving greater weight to the interests of members of their own race when there is a clash between their interests and the interests of those of another race...Similarly, those I would call 'speciesists' give greater weight to the interests of members of their own species when there is a clash between their interests and the interests of those of other species (1993, p. 58).

For Singer, the type of creature you are makes no contribution to the inventory of morally significant properties you have when deciding whether to count your interests, and how much weight to give to them. Species membership is morally arbitrary, or uninteresting, as is sex and race. For Singer, the properties of species, race, and sex conceal morally significant properties rather than present us with morally significant properties.

I have long found these arguments persuasive and yet I find it difficult to shake off what I refer to throughout this thesis as the 'Common-Sense Intuition': that the interests of humans count for more than those of non-human animals and that, for example, it is worse to perform medical experiments on a severely cognitively impaired human than on a comparably cognitively endowed non-human animal. Despite the fact that I spend the majority of this thesis arguing against arguments which attempt to justify this intuition, I take this intuition, and the search for a justification of it, as both my starting point and my motivation for the thesis. As I see it, there are two promising routes for defending the Common-Sense Intuition: we can either find some morally relevant property which all and only humans hold, or we can defend some form of speciesism which does not leave us exposed to parallel defences of malicious forms of prejudice such as racism and sexism.

Before I look specifically at attempts to defend this intuition, I spend the first two chapters considering problems which follow when we take animal rights seriously and, specifically, Tom Regan's argument that humans and animals possess equal inherent value and an equal right to be treated with respect. In chapter one, I discuss Regan's account of animal rights and

look at the predation argument which challenges it. According to Tom Regan (2004), animals have a right not to be killed. The 'Predation Argument' holds that if humans are often guilty of violating this right, then predators, too, violate the rights of other animals when they kill and eat them. This entails, the argument continues, that we should intervene to prevent predation and, as such, is often seen as a *reductio ad absurdum* of the animal rights position. I argue that Regan's attempted defence of this problem is unsuccessful, before going on to provide a more robust defence of his position. On Regan's view, predation does not involve injustice because the wolf is not a moral agent and thus cannot violate the rights of others. On the view I present, however, intervening with predation in the wild would involve injustice and for much the same reason: namely, that as a moral patient the wolf does not violate another's rights and, as such, would have its rights violated if attacked in an act of other-defence by a moral agent. However, this reveals an apparent tension in Regan's account when dealing with the permissibility of defending oneself from an animal attack. To resolve this tension he must discard at least one of two prior commitments. However, making use of the self-defence literature on defending oneself against a (human) innocent attacker, I argue that neither commitment can be easily amputated without causing Regan serious problems; rejecting the Permissibility Claim leads to positing an unintuitive duty of martyrdom, while rejecting the Agency Requirement causes him to fall prey to the Predation Argument once more. He is therefore faced with a serious dilemma.

In chapter two, I argue that given the entrenched nature of the dilemma presented in chapter one and how keen the point of each horn is, perhaps what we need to do is to find a way of keeping both of Regan's commitments in such a manner that Regan can avoid a charge of inconsistency. Here I consider the possibility of rejecting the dilemma and attempting to reconcile both of Regan's commitments within his theory by using well-known arguments within the self-defence literature and positing innocent threat cases as cases of symmetrical self-defence. This initially appears to be a promising avenue for Regan, as it at least allows for

self-defence against an attacking animal. However, it is clear that there remains work to be done with regard to the problems caused by the permissions operating around other-defence in such cases which would leave Regan's account defending claims which are no less absurd than that implied by the predation argument – that we have a duty to police nature.

In chapter three, I look at one promising attempt to justify the Common-Sense Intuition by holding that it seems plausible to suppose that a cognitively impaired human is unfortunate in a way that similarly endowed animals are not. Here 'fortune' is used to refer to how well or badly an individual's overall life is faring. Thus, an individual is unfortunate when her well-being falls short of an expected level. This chapter looks at two popular accounts of misfortune: the *Species Norm Account* and the *Intrinsic Potential Account*. The chapter takes as its starting point the assumption that if an account of misfortune is to be considered successful, it must clear the following two hurdles. Firstly, it should avoid being undermined by well-known counterexamples, namely, Jeff McMahan's *Superchimp* cases. Secondly, a successful account of misfortune must preserve and explain our widely and strongly held intuitions. Specifically, we want an account of misfortune to explain the following deeply held intuitions: that dogs are not unfortunate for not being persons while cognitively impaired humans *are* unfortunate for not being persons. This task, however, is far from straightforward. For, we want our judgements of misfortune to correlate with our verdicts of whether an individual is badly off or not. It will not do, for example, for an account of misfortune to conclude that an individual is unfortunate whose life intuitively appears to be going quite well overall. The problem, however, is that if we judge a dog not to be unfortunate for not having access to these levels of well-being and then apply the same criteria to the cognitively impaired human, it is difficult to maintain that she is badly off. For, if we only judge how well an individual's life is going by comparing her levels of well-being with those made possible by the individual's actual cognitive endowment, then it is difficult to maintain that a cognitively impaired human - who

has high levels of well-being by these limited standards - is not living a good life. It is clear, however, that this is not the result that we require from an account of misfortune. Rather, we want to claim that the cognitively impaired human *is* unfortunate for possessing the natural endowment that she has rather than one which would allow her to access the higher levels of well-being enjoyed by persons. But if the cognitively impaired are to be considered badly off for possessing the cognitive endowment that they do, then, presumably, non-human animals who share a similar endowment must be badly off in the same way. The difficulty, then, is to find a non-arbitrary way of restricting the scope of our judgements of misfortune, when the individual in question falls short of personhood, to members of the human species. This chapter suggests that neither account of misfortune is able to fulfil both goals: successfully avoiding being undermined by well-known counterexamples, and explaining our widely held intuitions that non-human animals are *not* unfortunate for not being persons while cognitively impaired humans *are* unfortunate for not being persons. The Species Norm Account is only able to clear the second hurdle by stumbling at the first, while the Intrinsic Potential Account, clears the first hurdle, only to crash headfirst into the second one. I argue that failing to clear either hurdle is enough to undermine an account of misfortune and, as such, neither account is wholly successful.

Chapter four presents a new set of problems to be overcome if we are to justify the Common-Sense Intuition (that the interests of humans count for more than those of non-human animals). The chapter begins by defending Peter Singer's claim that those who defend speciesism leave themselves open to parallel defences of other forms of prejudice such as racism and sexism. As such, those who wish to defend some preference for all humans over all animals must provide a justification for doing so which does not leave them open to such troublesome analogies. Such accounts, I argue, can be separated into two main camps: those who attempt to offer a defence of speciesism and those who attempt to justify the Common-

Sense Intuition while avoiding the charge of speciesism. Considering these in turn, I maintain that both types of arguments are hamstrung by similar problems.

Finally, in chapter five, I come full circle and address my initial motivation for the thesis. Here I set out the components of a persuasive account of the truth conditions of generics from the philosophy of language literature in order to address the following worry: if the argumentative work presented in the earlier chapters is as persuasive as I suggest, why then do I still find it difficult to shake off the Common-Sense Intuition that there is something morally significant about humankind and that, for example, performing medical experiments on severely cognitively disabled humans is worse than performing similar experiments on animals? The obvious explanation is that it refuses to disappear because the intuition is pointing to some deeper truth; that there really is something morally significant about humankind. Given that the arguments presented in this thesis, among others, have failed to deal the requisite killer-blow to this pesky intuition, I can either accept this conclusion, or offer an alternative explanation for why these nagging intuitions persist despite acceptance of arguments which discredit the intuition. This chapter takes the last of these positions, aiming to explain the persistence of the Common-Sense Intuition in a way that is consistent with the arguments presented in earlier chapters.

It is here that I must offer a *mea culpa* and tell you what the thesis is not: The discerning reader will have noticed from the above that I have not been in pursuit of any great substantive unity within this thesis. This thesis is interested in problems relating to the moral status of animals and, as such, looks at those problems which I find to be the most interesting and problematic. With that out of the way, I can tell you what it is and why I think it is significant. In many ways, the thesis can be distilled to the two main areas within the animal ethics literature that I am most interested in: rights and discrimination. Within these areas, I think, lie the most intractable problems and those which are most in need of being brought into the light and

clearly defined. I have wrestled with these problems throughout this thesis and have attempted to make some progress with them. While there remain problems left unsolved, I feel that by getting to the heart of what these problems are, and by clearly defining what would be required to solve them, I have left the overall state of the debate in a better position than which I found it in, and have hopefully provided a roadmap to achieving a successful defence of the Common-Sense Intuition.

Chapter 1

Regan and the Problem of the Innocent Attacker: Falling Prey to the Predation Argument

Introduction

According to Tom Regan's *The Case for Animal Rights*, animals have a right not to be killed which humans violate when they kill them for food (2004). It has been argued that this commitment leads Regan to "absurd" conclusions when applied to predator-prey relationships (Callicott, 1989, p. 45). This so-called 'Predation Argument' holds that if animals have a right not to be killed and eaten by us, then prey animals have a right not to be killed and eaten by predators, which leads to the "bizarre obligation to segregate predators from prey animals, disrupt natural food cycles, and in effect 'police' nature" (Cochrane, 2013, p. 132). Regan, however, has claimed that his theory can accommodate such an attack by relying on the thought that only moral agents can violate the rights of others and, as predators are not moral agents, they cannot violate the rights of prey animals when they kill them.

This chapter sets out to show that while Regan's defence of this problem is unsuccessful, he does in fact have the necessary tools in his kit to successfully defend his position against this problem. The thought is that we have no duty to save the prey animal not because, as Regan has argued, there are no rights involved in the situation, but rather because there *is* a right involved, but just not the one posited by friends of the Predation Argument. That is, although Regan is correct to say that the predator cannot violate the prey's rights, given that it is not a moral agent, in saving the prey animal we, who *are* moral agents, would be violating the rights of the predator.

However, I want to show that this defence reveals a deeper worry for Regan in how he deals with the permissibility of defending oneself, or another, from an animal attack. Given

certain commitments made by Regan in *The Case for Animal Rights*, he is faced with a charge of inconsistency and he must discard at least one of his prior commitments to remain consistent. Discarding any one of these though, I argue, will be more than a little problematic.

The chapter will proceed as follows. In section 1, I outline the Predation Argument, Regan's response, and how this response ultimately fails. In section 2, I show how Regan can deal effectively with this problem. In section 3, I outline the inconsistency in Regan's theory that this problem reveals. In the fourth section, I explore three possibilities for regaining consistency and demonstrate how problematic each of these is for Regan's overall theory of animal rights.

1. The Predation Argument

According to Regan, animals (or at least, mammals of a year or more) are the possessors of moral rights, because they, like humans, are subjects-of-a-life and all such subjects have equal inherent value. As beings with inherent value, human and nonhuman animals have, as a matter of justice, a basic moral right to respectful treatment. Encompassed within this general right is the more specific right not to be killed. Furthermore, Regan holds that we have a prima facie duty of assistance owed towards those individuals whose rights are being violated by others and that this duty is greater when the individual in question is unable to ensure that her rights are respected (Regan, 2004, p. 285).

The following problem presents itself on this view of animal rights. If we are to take seriously an animal's right not to be killed, as Regan clearly does, then we ought to try to save prey animals from having their rights violated when they are killed by predators in the wild. This would then have the odd consequence that we ought to implement some form of wildlife policing policy preventing predators from killing their prey whenever it is within our power to do so. This, in effect, is the Predation Argument.

Regan, however, argues that only moral *agents* can have duties towards others because, “only these individuals have the cognitive and other abilities necessary for being held morally accountable for what they do or fail to do” (Regan, 2004, p. 285).

Consider the following:

The Wolf and the Sheep: Tom is out for his morning walk and comes across a sheep grazing in a field. He spots a wolf approaching the sheep. The sheep does not see the wolf and will be killed and eaten if Tom does not intervene.

Regan argues that in such a case Tom has no duty to save the sheep from the wolf. He holds that just like most animals, wolves are moral patients, not moral agents, and thus cannot be meaningfully said to have duties towards others, and in particular, cannot have the duty not to kill the sheep. Therefore, it cannot be claimed that the wolf violates the sheep’s right not to be killed as moral patients are incapable of violating the rights of others. As such, the wolf cannot wrong the sheep by attacking it even if the wolf harms it in doing so; the wolf remains innocent of any wrongdoing. In short, agency is required for violation.

However, this attempt to save Regan’s theory from the conclusion drawn by the Predation Argument fails due to what Regan thinks about duties of assistance. While it might be true that moral patients cannot violate the rights of others and that wolves and other predators are moral patients, rights violations are not the only thing of importance in our moral systems. We also have duties of assistance, which although are often grounded in preventing rights violations, can also be grounded in the *prima facie* duty of beneficence. Regan himself recognises this, holding that humans owe this duty to each other independently of the demands of justice. Thus, just as we would owe it to a hiker to warn him of a boulder falling his way, despite the fact that it is not a matter of justice that we do so since the boulder cannot violate the rights of anyone, Regan maintains that if the wolf were attacking not a sheep but a human child (who is also a

moral patient), we would have a prima facie duty to intervene, grounded in a duty of beneficence (2004, p. xxxvi).¹ Given that he holds that both human and nonhuman animals have equal inherent value, why should he not think that we owe such duties to animals in similar situations rather than just reserving our assistance duties towards animals for when rights violations are occurring?

The reason, it would seem, is due to another obligation with regards to *wild* animals which takes precedence over our duties of assistance to them when their rights are not being violated. This is an obligation to “let them be” which is grounded in a “recognition of their general competence to get on with the business of living” (Regan, 2004, p. xxxvii). He thinks that we best honour the competence of both predator and prey by allowing them to exercise their natural abilities despite their competing needs. Animals, then, can take care of themselves in the wild and do not require our help to survive. Donaldson and Kymlicka make a similar point holding that “wild animals are competent to address the challenges they face, including the challenges of predation and food cycles, and have evolved ways of life and flourishing that are intimately tied up with meeting these challenges” (2013, p. 154). A human child, on the other hand, *does* need our help to survive and, as such, we do not best honour the child’s competence by “letting her be”.

Two points should be mentioned here. The first is that it is not obvious that animals are in fact competent “to get on with the business of living”. The fact is that life for many wild animals, and especially for those animals considered prey, is lived in continual fear and danger of violent death. It is thus difficult to see how allowing an animal to be torn to pieces and devoured is respecting its competence and allowing it to flourish. After all, how can permitting

¹ It is important to make it clear here that for Regan, duties of justice are concerned solely with rights violations, where an individual is wronged (by a moral agent). While not uninteresting, a full discussion of the debate over duties of justice versus duties of beneficence, their respective normative rankings, and which should be properly said to ground duties of assistance, would lead me too far astray in this chapter.

this to happen to an animal be to provide “them with the opportunity to live their own life, by their own lights, as best they can” (Regan, 2004, p. 357)? It would seem that Regan loses sight of the individual here when he invokes this argument from competence for, as McMahan argues, while it is true that prey and predator *species* often manage to persist and survive, it is false that *individual* prey do not need our help to survive (McMahan, forthcoming, pp. 14-15). Such animals would do much better than they do at present if we were to protect them against being killed by predators. So given that, in some species, only a small percentage of those born manage to make it to adulthood, it does not seem possible to claim that they are competent in their ability to survive and flourish; indeed, quite the opposite would appear to be true.

Secondly, and as Ebert and Machan (2012, p. 149) have recently claimed, it is unclear why we should accept that permitting a sheep to be torn apart and devoured by a wolf is honouring its competence, but saving a human child from the same fate is honouring hers. Indeed, there appears to be some form of speciesist discrimination going on in Regan’s thinking here as there do not seem to be any morally relevant differences from the victim’s perspective between a child being eaten by a wolf and a sheep being eaten. It is in neither the child’s nor the sheep’s best interests to be ripped apart by a wolf, nor do either of them have the opportunity to exercise their natural capacities when this happens, or get to “live their own life, by their own lights, as best they can”. Unless Regan can point to some characteristic that the child has, but the sheep lacks, which is relevant to why it is in the sheep’s interests to be killed by a wolf, but not in the child’s, and is not just based on the fact that the sheep fails to be a member of the right species, then he is open to a charge of speciesism here.²

² Moreover, this line of thought seems troublesome for Regan when we consider our duty of aid to an adult human when she is being attacked by a wolf. Unlike a human child, a normal adult human does not need our help to survive and is competent to live her life (Simmons, 2009, p. 20). It is thus unclear why we should intervene to save an adult human from being killed by a wolf if the primary reason that we should save a child, but not a sheep, is that the child needs our help to survive and that her competence is not best honoured by “letting her be”.

Perhaps, though, a more charitable way to present Regan's 'competence argument', which avoids the charge of speciesism, is to adopt a more holistic perspective where predation is a part of a larger whole. Here we might think that the flourishing of the overall system is necessary for the flourishing of individual animals. For instance, Donaldson and Kymlicka hold that while it is true that individual animals are killed and suffer, this is not a catastrophe because the communities go on existing and are competent in doing so (2013, p. 182). In a similar vein, Schlosberg has criticized accounts of flourishing which focus on individual flourishing in abstraction from the flourishing of species and ecosystems as being too narrow. He holds that it does not make sense to consider an individual's flourishing in isolation from "the environment in which this flourishing is to occur" and that "part of the flourishing of animals is to be the protein for other life forms" (2007, pp. 148-151).

The problem with this, though, is that it does not seem possible to have justice for individual animals, species, *and* ecosystems simultaneously as Schlosberg wants to claim. As we have seen, flourishing for an individual wolf, if this is taken to mean exercising the capability to hunt and kill deer, is simply not compatible with the flourishing of the individual deer which is being killed and eaten. Moreover, even if we accept that deer have a wolf-protein function which not only allows individual wolves, the wolf species, *and* the deer species to flourish, this is still in conflict with the flourishing of individual deer. For, it might well be that having this protein function allows the *species* deer to flourish, as it strengthens the herd by 'weeding' out the weak and the infirmed, and by preventing overpopulation, but, as Elizabeth Cripps (2010) has forcefully argued, it still remains highly implausible that it is conducive to the flourishing of the individual deer that it be killed and eaten by a wolf. As such, it is clear that doing justice to the individual deer, by preventing it from being killed by predators, means not only doing injustice to the individual wolf, but also to the species deer (Cripps, 2010, pp. 11-12). Similarly, protecting the flourishing of an ecosystem often conflicts with the flourishing of not

just individual animals but also with species, such as the killing of rabbits to protect the Australian bush, and the encouragement of land owners in Scotland to kill grey squirrels so as to protect the native red squirrel from extinction (Cripps, 2010, p. 11; Cochrane, 2013, p. 137).³

Thus, the problem for Schlosberg is that he wants to have justice for individuals, species, *and* ecosystems. But justice for each one of these three categories is incompatible with justice for the other two. Furthermore, we should keep in mind that what Regan wants is justice for *individual* animals and what the above demonstrates is that such a goal is incompatible with an environmental holistic viewpoint. To adopt such a viewpoint would be to “simply abandon ethical individualism” (Wissenburg, 2011, p. 404): something for which rights seem specifically designed to protect. Moreover, for Regan, the flourishing of a given species is only important insofar as it is conducive to the flourishing of the individual members of that species (2004, pp. 359-61). For example, when discussing how the rights view should think about endangered species, he argues that “the reason we ought to save the members of an endangered species of animals is not because the *species* is endangered but because the *individual* animals have valid claims and thus rights” (2004, p. 360 emphasis added). As such, Regan, unlike most environmentalists, is prepared to sacrifice not only the flourishing the ecosystem but also that of the species in aid of the flourishing of the individual animal.

2. *A More Plausible Defence*

While Regan’s attempted defence of his theory is, I think, ultimately unsuccessful, Regan does have the necessary tools in his theory to mount a successful defence of it against the Predation Argument. Regan is on the right track when he argues that the reason we cannot save the sheep from the wolf is because the wolf is not a moral agent and thus not capable of violating the

³ Note that these are also types of intervention in wildlife and so such a policy does not adhere with Regan’s one of “letting them be”.

rights of the sheep. However, he is mistaken in thinking that there are no rights at all at stake in *The Wolf and the Sheep*. For while neither the wolf nor the sheep is a moral agent and thus while they cannot have duties *to* others, they can have rights *against* others; namely, against moral agents. Tom is a moral agent and, as such, if he were to intervene he would violate the rights of the wolf in doing so. For let us suppose that Tom has exhausted all non-violent means to prevent the wolf from killing the sheep and can only save it by harming or killing the wolf. As the wolf is not threatening the rights of the sheep and, as such, is innocent of any wrongdoing, Tom's intervention cannot be a matter of justice as justice only demands that we try to prevent harm caused by violations (due to agents) but not harm caused in other ways (due, for example, to other moral patients). Thus, the justification for Tom's intervention must be a matter of beneficence. But Regan has claimed in no uncertain terms that "the demands of justice always take precedence over the claims of beneficence" (2004, p. xxvii) as "no one has a right to be protected against being harmed if the protection in question involves violating the rights of others" (2004, p. 346). As such, the Predation Argument fails because we are not forced to accept the conclusion that we owe a duty of assistance to prey animals to prevent them from being attacked by predators in the wild.

To summarise then, on Regan's view, predation in the wild does not involve injustice because the wolf is not a moral agent and thus cannot violate the rights of others. On the view I have presented, however, intervening with predation in the wild *would* involve injustice and for much the same reason: namely, that as a moral patient the wolf does not violate the rights of any other and, as such, would have its rights violated by a moral agent if attacked in an act of other-defence.

3. *A Deeper Worry*

Now, consider the following case:

Bear Attack: Tom is out for his morning walk and comes across a bear charging for a stranger, Peter, in an attempt to kill and eat him. Tom has a rifle with him capable of killing a bear and has no other means of preventing the bear from attacking Peter.

May Peter permissibly kill the bear in self-defence? May Tom permissibly kill the bear in other-defence? In discussing a similar case to this, Regan holds that despite the fact that the animal is guilty of no moral offence in attacking us, it “poses a distinct threat and we do no wrong if we harm the animal in the course of defending ourselves” (2004, p. 296). But notice now something which became clear in the context of defending Regan’s theory from the Predation Argument. Because there is no rights violation in the offing when the bear attacks Peter, and because Tom’s duty of beneficence to Peter cannot take precedence over his duties of justice to the bear, it seems that Regan should be forced to say that Tom is not permitted to kill the bear to prevent it from attacking and killing Peter. For, while the bear, as a moral patient, is unable to violate Peter’s rights, Tom, as a moral agent, *would* violate the bear’s rights not to be harmed should he do so in attempting to save Tom. Similarly, on an account that takes the rights of individual animals seriously, it would appear that Peter is also not permitted to kill the bear in self-defence as the bear is not threatening to violate his right not to be killed despite the fact that if Peter does nothing the bear will in fact kill him. The bear is innocent and thus will not wrong Peter if it kills him and, as such, has done nothing to make itself liable to defensive harm.⁴ What then grounds Peter’s right or permission to defend himself against the bear? If we are to take the bear’s right not to be killed seriously then we are owed a story

⁴ Of course, we might think that neither has Peter done anything to make *himself* liable to defensive harm. We might then be tempted to argue Peter has a right, or at least a permission, to use defensive force grounded in the fact of his own non-liability. We can attempt to articulate the structure of this victim-centred right more clearly: On such a view, the right to self-defence of the defendant, D, is grounded in D’s non-liability. As such, whenever D is put in a self-defensive situation by an attacker, A, we don’t need to establish: ‘Is A violating D’s right?’ We need only ask: ‘Is D liable to be killed by A?’. While it is not clear that Regan holds such a view, in the following chapter I attempt to flesh out his account by adopting this argument. While plausible, such a view, I argue, is ultimately unable to bring Regan the consistency he needs. I leave this aside for now.

as to why Peter does not violate the bear's right when he uses defensive force against it or, if he does, why such an action is permissible in these circumstances.⁵

Perhaps one reason that Regan thinks that such defensive action is permissible when a bear is attacking a human, despite the fact that he thinks that the bear does nothing wrong and violates no rights, lies in his treatment of the moral worth of individuals. He puts forward the following case:

The Lifeboat: Four normal adult humans and a dog are stranded on a lifeboat. The lifeboat can only accommodate four individuals and, thus, one must be thrown overboard.

In such situations, Regan holds, it is the dog that should be thrown out of the boat as normal adult humans have lives which, all else being equal, are more valuable than that of dogs, and animals more generally. This is because death for one of the humans would be a greater harm than death for the animal as "the loss which death represents is a function of the number and variety of possible sources of satisfaction" which it precludes, and death for a normal adult human represents a greater loss of possibility than for an animal (2004, p. xxix). However, there is a morally salient difference between *Bear Attack* and *The Lifeboat*. In the latter we are forced to choose between infringing one of the humans' rights on the one hand, and infringing the dog's right on the other. No matter what we decide, someone's right not to be killed will be infringed and thus some other theory must be applied to help us decide between conflicts of rights. In the former case, though, what we are deciding between is not a case of competing right claims, but rather the violation of the bear's right not to be killed on the one hand, and the harm done to Peter on the other. Yet if we are to take Regan's claim that animals have an equal right not to be killed seriously, then there seems to be no grounds for overriding the bear's right

⁵ This is part of a larger debate about whether it is ever permissible to kill an innocent attacker or threat, but is made much more complicated here by Regan's commitment to animal rights. I turn to this in more detail in the following section.

not to be killed. Peter's life *is* threatened, and in other circumstances we would owe him a duty of assistance to help him but, as Regan argues, we cannot do so at the expense of another individual's rights.

Regan faces a charge of inconsistency here as it would appear that he cannot hold both that moral patients cannot violate the rights of others (the very claim that allowed him to escape the Predation Argument) *and* that it is permissible to kill moral patients in self-defence.⁶ It seems that in order to regain consistency within his theory of animal rights he must reject one of these two commitments. However, I argue in the following section that rejecting either of them leads Regan into trouble; rejecting the latter leads to positing an unintuitive duty of martyrdom, while rejecting the former causes him to fall prey to the Predation Argument once more. He is thus faced with a serious dilemma.

4. *Regaining Consistency*

Here are the two commitments Regan makes in *The Case for Animal Rights* stated again:

Agency Requirement: Moral patients cannot violate the rights of others (agency is necessary for violation);

Permissibility Claim: It is permissible to kill moral patients in self-defence.

If we add to these the following two claims:

Animal Rights: Humans and animals possess equal inherent value and an equal right to be treated with respect;

⁶ Whether the defence I offered in section 2 is successful or not, the inconsistency in Regan's theory remains due to his prior commitments.

Justice Defeats Beneficence: Duties of justice, all things being equal, take precedence over duties of beneficence.

It then becomes clear that there is an inconsistency in Regan's theory here and one of the above commitments must be discarded. Here, I'm going to assume that it will cost Regan less to get rid of one of the former two claims rather than one of the latter two. As such, I will focus only on the possibility of one of the former two claims being sacrificed.

4.1. Disposing of the Agency Requirement: Falling Prey to the Predation

Argument

I begin with the first horn of the dilemma, the Agency Requirement. Other rights theorists have abandoned this requirement when dealing with the issue of self-defence. So one strategy is to look at such an account and to see if Regan can adopt a similar line. For example, Judith Jarvis Thomson has argued that agency is not required for the violation of rights and, as such, that it is permissible to kill a moral patient in self- and other-defence. While the first half of this claim may sound strange, the second half is at least quite intuitive to most people.

She gives the following example:

Innocent Threat: You are lounging on a deck chair when a gust of wind blows a fat man off a cliff above you. He will certainly kill you if he lands on you, although he will survive. Your only option to save yourself is to move your awning in order to deflect the falling fat man away from you onto some rocks, but doing so will certainly kill him (1991, p. 287).

Thomson thinks that you are not liable to be killed by the fat man and, as such, you have a right that he not kill you; a right that he will violate by landing on you. Thus, it is permissible for you to move the awning. She argues that fault makes no morally relevant difference in such cases so that despite the fact that the fat man is innocent, it remains the case that he will violate

your right not to be killed and therefore lacks a right not to be killed himself. On this view, then, violating another's right to be killed causes one to lack that same right.⁷ Indeed, such a view would seem to fit with the brief comments made by Regan when discussing innocent attacker cases. Regan himself suggests that because we are innocent of any wrongdoing when we are attacked by an innocent attacker, we are permitted to defend ourselves even if this requires harming the attacker, provided non-violent means have been exhausted (Regan, 2004, p. 293). So perhaps we can flesh out Regan's view by adopting Thomson's arguments and hold that because Peter is not liable to be killed by the bear, the bear violates his right not to be killed when it attacks him and thus lacks a right not to be killed itself. Peter therefore does not violate any of the bear's rights when he kills it in self-defence.

One problem for both Thomson and Regan with holding that agency is not required for rights violations, is that they both view rights as having correlative duties. To say, then, that we have a right against a moral patient that he not kill us, is to say that he has a duty not to kill us. However, it seems more than a little odd to think of moral patients as having any kind of duties at all towards anyone. On Regan's view, "I cannot have a valid claim (i.e. a right) if I do not have a valid claim-against someone, and I cannot have a valid claim-against someone if that individual does not have a duty to me to do or forbear doing the act I claim is owed to me" (2004, p. 272-3). He claims that whether an individual can in fact have a duty to another "depends in part on whether it is within the powers and capacities of those against whom the claim is made to do or forbear doing what is claimed [as]...to make claims to something against someone cannot be valid if they cannot be satisfied" (2004, p. 272). Thomson, too, holds that we cannot strictly be said to have a right unless we can point to someone whom we have the

⁷ Thomson also thinks that the fat man has *forfeited* his right not to be killed as fault is similarly not required for forfeiture. However, she notes that even if we are reluctant to say that he forfeits his right in this situation, the fat man, in any case, *lacks* the right not to be killed here (1991, 301). While beyond the scope of this chapter, we might wonder whether Thomson needs this further claim for her argument: that the innocent attacker loses the right not to be killed by violating another's right not to be killed.

right against and is thus under a duty to accord our claim; “having a duty as a correlative is...central to our notion of rights” (1990, p. 40). On Regan’s account, then, to deny that agency is required for the violation of rights would be to accept that animals can have duties towards us, duties which are not “within the powers and capacities” of them to do or forbear from doing. It seems, then, that neither Thomson nor Regan are able to punch their way out of this Hohfeldian circle.

That said, even if we could accept that moral patients, and non-human animals specifically, can have duties towards others, there is a much larger problem that Regan, but not Thomson, must face; namely, the Predation Argument discussed in section 1. If Regan is to abandon the Agency Requirement (that moral patients cannot violate the rights of others), his commitment to a strong animal rights position would seem to force him to accept the claim that we must police wildlife so as to prevent predators from killing prey animals and this is plainly a position that he is keen to avoid. Indeed, if he was to abandon the Agency Requirement, he cannot even appeal to his competence argument discussed in section 1 as to do so would be to ignore the rights of the sheep and to treat it as a “mere receptacle of value” after all. For, the competence argument can only work if there are no rights-violations at play. But if the wolf does in fact violate the rights of the sheep, then protecting the sheep’s rights must take precedence over respecting the competence of the sheep and the wolf. To fail to do so would be to fail to treat the sheep as a being with inherent value and to fail to respect its rights.

Perhaps, however, Regan can adopt some form of the flourishing or competence approach discussed in section 1, and contend once more that there are no rights violations occurring in natural predation even though agency is not required for violation. The thought would be that although the sheep has a right not to be killed, the content of such a right does not include a right not to be killed by a wolf or any non-human animal. For instance, Alasdair Cochrane has argued that when all things are considered the prey animal’s interest in continued life is not

sufficient to ground a right that they not be killed by predators. This is because the wolf is not able to lead a wolf-like life and flourish without killing prey. Thus, for Cochrane, because the killing of prey is necessary for survival, not just for the predator but also for “scavengers, who depend on such kills” (2013, p. 134) the “right course of action is to adopt a general policy of non-interference” (2012, pp. 94-5).⁸ However, because “the survival of others is not at stake” (2012, p. 94) when humans kill animals to eat them, the animal’s interest in life *is* sufficient to establish a right that they not be killed by us. If this is right, and the defence I offered in section 2 is correct, then we also have a positive reason not to intervene in predation as to do so would be to cause the violation of the predator’s rights.

This kind of rights specification appears ad hoc and question begging for a number of reasons. As Steinhoff suggests against rights specification theories in general, “stipulatively specifying rights such that they conform to one’s philosophical predilections does not amount to an argument” (2016 p. 1020). For instance, consider the fact that, on Cochrane’s account, what determines the content of the sheep’s right not to be killed in this particular case is the flourishing of animals, or as he puts it, the fact that the survival of another animal is at stake. But if what is important is the flourishing of individual animals, or at least if the flourishing of animals is important enough to place restrictions on an individual’s right not to be killed, then surely, on any rights account, the prey animal’s flourishing should be more important than the predator’s, given that it is the content of the prey animal’s right that is in question. Indeed, it would seem that the prey animal’s flourishing would arguably be important enough to ground a right not to be killed against any predator since in order for it to flourish and to lead a sheep-like life, it is necessary for it not to be killed and eaten by a predator. Thus, the fact that the wolf cannot lead a wolf-like life without killing sheep seems irrelevant to the question of

⁸ Another possibility might be to accept that rights are violated, but that alternative intervention-based systems would self-defeatingly worsen the prospects for all these creatures. It’s not clear that Regan could take advantage of this, though, given the importance he places on individual rights.

whether the sheep's interests are sufficient to ground a right not to be killed. What matters for the establishment of rights, at least on accounts like that of Regan's and Cochrane's, is that the sheep has a strong interest in not being caused the terror, agony, and death that accompanies being eaten by a wolf; an interest that is strong enough, if any is, to establish a right not to be treated in this way.

What is more, even if we could accept that the content of the sheep's right is such that it does not possess a right not to be killed by predators in the wild, it is not clear that this can provide Regan with the relief that he seeks in the *Bear Attack* case. It may be true that this, in conjunction with the defence offered in section 2, would help him to avoid the Predation Argument but it offers him no help with the current problem. For, we are considering the rejection of the Agency Requirement (that moral patients cannot violate the rights of others) as a means of allowing Regan to keep the Permissibility Claim (that it is permissible to kill moral patients in self-defence). That is, we want to be able to say that the bear violates Peter's rights when it attempts to kill him and that it is *this* fact which makes the defensive killing of the bear permissible. But if the reason that the sheep holds a right against us not to be killed, but not against the wolf, is that we do not need to eat it in order to flourish, then why should the content of Peter's right not to be killed not be the same as the sheep's? That is, why would humans have a general right not to be killed which includes the specific right not to be killed by predators, instead of just a right not to be killed by other beings which do not need to kill us in order to survive? There appears to be no obvious reason that the content of the sheep's right not to be killed would take this form but that ours should be somehow different and should protect us against predators.

Once again, this would seem to involve some kind of discrimination on the basis of species membership for there appears to be no morally relevant factor which would suggest that the sheep does not possess a right not to be killed by the wolf and, as such, we are unable to save

it in an act of other-defence, but that Peter does have a right not to be killed by the bear and that it *is* permissible to protect him in other-defence. Why should prey animals be sacrificed on this account for the flourishing of predators but not humans? It also makes little sense to claim that these animals are somehow the ‘natural’ prey of predators while we are not, as this seems merely to be a function of location and the fact that we, as a species, have attempted to carve out pieces of the world which eliminate the threat of predation.⁹

If the reasoning above is correct and the bear is unable to violate our right not to be killed because we do not possess such a right against the bear, then although rejecting the Agency Requirement might provide Regan with some protection against the Predation Argument, it offers him little or no help in regaining consistency. Again Regan is left with a similar kind of dilemma facing him here: he can avoid the Predation Argument but only at the cost of sacrificing the intuitively appealing feature of his account that animals are liable to defensive killing. Thus, Regan is back to square one and gains little by abandoning the Agency Requirement, as if the bear does not violate Peter’s rights when it attacks him then he is still left with no justification for violating the bear’s rights when he kills it in self-defence.

4.2. Disposing of the Permissibility Claim: Embracing a Duty of Martyrdom

That leaves the possibility of abandoning the Permissibility Claim, the second horn of the dilemma. And although problematic, perhaps relinquishing the Permissibility Claim *is* Regan’s best option, or his least bad one at any rate, to regain consistency. Of course this will have what many take to be the extremely unintuitive consequence that we cannot permissibly defend ourselves against animal attacks if it means harming the animal in doing so. However,

⁹ See Hatley (2004) for a defence of the thought that we (humans) are the ‘natural’ prey of large predators in certain environments.

perhaps this is not quite as unintuitive as it first sounds. After all, others have made this move with regards to what we can do to defend ourselves against a human moral patient, that is, a human innocent attacker or threat, and McMahan, for example, put forward similar arguments for the conclusion that it is impermissible to kill a moral patient in self-defence. They argue that in order to be liable to be killed or to lack the right not to be killed, the threat must have “moral responsibility” and that this requires agency (Otsuka, 1994, p. 90; McMahan, 1994, p. 276). Otsuka claims that we would not think it possible for a stone to violate your rights if the stone constituted the threat in *Innocent Threat* rather than a fat man. Likewise, he continues, there seems no reason to think that “the rights-violating power of such a human object could be any greater than the rights-violating power of a chunk of granite” (1994, p. 79). On this view, then, despite the fact that the fat man and the stone have a different moral status insofar as what may be permissibly done to them, there is no morally salient difference between them when it comes to whether or not they can violate the rights of an individual. McMahan’s Responsibility Account, too, contends that it is simply not enough that the fat man harms you by falling on you, as to make it permissible to kill him he must also wrong you and, in doing so, violate your rights (2005, p. 401).

However, notice what happens when we add a commitment to animal rights to such accounts. We are left with the same unintuitive conclusion that Regan is now facing; that we cannot harm an animal in self-defence. This would seem to impose on all moral agents a ‘duty of martyrdom’ when it comes to innocent aggressors and threats (Otsuka, 1994, p. 67). However, just like Thomson, neither McMahan nor Otsuka are committed to a strong animal rights position. Thus, they are not forced to commit themselves to the conclusion that it is impermissible to kill an animal in self-defence and, as such, can reject the Permissibility Claim much more readily. Regan, though, being committed to such a position, is forced to accept this unappealing conclusion if he decides to rid himself of the Permissibility Claim.

Moreover, such a duty would seem to go against our common-sense intuitions about self-defence and what morality requires one to do. As Kaufman holds, one way to see this is by asking what type of penalties it would be appropriate to visit on those who failed to martyr themselves to an innocent aggressor or threat (2010, p. 94). The implausibility of inflicting any punishment at all on those who refuse to save an innocent lethal threat at the expense of their own lives, Kaufman continues, is a mark against those accounts which require such action. Indeed, this appears only to become more implausible when we consider that the threat in question is a nonhuman animal. Regan himself has argued that, even in the human case, because accepting a duty of martyrdom “is so at odds with our considered beliefs, in a broad range of cases...we have principled reasons to reject” it (2004, 289).

While it may be true that Otsuka and McMahan accept this duty of martyrdom when it comes to human moral patients, Otsuka at least, is clear that he wants to avoid accepting this in the case of an attacking animal (Otsuka, 1994, pp. 92-3).¹⁰ Otsuka does this by relying on the difference between how valuable a normal adult human’s life is compared with that of an animal. He argues that while it is impermissible to kill a human moral patient in self-defence, it is in fact permissible for Peter to kill the bear in *Bear Attack*. This, he maintains, is because Peter is a human *person*, “whereas it is merely a grizzly bear, and human persons are worth more than grizzly bears” (1994, p. 92). One thing to note here, though, is that if we could show that some animals are persons, Otsuka would seem committed to accepting, as a consequence of his theory, that we are not permitted to kill them in self-defence. So, for example, if it’s right that primates and cetaceans should be considered persons, and yet not quite moral agents, then if Peter was instead attacked by a bonobo say, rather than a bear, Otsuka would be forced to

¹⁰ McMahan (2010; forthcoming) has claimed that we have reason to seek the extinction of carnivorous species and to replace them with new herbivorous species, provided this could be done without significant ecological upheaval. However, this does not show that he thinks that it is impermissible to defend oneself against an attacking animal, only that he thinks that it would be better if there were less pain and suffering in nature.

accept that Peter is not permitted to use defensive harm against the attacking animal despite the fact that he would be permitted if it was an animal nonperson. This seems a strange consequence.¹¹ In any case, Regan too, as discussed in *The Lifeboat* above, thinks that the life of a normal adult human is worth more than the life of an animal, all else being equal. Perhaps, then, it is open to Regan to adopt Otsuka's approach for the permissibility of defending oneself against an animal attack.

Given Regan's commitment to animal rights, though, all else is not equal and, as such, adopting such a position is not feasible. On his view, it is irrelevant how much more valuable normal adult humans (or persons for that matter) are in comparison to animals, when there is a right's violation in the offing and there are no competing rights at stake (and given that we are considering the rejection of the Permissibility Claim rather than the rejection of the Agency Requirement, there *can* be no competing rights at stake). Again, to fail to treat such considerations as irrelevant in this situation would be to fail to treat the bear as a being with equal inherent value and to fail to respect its rights which it is due as "a matter of strict justice" (2004, p. 329), treating it instead as "a mere receptacle of value" (2004, p. 289) after all – the very treatment that Regan has set out in *The Case for Animal Rights* to protect animals from.

5. Conclusion

I have argued, in section 3, that there is an inconsistency in Regan's thinking in *The Case for Animal Rights* and that he faces a serious dilemma in trying to regain consistency for his theory of animal rights. Furthermore, in section 4, I argued that neither horn of this dilemma is an attractive option for Regan. On the one hand, abandoning the commitment to the requirement of agency for the violation of rights causes Regan's theory to fall prey to the Predation

¹¹ Although it may very well be a consequence that Otsuka is willing to accept. He might explain away the feeling of oddness by pointing either to the relative novelty of the thought that bonobos are really persons, or a residual speciesism, which is surely bound to skew our intuitions.

Argument, an argument for which, as discussed in sections 1 and 2, Regan seems only able to satisfactorily reply to if he holds onto the Agency Requirement. On the other hand, abandoning the second of his commitments, that it is permissible to kill a moral patient in self-defence, leads Regan to posit a duty of martyrdom in accepting the unintuitive conclusion that we cannot defend ourselves against an animal that is trying to kill us if it means harming it, despite the fact that we are innocent. It would appear that neither can be easily amputated from Regan's account without causing serious problems elsewhere for his overall theory of animal rights. Perhaps one of these options is more attractive, or less unattractive, than the other, but if so we require some reason for thinking that the conclusions which follow from abandoning either of them are not as dire as they seem to be for his account.

In the following chapter, I look at the possibility of rejecting the dilemma and attempting to reconcile both of Regan's commitments within his theory by positing innocent threat cases as cases of symmetrical self-defence.

Works Cited

- Callicott, J. B., 1989. Review of Tom Regan, *The Case for Animal Rights*. *Defense of the Land Ethic: Essays in Environmental Philosophy*, pp. 39-48.
- Cochrane, A., 2013. Cosmozoopolis: The Case Against Group-Differentiated Animal Rights. *Law, Ethics, and Philosophy*, Volume 1, pp. 127-141.
- Cochrane, A., 2012. *Animal Rights Without Liberation: Applied Ethics and Human Obligations*. New York: Columbia University Press.
- Cripps, E., 2010. Saving the Polar Bear, Saving the World: Can the Capabilities Approach do Justice to Humans, Animals and Ecosystems?. *Res Publica*, Volume 16, pp. 1-22.
- Donaldson, S. & Kymlicka, W., 2011. *Zoopolis: A Political Theory of Animal Rights*. Oxford: Oxford University Press.
- Donaldson, S. & Kymlicka, W., 2013. A Defense of Animal Citizens and Sovereigns. *Law, Ethics, and Philosophy*, 1(1), pp. 143-160.
- Ebert, R. & Machan, T. R., 2012. Innocent Threats and the Moral Problem of Carnivorous Animals. *Journal of Applied Philosophy*, 29(2), pp. 146-159.
- Hatley, J., 2004. The Uncanny Goodness of Being Edible to Bears. In: Foltz, V & Frodeman, R, eds. *Rethinking Nature: Essays in Environmental Philosophy*. Bloomington: Indiana University Press, pp. 13-31.
- Hosein, A., 2014. Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?. In: H. Frowe & G. Lang, eds. *How We Fight: Ethics in War*. Oxford: Oxford University Press, pp. 87-103.
- Kaufman, W., 2010. Self-Defense, Innocent Aggressors, and the Duty of Martyrdom. *Pacific Philosophical Quarterly*, Volume 91, pp. 78-96.
- Mapel, D. R., 2010. Moral Liability to Defensive Killing and Symmetrical Self-Defence. *The Journal of Political Philosophy*, 18(2), pp. 198-217.
- McMahan, J., forthcoming. The Moral Problem of Predation. In: A. Chignell, T. Cuneo & M. Halteman, eds. *Philosophy Comes to Dinner*. London: Routledge, pp. 1-21.
- McMahan, J., 2014. Self-Defence Against Justified Threateners. In: H. Frowe & G. Lang, eds. *How We Fight: Ethics in War*. Oxford: Oxford University Press, pp. 104-137.
- McMahan, J., 1994. Self-Defence and the Problem of the Innocent Attacker. *Ethics*, 104(2), pp. 252-290.

Otsuka, M., 1994. Killing the Innocent in Self-Defense. *Philosophy and Public Affairs*, 23(1), pp. 74-94.

Quong, J., 2009. Killing in Self-Defense. *Ethics*, 119(3), pp. 507-537.

Schlosberg, D., 2007. *Defining Environmental Justice: Theories, Movements and Nature*. New York: Oxford University Press.

Simmons, A., 2009. Animals, Predators, The Right to Life, and the Duty to Save Lives. *Ethics and the Environment*, 14(1), pp. 15-27.

Steinhoff, U., 2016 The Liability of Justified Attackers. *Ethical Theory and Moral Practice* 19, pp. 1015–1030.

Thomson, J. J., 1990. *The Realm of Rights*. Cambridge: Harvard University Press.

Thomson, J. J., 1991. Self-Defense. *Philosophy and Public Affairs*, 20(4), pp. 283-310.

Chapter 2

Regan and the Problem of the Innocent Attacker: Having One's Prey and Eating It Too

Introduction

In the previous chapter I argued that defending Regan's account of animal rights from the predation argument – which holds that if animals have a right not to be killed and eaten by us, then prey animals have a right not to be killed and eaten by predators, which creates an unappealing obligation to police nature - reveals a deeper worry for Regan in how he deals with the permissibility of defending oneself, or another, from an animal attack in cases like the following:

Bear Attack: Tom is out for his morning walk and comes across a bear charging towards a stranger, Peter, in an attempt to kill and eat him. Tom has a rifle with him capable of killing a bear and has no other means of preventing the bear from attacking Peter.

I argued that Regan is faced with a charge of inconsistency and he must discard at least one of his prior commitments to remain consistent. Given that the options of rejecting either the Agency Requirement (moral patients cannot violate the rights of others: agency is necessary for violation) or the Permissibility Claim (it is permissible to kill moral patients in self-defence) seem similarly dire for Regan's account, perhaps what we need to do is to find a way of keeping both of these commitments in such a manner that Regan can avoid a charge of inconsistency.

In what follows, I explore one such option: classifying cases like *Bear Attack* and *Innocent Threat* as what McMahan refers to as symmetrical cases of self-defence. Here is *Innocent Threat* stated again:

Innocent Threat: You are lounging on a deck chair when a gust of wind blows a fat man off a cliff above you. He will certainly kill you if he lands on you, although he will survive. Your only option to save yourself is to move your awning in order to deflect the falling fat man away from you onto some rocks, but doing so will certainly kill him (Thomson 1991, p. 287).

If we can classify such cases as symmetrical cases of self-defence, then Regan could claim that Peter has a special agent-relative permission to use defensive force against the bear, but that the bear is also able to resist Peter's defensive actions with defensive force of its own.¹² While McMahan maintains that such cases are asymmetrical cases of self-defence, I argue that by exploiting an apparent gap between McMahan's conditions for when an attacker is liable to defensive force and his conditions for justified self-defence, we can establish that some innocent attacker and threat cases are in fact symmetrical cases of self-defence. However, I present a problem for this view with regard to third-party permissions and conclude that even supposing that *Bear Attack* is such a case, this still does not help Regan to regain consistency.

I will proceed as follows: the first section will set up the argument by presenting McMahan's 'justification defeats liability' (JDL) principle and explaining its significance to symmetrical cases of self-defence. In the second section, I discuss two recent challenges to this principle, before demonstrating in the third section how we can frame other innocent threat cases as symmetrical cases of self-defence by exploiting a gap in McMahan's account.

1. *Justification Defeats Liability*

We can take note of three necessary conditions on McMahan's account for liability to defensive killing. The threat must be 'morally responsible' for wronging another; 'causally

¹² Of course, a bear is going to do what a bear is going to do, and so while it may not make sense to talk about the bear having a *permission* to use defensive force, the significance of framing *Bear Attack* as a symmetrical case of self-defence is that we can endow *Peter* with a permission to use defensive force.

responsible', in the sense that he is causally connected to the harmful threat in the right kind of way; and the threat he poses must be 'objectively unjust' in that there is no objective overall justification for the threat posed (2005, pp. 394-6). Only when an attacker meets all three of these conditions is he liable to defensive killing. Neither the fat man nor the bear, however, meet the moral responsibility condition and are thus not liable to be killed by you when they threaten your life.

David Mapel (2010), though, has noted that there appears to be a gap between McMahan's conditions for when an attacker is liable to defensive killing and when self-defence is justified, unlike rights-based accounts of self-defence where what often justifies your defensive action is the fact that the attacker has threatened to violate your rights and has thus caused herself to lack those same rights, or to become liable to defensive harm. In other words, McMahan's conditions for liability do not seem to be necessary for justified self-defence and, as such, Mapel thinks that we are left with the "problem of determining when and how liability to defensive killing is a morally significant consideration" (2010, p. 200). To see this, consider the following case presented by McMahan.

Tactical Bomber: A pilot fighting in a just war attacks a military facility, foreseeing that the explosion will unavoidably kill five innocent civilians as a side-effect of his overall justified action which will save 100 people. As he approaches, the civilians learn of his mission. They cannot flee in time but they have access to an anti-aircraft gun (2005, p. 389).

McMahan contends that the pilot may permissibly kill the civilians even though they are not liable to be killed and have a right not to be killed by him, because his justification is sufficient to override the rights of the civilians. Thus, we can say that the tactical bomber infringes the rights of the civilians rather than violates them. McMahan also holds that the civilians are not liable because they are innocent of any wrongdoing and are not required to lay down their lives for the success of the bomber's mission. So far, so plausible. However,

because the civilians are not liable to be killed by the pilot, McMahan thinks that “it is hard to believe that the civilians may not kill the tactical bomber in self-defence” (2005, p. 389). The pilot, though, does not meet all of McMahan’s conditions for liability to defensive killing as he has an overall justification for killing the civilians. As such, he is not liable to be killed by them and he too retains his right not to be killed. Thus, on this account, the pilot does not make himself liable to be killed in attacking the civilians even though they are not liable to be killed by him because “justification defeats liability” (2005, p. 399). As the bomber is not liable to be killed by the civilians, he will be wronged by their defensive attack and is thus permitted to attack the civilians in pre-emptive self-defence.

The following question presents itself here: why is it that the civilians are permitted to kill the pilot in self-defence in the first place if he is not liable to be killed by them? The bomber is permitted to kill them because he has an objective, all things considered justification for doing so as a side-effect of his mission. The civilians, however, have no such objective justification; indeed, by hypothesis, if they kill the pilot in self-defence they will make things objectively worse given that the pilot’s mission is of sufficient importance to justify his killing them. On McMahan’s account, though, the civilians have a special form of justification for killing which permits people a “necessary and proportionate defence against being wronged either by unjustified or *justified* action” (2005, p. 399). Thus, *Tactical Bomber* is an example of a case of *symmetrical* self-defence. As such, it seems that the conditions for liability to defensive killing, on this account, come apart from the conditions for justified self-defence. Instead, what justifies defensive force is being *wronged*, whether by unjustified *or* justified action.¹³ Unlike Thomson’s account, then, this account shifts the focus away from some relevant factor about the threat, which makes him liable to be killed in self-defence, to some

¹³ McMahan (2014) attempts to close this gap by removing the defensive permissions of the civilians in later work. I discuss this more below.

relevant factor about the victim, which makes it the case that she is justified in killing in self-defence.¹⁴ As such, Mapel thinks that the importance of McMahan's conditions for when an attacker is liable to defensive killing lies not in justifying the primary defensive action of the victim, but rather in "justifying the secondary, responsive self-defence of the attacker or threat" (2010, p. 200).

2. *Closing the Gap*

Before moving on to look at how this gap can be exploited, and given the importance that symmetrical cases of self-defence play in the arguments which follow, it will be prudent to first address two recent challenges to McMahan's JDL principle which, if successful, would in effect close this gap. For, if we are to employ this principle so as to provide Regan with some reprieve from the dilemma presented in the previous chapter by framing *Bear Attack* as a case of symmetrical self-defence then we should ensure that it is fit for purpose.

The first challenge presented by Steinhoff (2008: 2012) argues that there is a moral asymmetry between the bomber and the villagers. He thinks that this becomes apparent when we look at the compensation that is owed to each of the parties in the Tactical Bomber example. Steinhoff argues that the relatives of the innocent villagers who are killed by the bombers attack can demand compensation from the bomber, or at least from the nation that issued the bomber's orders. Should the villagers manage to save themselves by killing the bomber before he can drop his bombs however, the relatives of the bomber would not be able to demand compensation from the villagers. This difference in liability to compensation demonstrates that there is a real moral asymmetry between the bomber and the villagers and is the reason,

¹⁴ Here, and throughout this chapter, I refer to Thomson's work on self-defence in *The Realm of Rights* (1990). Thomson puts forward a notably different account of self-defence in her other work. For this view see Thomson (1991).

Steinhoff contends, that the bomber is liable to be killed by the villagers but that the villagers are not liable to be killed by the bomber (2008, p. 223-4).

We might wonder though, if it immediately follows from the fact that the bomber is liable to compensation, that he is then liable to defensive killing? McMahan (2014) holds that there is a difference between liability to compensation and liability to defensive harm. It would seem that McMahan can accept that this difference in liability to compensation does demonstrate a moral asymmetry between the villagers and the bombers, without accepting the further claim that this moral asymmetry stretches further than being liable to compensate. That is, he does not have to accept that there is a similar moral asymmetry when it comes to liability to defensive harm. Consider the following case that McMahan puts forward to demonstrate this difference between liability to compensate and liability to defensive harm:

The Diabetic: Tom comes across a stranger who, through no fault or negligence of her own, has just slipped into a diabetic coma. The only supply of insulin that Tom can get to before the diabetic dies is located in a nearby house. The owner of the house is a stranger to both Tom and the diabetic, and is not at home. Tom breaks into the house to get the insulin and save the diabetic (2014, p. 114).

Although it may be true that Tom is liable to pay compensation to the owner of the house, both for the insulin taken and for the damage caused when breaking in, McMahan argues that it does not follow from this that Tom is liable to defensive harm. For, given that Tom acts with a moral justification which allows him to infringe the owner's property rights, had the owner been at home at the time and not consented to Tom taking the insulin, he would not have been permitted to use defensive force against Tom to prevent him from taking the insulin to save the

diabetic.¹⁵ And this, McMahan thinks, is the main difference between liability to compensation and liability to defensive harm, as to hold someone liable to the first type of liability is not to say that anyone is permitted to prevent his justified action, whereas to hold someone liable to the second type of liability *is* to say that others are permitted to prevent the action (2014, p. 119). Thus, even if Steinhoff is correct and there is a moral asymmetry between justified threateners and their victims in one area, this does not show that there is a moral asymmetry between them in all areas.¹⁶

The second challenge to McMahan's JDL principle also argues that there is a moral asymmetry between justified threateners and their victims. Adam Hosein (2014) puts forward a Thomsonian account of liability to defensive harm, extending it to cover both rights violations *and* infringements. He argues that an individual becomes liable to be harmed by threatening to wrong someone, and an individual threatens to wrong someone when they threaten to violate *or* infringe that person's rights. Furthermore, Hosein maintains that individuals are justified in

¹⁵ One possible difference between the cases, however, is that the owner himself would have been under a duty to provide the insulin had he been in a position to do so. But the civilians are under no such duty to sacrifice themselves in order to expedite the bomber's mission. I leave this aside for now but return to it below.

¹⁶ However, we might also dispute that any moral asymmetry exists between the threatener and the liability to compensation. McMahan (2014) notes that even if we do agree that compensation is owed in certain cases, it does not follow that anyone is liable to provide such compensation. Furthermore, it does not follow that it is the justified threatener who owes compensation to the owner. It hardly seems fair to think that the costs of the diabetic's rescue should fall on Tom, who has already devoted his time and energy to the rescue of a stranger. We might instead think that the costs should be borne by the individual who has derived the benefits from the rescue. McMahan notes that this isn't ideal, but this is still a better outcome for the diabetic than had she not been rescued and it is still fairer than having the costs go to Tom or to the owner of the house. Or we might think that, had the owner been home at the time, then he would have been morally obligated to have offered up the insulin and, as such, Tom was merely fulfilling another person's enforceable duty (however we might think that the owner is still owed compensation for the damage caused by breaking into the house). Perhaps a fairer still way to distribute the costs would be some form of social cooperation scheme where everyone in society is required to contribute to compensating people who suffer a loss through sheer bad luck. Also, even if we think that the costs of the rescue should fall on the diabetic alone, as she has derived the benefits from it, we might still think that while she owes compensation to the owner, she is not liable to compensate the owner. That is, given that the diabetic was in a coma, something for which she was neither at fault for nor negligent for, when Tom decided to break into the house, it seems that the diabetic can have done nothing at all which could have made her liable to pay compensation. In any case, even if it was the case that the diabetic had done something to make herself liable to compensation, for example if she had broken into the insulin owner's house herself, we still would not think that she had made herself liable to defensive harm.

defending themselves to prevent someone from wronging them or others. Thus, he argues that the tactical bomber becomes liable to defensive harm from the citizens when he threatens to wrong them by infringing their rights, thus closing the gap between liability to defensive harm and justified self-defence.

However, before engaging with Hosein's arguments in detail, we should consider the following: if the bomber is in fact liable to be harmed by the citizens then he can have no permission to defend himself against those citizens when they use defensive force against him. For, as McMahan (2014) has argued, if an individual is liable to be harmed by someone for a particular reason, she can then have no right to defend herself against being harmed by that person for that reason. Yet, Hosein holds that if the bomber's justification is strong enough to justify his killing of the citizens so that he might successfully complete his mission and, by hypothesis, it is, then it is surely strong enough to justify his killing them in self-defence in order that he may accomplish his mission. Of course, this justification will only hold so long as it is necessary for the bomber to kill the citizens in order to carry out his mission. As such, if the citizens were to attack the bomber after he has successfully completed his mission, he could have no grounds for self-defence.¹⁷ But this, Hosein argues, is exactly what we should expect from a theory of self-defence (2014, pp. 99-100).¹⁸

McMahan's JDL principle does, however, have considerable intuitive force. After all, if there are strong moral reasons for your action, it seems implausible to think that you still leave yourself open to losing some of your rights as a consequence of performing this morally good

¹⁷ While intuitively plausible, this view has been challenged. See Frowe (2018) for one recent example of this.

¹⁸ One problem worth highlighting here is that if the bomber is liable to be killed in order to save the citizens, then what is to stop a disinterested bystander from permissibly killing the bomber in defence of the citizens? For, if the bomber is liable to be harmed because he threatens to wrong the citizens by infringing their rights, then the citizens would appear to have an agent-neutral permission to use defensive force against him. I set this problem of third-party permissions aside for now but return to it in the following section.

action. Liability to harm seems like a substantial cost to pay for performing a morally good action. Thus, given the intuitive attractiveness of this principle, Hosein needs to demonstrate why we should think that it is false. He attempts to do this by presenting the following justified threatener case.

Car Crusher: A car crusher machine malfunctions and Oliver, who is trapped in it, will be crushed and killed if the machine is not switched off. Malcolm nearby sees all this and realises that the only way to prevent Oliver from being crushed is to place Nicola's arm into the gears. This will cause Nicola's arm to be permanently damaged (2014, pp. 92-3).

Hosein thinks that even though Malcolm will be infringing Nicola's right not to have her arm badly damaged, it is all the same permissible for him to do so given the benefit that will be produced. He thus thinks that Malcolm, just like the tactical bomber, is a *justified* threatener. Suppose now, that Malcolm can save Oliver by placing his own arm into the machine. Hosein thinks that, intuitively, if Malcolm decides to save Oliver's life when he is not morally required to do so, and he can bear the costs of doing so himself, then he *must* bear them rather than transferring them to Nicola. Hosein thinks that what explains this intuition is the fact that it would be more *just* for Malcolm to bear the costs than an impartial bystander as it would involve less "*wronging*". Malcolm, he argues, would wrong Nicola by forcing her to bear the costs, whereas he would wrong no one by bearing them himself (2014, p. 93). Therefore, *justice* requires that Malcolm puts his own arm into the gears rather than Nicola's. Hosein calls this "the principle of just beneficence". This principle says that while it is permissible for Malcolm to decline to save Oliver by doing nothing, if he does decide to save him he "must do the beneficent action that best conforms to the requirements of justice" (2014, p. 95).

Suppose that, all the same, Malcolm attempts to place Nicola's arm into the machine. Hosein thinks that it is permissible for Nicola to permissibly defend herself by placing Malcolm's arm into the machine and transferring the harm from herself to him. Moreover, he

thinks that it is *not* permissible for Malcolm to prevent Nicola from doing so. Therefore, the presence of the option to save Oliver by harming himself changes the permissions operating in this case. That is, the fact that Malcolm has an option to save Oliver without wronging someone makes the option to save Oliver by wronging Nicola morally impermissible, even though it would have been permissible had it been the only way to save Oliver's life.

Hosein argues that as there is a moral asymmetry between Malcolm and Nicola, this case provides a counter-example to the JDL principle as it is clear that justification does not defeat liability here. The most plausible explanation for this, he maintains, is that Malcolm has made himself liable to be harmed by Nicola by threatening to wrong her, even though he was justified in doing so. This demonstrates, Hosein argues, a moral asymmetry between justified threateners and their victims in general. This moral asymmetry is one of liability to defensive harm. Justified threateners, then, become liable to defensive harm when they threaten to wrong their victims while carrying out their justified actions and, as such, their victims are permitted to use defensive force against them. Given that justified threateners are liable in this way, they are not permitted to defend themselves against their victim's defensive actions. Therefore, Hosein continues, the tactical bomber also becomes liable to be harmed by the citizens in the course of threatening to harm them.

There appear to be two promising routes to defend the JDL principle here: either by showing that justification does in fact defeat liability in this case, or by arguing that this case is significantly different from other justified threatener cases so that the fact that liability does not defeat justification in this case does not entail that justification does not defeat liability in others. I agree with Hosein that in *Car Crusher*, Malcolm is in fact liable to defensive harm and so I focus on the latter strategy. There are two relevant differences, I think, between this case and *Tactical Bomber*. I discuss them in turn.

The first is one that Hosein recognises. This is that in *Car Crusher*, the victim can defend herself against the justified threatener without preventing the benefit from accruing, whereas in *Tactical Bomber*, it is not possible for the citizens to defend themselves without thereby thwarting the morally good action. Indeed, if the citizens defend themselves from the bomber, by hypothesis, they effectively condemn the 100 people to death. As such, it might be possible to modify the JDL principle so that even though Tom becomes liable in the *Car Crusher* case, justified threateners like that of McMahan's tactical bomber do not. Hosein thinks that such a revised principle would take the following form:

(JDL*): An attacker who imposes a threat of harm that will wrong another person does not thereby become liable to harm if (a) the threat he poses is justified and (b) harming him would prevent him from doing the good he would otherwise do (2014, p. 96).

Hosein, however, thinks that this revised principle is ad hoc and runs into the following problem. If harming Malcolm would cause Oliver not to be saved, then this revised principle suggests that Malcolm would not then be liable to defensive harm. Thus, whether Malcolm has a right not to be harmed depends on facts about a third person, Oliver, rather than on facts about Malcolm. That is, the revised JDL principle entails that whether Malcolm has a right not to be harmed depends on whether it would be bad for Oliver if Malcolm was harmed. This, Hosein thinks, is implausible as it conflicts with fundamental ideas about rights; specifically, "whether someone has a basic right in the first place seems to depend only on facts about that person" (2014, p. 96-7). As such, rather than giving up on some of our fundamental ideas, he holds that we should abandon JDL* instead.

Is it as damaging as Hosein suggests for an account to hold that an individual's permission for self-defence can be determined, in certain cases, by whether this would be bad for a third person? Hosen argues that "whether a basic right may be infringed depends on facts about

people other than the rights holder. My point is just that whether someone has a basic right in the first place seems to depend only on facts about that person” (2014, p. 97). But we might think that there is an additional factor here, namely, whether someone lacks a specific right against a specific person might depend on certain facts about other people. We might understand liability to be about lacking a right against a specific person for a specific reason and not about whether a person has a basic right or not. Thus, we might think that Malcolm retains his basic general right not to be harmed in the *Car Crusher* case, but that he loses his specific right not to be harmed by a particular person, Nicola, in this particular instance. It seems to make sense that whether or not he loses his right to be harmed by Nicola depends on facts about her rather than about Malcolm. But, if facts about someone other than Malcolm can have a bearing on whether or not he lacks a right not to be harmed then why not think that facts about Oliver can influence whether Malcolm lacks a right not to be harmed in this particular instance? Indeed, it seems likely that whether the tactical bomber has the right to self-defence against the villagers seems to be partly determined by facts about third-parties. That is, as Hosein himself admits, whether the bomber has the right to defend himself against the villagers’ attack turns on whether the 100 people have been saved or not. If they have not yet been saved then the bomber may permissibly defend himself against the villagers’ defensive force. If the people have been saved, however, then it is no longer permissible for the bomber to defend himself against the villagers. If this is right then it would appear that Hosein cannot reject JDL* as false on these grounds.

Perhaps, however, Hosein would reply that because the tactical bomber is liable, it is not the case that he has a right to self-defence against the villagers at any stage. What he had instead, was an impartial justification for killing the villagers in order that he might successfully carry out his mission. This justification allows him to pre-emptively kill the villagers to ensure the success of his mission. It is not the case, then, that he has a right to self-defence when the

100 people still need to be saved and then loses this right whenever the 100 people have been saved. Rather, he has a justification for harming the villagers when the 100 people need to be saved, and he then loses this justification after they have been saved.

So much for the first difference. There is, however, a further problem with Hosein's argument which suggests that even if justified threateners like Malcolm do become liable, those like the tactical bomber do not. This is highlighted by another morally salient difference between the *Car Crusher* case and the *Tactical Bomber* case. This is that, in the latter case, it is not possible for the bomber to save the 100 people by absorbing the costs himself rather than inflicting them on the citizens. In *Car Crusher*, however, Malcolm *can* save Oliver by absorbing the costs himself rather than inflicting them on Nicola. In what follows, I argue that this difference is enough to undermine Hosein's argument that his *Car Crusher* case serves as a counter-example to the JDL principle for it suggests that Malcolm is not in fact a *justified* threatener at all.

To see this, first recall that Hosein thinks that the tactical bomber is permitted to repel the citizens' defensive actions just so long as he retains his justification for carrying out his mission, that is, so long as the 100 people still need to be saved. If the bomber is permitted to defend himself against the citizens' defensive actions because his mission is of sufficient importance to justify him doing so, then we should also hold, on pain of inconsistency, that in the case where Malcolm can only save Oliver by harming Nicola that he is likewise permitted to defend himself against Nicola's defensive actions just so long as it is necessary to do so to save Oliver.

To see this, suppose that the only way for Malcolm to save Oliver's life is to put a crowbar into the machine, which will foreseeably send gear debris shooting out which will, in turn, badly damage Nicola's arm. It seems that if Malcolm's actions are justified in such a case, then although Nicola can permissibly pre-emptively defend herself, Malcolm can permissibly

repel such an action. For, just like in the tactical bomber case, if Malcolm's justification is strong enough to justify his harming Nicola in the first instance, then it is surely strong enough to permit him to defend himself so that he can successfully save Oliver. If, however, Nicola can hold up a shield, thus deflecting the debris away from her own arm and towards Malcolm's instead, after Oliver has already been saved, it would no longer be permissible for Malcolm to defend himself (by harming Nicola) as he no longer has a justification for doing so. Similarly, in the *Tactical Bomber* case, Hosein thinks that after the bomber has completed his mission, if the citizens could then shoot down his plane in such a way as to use it as a shield against the approaching fire, then he would not be permitted to defend himself against such actions, given that he no longer possesses a justification to do so.

It is clear that in this modified version of the *Car Crusher* case, there is no asymmetry between justified threateners - like Malcolm and the tactical bomber - and their victims. Justified threateners, then, do not become liable to defensive harm in such cases. It seems then, that the fact that Malcolm can bear the costs of saving Oliver himself, in the original *Car Crusher* case, changes the permissions operating around self-defence. This suggests that even if Hosein is right that there is a moral asymmetry in liability between justified threateners and their victims in cases like *Car Crusher*, this does not show that there is a moral asymmetry between justified threateners and their victims in cases like *Tactical Bomber*. And if this is right, then Hosein has not conclusively shown that JDL is false.

Given this, I think that we can rescue JDL by modifying it further to account for the difference suggested above in the following way:

JDL**: An attacker who imposes a threat of harm that will wrong another person does not thereby become liable to harm if (a) the threat he poses is justified, (b) harming him

would prevent him from doing the good he would otherwise do, *and* (c) it is not possible for him to absorb the costs himself rather than inflict them on his victim(s).¹⁹

Under this revision, then, JDL** suggests the following: if the facts of the Tactical Bomber case were instead that the bomber had a choice of carrying out his mission by bombing his target and killing the five citizens, or accomplishing his mission without threatening the nearby citizens by flying his plane directly into the target, then if he chose the former option he could not claim to be shielded from liability and would not be permitted to repel the citizens' defensive action. This is not an ad hoc addition to the JDL principle because, as Hosein himself has claimed, to harm the citizens in such a case would be to perform an *additional* wrong on top of the wrong of infringing their right not to be killed. For, given that the tactical bomber can absorb the costs of saving the 100 people himself without transferring these to the citizens, if he were to choose to do otherwise then he would be violating Hosein's "principle of just beneficence" (2014, p. 94). This principle dictates that, while it is permissible for the bomber to decline to save the 100 people by doing nothing, if he decides to be beneficent and save them, he must do so in the manner that justice requires, namely, by absorbing the costs himself. For, if he carries out his mission by dropping his bomb and killing the citizens, he has wronged five people, whereas if he completes his mission by flying his plane directly into the target he has wronged no one. Therefore, he should do the action which involves the least amount of wrong. As Hosein claims, all else being equal, we expect "that people should bear the costs of their own beneficence rather than displacing these costs onto others" (2014, p. 95). Similarly, in the *Car Crusher* case, it would be more just for Malcolm to place his own arm into the machine rather than to transfer the costs of his beneficence to Nicola. In this case, where Malcolm can absorb the costs, he wrongs Nicola in a different way than the way in which he

¹⁹ With the exception of (c), this formulation is Hosein's (2014, p. 96).

wrongs her when the only way that he can save Oliver is by putting her arm into the machine. Specifically, he wrongs her by doing something which does not conform to justice.

It is this fact which highlights that the difference between how Malcolm wrongs Nicola in the original case, when damaging her arm is the only way to save Peter, and how he wrongs her in the revised case, is that the former wrong is a permissible one, whereas the latter wrong is an impermissible one. This suggests that Malcolm no longer acts with justification in the revised case. As such, in cases where the threatener can bear the burdens of saving others himself, rather than placing them on others, then the permissions surrounding self-defence change entirely. That is, in the revised case, Malcolm cannot rightly be considered as a *justified* threatener at all. Therefore, we should challenge the stipulation in Hosein's revised *Car Crusher* example that Malcolm's action is justified in the first place. Given the presence of the option to absorb the costs of saving Oliver himself, it is clear that Malcolm's actions are no longer overall justified. If so, then the reason that Malcolm becomes liable to defensive force from Nicola is not that his justification fails to defeat liability as Hosein argues, but rather that his actions are not justified in the first place.

If this right, then *Car Crusher* is not a counter-example to the JDL principle because it is not a case where a justified threatener becomes liable despite having a justification for his actions. A requirement for the justification of any defensive action is that it fulfil a necessity condition and it is clear that if Malcolm can save Oliver by placing his own arm into the machine, then it is not necessary to harm Nicola. Indeed, with this in mind, it might well be the case that condition (c) is an unnecessary addition to the formulation of JDL as this simply is the necessity condition for *justified* self-defence (or in this case, other-defence). That is, for any action to be considered justified which will secure a good by harming innocent others who are not liable to be harmed, it must be the case that it is necessary to perform that action rather than another action which will involve less wronging. If it is the case that the action will

involve *more* wrong than another action which will achieve the same benefit, then that former action lacks justification. In the original *Car Crusher* case, what makes it an *infringement* of the Nicola's rights rather than a *violation*, is that there exists no alternative way for Malcolm to save Oliver which involves less wrongdoing.²⁰

Hosein might reply here that, as he has presented *Car Crusher*, Malcolm's action is, by hypothesis, permissible and, even if we disagree with his intuition that it is permissible for Malcolm to place Nicola's arm in the machine, we can reduce the harm done to Nicola "until it becomes a justified action" (2014, p. 98). So, for example, if Malcolm can save Oliver's life merely by pinching Nicola's finger, we might think that he is permitted to do so.²¹ However, even in this case, if it is possible for Malcolm to absorb the costs himself, by pinching his own finger say, then it is simply not the case that his action is justified. For Hosein's argument against JDL to be successful, he needs to maintain that Nicola is permitted to defend herself against Malcolm's actions but that Malcolm is not permitted to repel Nicola's defensive force. To summarise, Hosein is right that what best explains the asymmetry in his *Car Crusher* case is the fact that Malcolm has become liable to be harmed by Nicola. However, he is mistaken to think that this is an asymmetry which exists between justified threateners and their victims. Rather, it is the very familiar asymmetry which exists between *unjust* threateners and their victims.

²⁰ A further relevant difference between *Car Crusher* and *Tactical Bomber* is that Malcolm seems to use Nicola merely as a *means* of saving Oliver, whereas the tactical bomber only kills the citizens as a side effect of saving the 100 people. Indeed, it would be much better for the bomber if the citizens were not present at all. It would not, however, be better for Malcolm if Nicola were not present because he would then be unable to save Oliver's life. It is not obvious that it is permissible for the bomber to defend himself against the citizens' defensive actions in a revised case where he uses them to complete his mission. For example, see McMahan's 'Terror Bomber' case (1994) where the bomber kills the citizens directly to inspire fear and terror in the hope of forcing the enemy government to surrender thus saving more lives overall. In this case also, there is good reason to think that the bomber is not permitted to defend himself against the citizens because he lacks justification for his actions. For a real-life case of this see Malcolm Gladwell's *Revisionist History* episodes on the life of Air Force general Curtis LeMay.

²¹ We might think that this looks like a cost of rescue which could be plausibly imposed on Nicola: Malcolm would merely be enforcing a duty she already had. However, it is unclear why we should think that this duty would fall on Nicola here, rather than Malcolm.

3. *Exploiting the Gap*

With this challenge to the JDL principle now addressed, and the gap secured, we can now return to McMahan's Responsibility Account. Recall that, on this account, the conditions for liability to defensive harm, seem to come apart from the conditions for justified self-defence and, thus, the importance of McMahan's conditions for when an attacker is liable to defensive killing seems to lie not in justifying the primary defensive action of the victim, but rather in justifying the secondary, responsive self-defence of the attacker or threat. With this in mind, we might legitimately wonder whether the civilians, on this account, actually make *themselves* liable to be attacked by the pilot when they attack him in self-defence despite the fact that their defensive action is justified. For consider the following. The civilians are not liable to be killed by the pilot, but the pilot does not become liable to be killed when he attacks them because justification defeats liability on the Responsibility Account. The primary defensive action of the civilians is justified by the fact that they would be wronged by the pilot if he killed them in completing his mission because they are not liable. As such, they can attack the pilot even though he is also not liable to be killed by them. This is what makes *their* defensive action an infringement rather than a violation of the pilot's rights. Similarly, as the bomber is not liable to be killed and will be wronged if the civilians kill him, his secondary self-defence is justified. The question then is whether the civilians' primary defensive action now makes them liable to defensive killing, an issue which is separate from whether their defensive action is justified or not. The civilians, it would seem, now meet all three of McMahan's conditions for liability to defensive killing (that is, they are morally responsible, causally responsible, and

they lack an objective justification for the threat they pose).²² It is true that they would be wronged by the pilot if he were to kill them in carrying out his mission, but this only provides them with the justification for killing the pilot in self-defence, and is not relevant to whether or not they make themselves liable to defensive killing.

This question about the civilians' liability to defensive killing may seem relatively trivial as the bomber already has his justification for killing the civilians in self-defence regardless of whether they are liable to defensive killing, on the basis of being wronged by them. That is, if the civilians are not liable, the bomber is still justified in infringing their rights and killing them in self-defence. McMahan, though, thinks that in symmetrical cases of self-defence, such as this one, *other*-defence is not permitted because both parties are not liable to be killed and only have a special agent-relative justification for killing the other in self-defence, a justification which does not extend to third-parties. However, should it be the case that the civilians, in attacking the pilot in self-defence, have now made themselves liable to defensive killing, disinterested third-parties *would* be justified in killing them in defence of the bomber. This is because the reasons for using defensive force against them now appear to be agent-neutral. So, while it may be true that that the civilians cannot be killed in other-defence of the pilot for the purpose of ensuring that he completes his mission, because they are not liable to be killed for this purpose, they *can* be killed in other-defence for the purpose of saving the pilot from their defensive attack as there seems no reason to think that the civilians have not now made themselves liable to defensive killing, at least under the conditions that McMahan has provided.

²² Admittedly, this last condition is a little murky in that defensive action undertaken from individuals who are not liable to be attacked does not seem to be straightforwardly objectively unjust. However, both the fact that the civilians have only agent-relative reasons for defending themselves against the bomber and that they will make things objectively worse overall by defending themselves highlight, I think, their lack of an objective justification to use defensive force against the bomber.

Furthermore, because the civilians have now made themselves liable to defensive killing, they would have no justification for defending themselves against a third-party.²³

McMahan, then, in order to avoid this highly unintuitive outcome, would need to claim that justification is not the only thing that defeats liability; being wronged also negates liability. That is, just as McMahan holds that *being wronged* provides the civilians with a justification for defensive force, to avoid the above unintuitive conclusion, he will also need to hold that being wronged provides the civilians with the necessary protection against liability to defensive harm. If so, we might think that there are other things which defeat liability and thus other cases which qualify as cases of symmetrical self-defence. Perhaps the following example presented by McMahan is also such a case:

Mobile Phone Operator: A Villain tampers with a bystander's phone so that if he answers a call it will trigger a bomb tied to your chest. The villain then calls the bystander. The only way you can save yourself is by killing the bystander (2005, p. 397).

McMahan argues that although the mobile phone operator "poses an unjust threat" to a non-liable victim when he answers the call and is "a morally responsible agent whose pressing the button would be a free, voluntary act", the operator is not responsible for killing you (2005, p. 397). This is because a necessary condition of moral responsibility is that of 'foreseeability';

²³ See McMahan (2014, p.129-130) for an argument that withdraws the permission for the civilians to defend themselves against the bomber. Here he argues that the permission of self-defence in such a case would imply that "acts of self-preservation can be permissible when they kill as a side-effect, more nonliable people than they save". Despite the arguments put forward above in favour of the civilian's permission for defensive force, it seems independently unintuitive to hold that the citizens are unable to defend themselves against the bomber. In any case, McMahan is clear that he still thinks that cases of symmetrical self-defence do exist. See McMahan (2014) for his examples of such cases. As such, even if his later position is correct and *Tactical Bomber* is not a case of symmetrical self-defence, the fact that such cases exist is enough motivation for attempting to regain consistency within Regan's account by framing *Bear Attack* as such a case. Given this, and how unintuitive it is to think that the civilians are unable to defend themselves, I focus primarily on McMahan's earlier account in this chapter.

that is, to be morally responsible for a threat, you must have a reasonable possibility of foreseeing that the action you are engaging in carries “a non-negligible risk of causing a significant unjust harm” (2005, p. 397). As the mobile phone operator lacks this foreseeability condition and thus lacks moral responsibility for his action, he is not liable to be killed by you in self-defence. As such, he can resist your defensive action with his own. Therefore, this case would appear to be another example of a symmetrical self-defence case. In other words, not only does overall justification defeat liability, but so too does lack of moral responsibility

We might then think that the *Innocent Threat* case presented in the previous chapter could also be a case of symmetrical self-defence. Here it is again:

Innocent Threat: You are lounging on a deck chair when a gust of wind blows a fat man off a cliff above you. He will certainly kill you if he lands on you, although he will survive. Your only option to save yourself is to move your awning in order to deflect the falling fat man away from you onto some rocks, but doing so will certainly kill him (Thomson 1991, p. 287).

After all we have already seen that both justification and lack of moral responsibility negate liability, and as the fat man lacks agency when falling it is clear that he also lacks moral responsibility. Thus, if you were to attack him in self-defence, he could defend himself against you because his lack of moral responsibility ensures that he is not liable to defensive killing from you, despite the fact that you are not liable to be killed by him.

So why not consider this a case of symmetrical self-defence?²⁴ The main reason, it would seem, is that McMahan does not think that it is permissible for you to kill the fat man in an act of primary self-defence. Thus, McMahan’s account of symmetrical self-defence has two distinct dimensions: justified self-defence, which focuses on something about the victim that

²⁴ See Steinhoff (2011) for one interesting account which, contrary to McMahan, holds that cases like *Innocent Threat* are in fact cases of symmetrical self-defence, while the *Tactical Bomber* case is not. This account relies on an individual’s right to self-defence transforming from a claim-right into a liberty-right.

makes primary defensive killing permissible, and liability to defensive killing, which focuses on something about the attacker that makes secondary defensive killing permissible. So while *Innocent Threat* satisfies the second dimension of symmetrical self-defence cases, McMahan thinks that it fails the first one because the fat man does not wrong you. But we might wonder why, from the attacker's perspective, lacking one of McMahan's conditions for liability to defensive killing shields you from being *permissibly wronged* by your victim's defensive actions, while lacking another does not, given that lacking any one of them shields you from *liability*. There appears to be a lack of consistency here. To see this consider the following: the bomber does not meet the condition 'lacks an objective overall justification' and so does not make himself liable when he kills the civilians. Similarly, the fat man does not meet the condition 'morally responsible' and also does not make himself liable to defensive killing when he kills you. However, not meeting the 'morally responsible' condition shields the fat man from being permissibly wronged by you in self-defence, whereas not meeting the 'lacks an objective overall justification' condition *does not* shield the bomber from being permissibly wronged by the civilians in self-defence. Why is it that, in McMahan's own words, "acting with justification does not...afford the same degree of immunity as lack of moral responsibility" (2005, p. 402)? Moreover, as Mapel has claimed, there is something strange about McMahan maintaining that self-defence is permissible against the bomber who has the strongest moral justification for harming another, but impermissible against the fat man who has no moral justification at all (2010, 209). It is true that the bomber acts with moral responsibility, while the fat man does not, but the bomber makes up for this fact by having an overall justification for killing the civilians, which the fat man lacks for killing you. This seems like a substantial price to pay for performing what is, by all accounts, a morally good and permissible action.²⁵

²⁵ Again, see McMahan (2014) for an argument that the bomber's justification does in fact shield him from defensive harm.

The reason, it would appear, that the fat man is shielded from being permissibly wronged by his victim is that *Innocent Threat* differs from *Tactical Bomber* in one crucial respect; namely, that the civilians in the latter case are *wronged* when they are killed by the pilot, whereas you are not when you are killed by the fat man. But then, so too, it would seem, does *Mobile Phone Operator* differ in this respect. For how could you be justified in killing the mobile phone operator in self-defence if he is not morally responsible for his action and thus does not wrong you? McMahan says that “it is possible that you could be justified in killing [him], but you could not plausibly claim that he had made himself liable to be killed” (2005, p. 397). If the operator does not make himself liable to be killed by you *and* he does not wrong you, then there must be other factors (besides being wronged) that are relevant to a victim’s justification for killing in self-defence when her attacker is not liable to defensive killing. For instance, Mapel argues that, in some cases, if you are faced with a threat of “unprovoked imminent harm” then you can permissibly wrong that non-liable threat (2010, p. 209). Thus, if we adopt this we can hold that, in both *Mobile Phone Operator* and *Innocent Threat*, you are permitted to use necessary and proportionate force against the threat in question despite the fact that they are not liable to defensive killing because they pose a threat of imminent lethal harm which was not unjustly provoked. Because the fat man and the mobile phone operator are not liable to defensive killing and you will thus wrong them in killing them, they are also permitted to use defensive force to resist your self-defence. This is an intuitively appealing solution as it allows us to say both that you are permitted to defend yourself against an innocent threat, and that because the threat is also innocent, he is permitted to defend himself against you if possible. Therefore, it seems that we can think of *Innocent Threat* as yet another case of symmetrical self-defence along with *Tactical Bomber*.

4. *Bear Attack Revisited*

If this is right, and *Innocent Threat* can be considered to be a case of symmetrical self-defence then perhaps Regan can adopt this view and hold that *Bear Attack* is also such a case, thus allowing him to retain both the Agency Requirement (moral patients cannot violate the rights of others: agency is necessary for violation) and the Permissibility Claim (it is permissible to kill moral patients in self-defence). In *Bear Attack* then, Peter would be permitted to kill the bear in self-defence, but not because the bear has violated Peter's rights and thus made it the case that the bear no longer has the right not to be killed - as this would require moral responsibility on the part of the bear - but rather because Peter is faced with something like an unprovoked, imminent lethal harm and is thus permitted to defend himself against the bear. Given that the bear has not made itself liable to be killed by Peter in attacking him, it too retains its right not to be killed which Peter will therefore infringe and, as such, he will wrong the bear when he attacks it in self-defence. Thus, the bear will be permitted to defend itself against Peter's defensive actions. Of course, given that the bear is not a moral agent and therefore cannot have any duties whatsoever to us or anyone else, it is already the case that the bear has no obligations to forbear from defending itself. The important thing here, though, is that Peter is permitted to attack the bear in self-defence regardless of the fact that the bear is not liable to be attacked by him and retains its right not to be killed by him. Peter merely *infringes*, rather than violates, the bear's right not to be killed. What this also means, though, is that it is not permissible for third-parties, like Tom, to engage in other-defence of Peter in cases like *Bear Attack*. This is because, as we have seen, Peter's permission of self-defence is a special *agent-relative* justification which does not extend to impartial third-parties as both him and the bear are innocent and are not liable to be killed.

That being said, one obstacle for the possibility of Regan adopting this view to regain consistency in cases like *Bear Attack* is the concern, mentioned earlier, that the innocent threat's victim has now made himself liable to defensive killing by attacking the innocent threat

in self-defence, given that on agent-relative accounts of self-defence there appears to be a gap between liability to defensive killing and justified self-defence. I noted earlier with regards to *Tactical Bomber* that the civilians seem to make themselves liable to defensive harm when they use defensive force in response to the bomber's attack but that McMahan might be able to resist this conclusion by holding that 'being wronged' also negates liability as well as having an 'overall justification' and lacking 'moral responsibility'. However, given that Peter has not been wronged, has moral responsibility, and lacks an *objective* justification for attacking the bear, I think we can say that Peter has now made himself liable to defensive killing when he attacks the bear in self-defence. Again, this is not important for the bear's permission to defend itself against Peter as it is not a moral agent and has no duties to us.²⁶ Rather, its importance lies in the third-party permissions for other-defence. While it is true that Tom cannot kill the bear in other-defence of Peter because the bear is not liable to defensive killing, it would seem that, in defending himself against the bear, Peter has now made himself liable to defensive killing for reasons which are agent-neutral. It would appear, therefore, that Tom now has a permission to shoot Peter, in other-defence of the bear, and not vice versa. This is not because Peter is *not* permitted to defend himself against the bear, but because in doing so, Peter makes himself liable to defensive killing, and this allows an impartial third-party to kill Peter in defence of the bear. But this is no solution to Regan's problem at all. Bad enough that an impartial bystander is not permitted to *defend* Peter from a bear attack, but it seems absurd to claim that Tom can kill Peter to prevent Peter from doing something that he was morally permitted to do. Indeed, this type of 'solution' to Regan's inconsistency problem is just as intuitively unappealing as dropping the Permissibility Claim from Regan's account and holding that Peter is not permitted to defend himself at all against an attacking bear. To claim that

²⁶ However, even if it was the case that the bear is a moral agent, the fact that it has not made itself liable to defensive killing ensures that Peter wrongs it when he defensively attacks it and would thus provide the bear with McMahan's special justification for self-defence.

morality does not require you to martyr yourself when a bear attacks you and you have the means to defend yourself against it, but that in defending yourself another person may kill you in defence of the animal that is attempting to kill you in an unprovoked attack, almost seems to border on inconsistency.

5. *Conclusion*

This chapter has considered the possibility of rejecting the dilemma presented for Regan in the previous chapter and attempting to reconcile both of Regan's commitments within his theory by positing innocent threat cases as cases of symmetrical self-defence. After defending McMahan's JDL principle I argued that by exploiting an apparent gap between McMahan's conditions for when an attacker is liable to defensive force and his conditions for justified self-defence, we can put forward a case that some innocent attacker and threat cases are in fact symmetrical cases of self-defence. This would then offer Regan the possibility of claiming that *Bear Attack* is a case of symmetrical self-defence: Peter has a special agent-relative permission to use defensive force against the bear *and* the bear is also able to resist Peter's defensive actions with defensive force of its own. While this initially appears to be a promising avenue for Regan, as it at least allows for self-defence against an attacking animal, it is clear that there remains work to be done with regard to the problems caused by the permissions operating around other-defence.

Works Cited

- Callicott, J. B., 1989. Review of Tom Regan, *The Case for Animal Rights*. *Defense of the Land Ethic: Essays in Environmental Philosophy*, pp. 39-48.
- Cochrane, A., 2013. Cosmozoopolis: The Case Against Group-Differentiated Animal Rights. *Law, Ethics, and Philosophy*, Volume 1, pp. 127-141.
- Cochrane, A., 2012. *Animal Rights Without Liberation: Applied Ethics and Human Obligations*. New York: Columbia University Press.
- Cripps, E., 2010. Saving the Polar Bear, Saving the World: Can the Capabilities Approach do Justice to Humans, Animals and Ecosystems?. *Res Publica*, Volume 16, pp. 1-22.
- Donaldson, S. & Kymlicka, W., 2011. *Zoopolis: A Political Theory of Animal Rights*. Oxford: Oxford University Press.
- Donaldson, S. & Kymlicka, W., 2013. A Defense of Animal Citizens and Sovereigns. *Law, Ethics, and Philosophy*, 1(1), pp. 143-160.
- Ebert, R. & Machan, T. R., 2012. Innocent Threats and the Moral Problem of Carnivorous Animals. *Journal of Applied Philosophy*, 29(2), pp. 146-159.
- Frowe, H., 2018. Lesser Evil Justifications for Harming: Why We're Required to Turn the Trolley'. *Philosophical Quarterly*, 68, 460-80.
- Hatley, J., 2004. The Uncanny Goodness of Being Edible to Bears. In: Foltz, V & Frodeman, R, eds. *Rethinking Nature: Essays in Environmental Philosophy*. Bloomington: Indiana University Press, pp. 13-31.
- Hosein, A., 2014. Are Justified Aggressors a Threat to the Rights Theory of Self-Defense?. In: H. Frowe & G. Lang, eds. *How We Fight: Ethics in War*. Oxford: Oxford University Press, pp. 87-103.
- Kaufman, W., 2010. Self-Defense, Innocent Aggressors, and the Duty of Martyrdom. *Pacific Philosophical Quarterly*, Volume 91, pp. 78-96.
- Mapel, D. R., 2010. Moral Liability to Defensive Killing and Symmetrical Self-Defence. *The Journal of Political Philosophy*, 18(2), pp. 198-217.
- McMahan, J., forthcoming. The Moral Problem of Predation. In: A. Chignell, T. Cuneo & M. Halteman, eds. *Philosophy Comes to Dinner*. London: Routledge, pp. 1-21.
- McMahan, J., 2014. Self-Defence Against Justified Threateners. In: H. Frowe & G. Lang, eds. *How We Fight: Ethics in War*. Oxford: Oxford University Press, pp. 104-137.

McMahan, J., 1994. Self-Defence and the Problem of the Innocent Attacker. *Ethics*, 104(2), pp. 252-290.

Otsuka, M., 1994. Killing the Innocent in Self-Defense. *Philosophy and Public Affairs*, 23(1), pp. 74-94.

Quong, J., 2009. Killing in Self-Defense. *Ethics*, 119(3), pp. 507-537.

Schlosberg, D., 2007. *Defining Environmental Justice: Theories, Movements and Nature*. New York: Oxford University Press.

Simmons, A., 2009. Animals, Predators, The Right to Life, and the Duty to Save Lives. *Ethics and the Environment*, 14(1), pp. 15-27.

Steinbock, U., 2016. The Liability of Justified Attackers. *Ethical Theory and Moral Practice* 19, pp. 1015–1030.

Steinbock, U., 2012. Rights, Liability, and the Moral Equality of Combatants. *The Journal of Ethics* 13, pp. 339-366.

Steinbock, U., 2011. McMahan, Symmetrical Defense, and the Moral Equality of Combatants. *The Journal of Ethics*.

Steinbock, U., 2008. Jeff McMahan on the Moral Equality of Combatants. *The Journal of Political Philosophy* 16, pp. 220-226.

Thomson, J. J., 1991. Self-Defense. *Philosophy and Public Affairs*, 20(4), pp. 283-310.

Thomson, J. J., 1990. *The Realm of Rights*. Cambridge: Harvard University Press.

Chapter 3

A Series of Unfortunate Accounts of Misfortune

Introduction

Many of us share the common-sense intuition that it is worse to perform medical experiments on a severely cognitively impaired human than on a comparably cognitively endowed non-human animal. How do we justify such differential treatment? A common way of doing so is to point to a morally significant property shared by humans but which animals lack. Attempts at picking out such a property that all humans share and no animals possess, though, have rarely been successful.²⁷ Instead, an alternative explanation for the common sense intuition is that it seems plausible to suppose that the cognitively impaired human is unfortunate in a way that similarly endowed animals are not. Here ‘fortune’ is used to refer to how well or badly an individual’s overall life is faring. Thus, an individual is unfortunate when her well-being falls short of an expected level. The comparison class used to determine an individual’s expected level of well-being and, thus, how unfortunate she is, differs from account to account. This chapter will look at two popular accounts of misfortune: the *Species Norm Account* and the *Intrinsic Potential Account*.²⁸

This chapter takes as its starting point the assumption that if an account of misfortune is to be considered successful, it must clear the following two hurdles. Firstly, it should avoid being undermined by well-known counterexamples, namely, Jeff McMahan’s *Superchimp* cases. Secondly, a successful account of misfortune must preserve and explain our widely and

²⁷ See Singer (1993).

²⁸ Although both accounts are credited to Jeff McMahan, only the latter is favoured by him (1996; 2002).

strongly held intuitions. Specifically, we want an account of misfortune to explain the following deeply held intuitions:

Fortunate Non-Humans: My springer spaniel, Rodney, is not unfortunate for not being a person, and;

Unfortunate Humans: Kenneth, a cognitively impaired human, *is* unfortunate for not being a person.

This task, however, is far from straightforward. For, we want our judgements of misfortune to correlate with our verdicts of whether an individual is badly off or not. It will not do, for example, for an account of misfortune to conclude that an individual is unfortunate whose life intuitively appears to be going quite well overall. That is, if we judge an individual to be unfortunate, then we want to be able to say that she is badly off and thus can be compensated accordingly. This can be explained by *Fortunate Non-Humans*: that is, the intuition that it seems counterintuitive to judge a dog who is living a good life to be unfortunate for not having access to the levels of well-being enjoyed by human persons. Thus, it is counterintuitive to judge a dog to be unfortunate when he does not seem to be badly off. The problem, however, is that if we judge a dog not to be unfortunate for not having access to these levels of well-being and then apply the same criteria to the cognitively impaired human, it is difficult to maintain that she is badly off. For, if we only judge how well an individual's life is going by comparing her levels of well-being with those made possible by the individual's actual cognitive endowment, then it is difficult to maintain that a cognitively impaired human - who has high levels of well-being by these limited standards - is not living a good life. It is clear, however, that this is not the result that we require from an account of misfortune. Rather, we want to claim that the cognitively impaired human *is* unfortunate for possessing the natural endowment that she has rather than one which would allow her to access the higher levels of

well-being enjoyed by persons. But if the cognitively impaired are to be considered badly off for possessing the cognitive endowment that they do, then, presumably, non-human animals who share a similar endowment must be badly off in the same way. The difficulty, then, is to find a non-arbitrary way of restricting the scope of our judgements of misfortune, when the individual in question falls short of personhood, to members of the human species.

This chapter suggests that neither account of misfortune is able to fulfil both goals: successfully avoiding being undermined by well-known counterexamples, and explaining our widely held intuitions that non-human animals are *not* unfortunate for not being persons while cognitively impaired humans *are* unfortunate for not being persons. The Species Norm Account is only able to clear the second hurdle by stumbling at the first, while McMahan's favoured account, the Intrinsic Potential Account, clears the first hurdle, only to crash headfirst into the second one. I argue that failing to clear either hurdle is enough to undermine an account of misfortune and, as such, neither account is wholly successful. The chapter will proceed as follows. The Species Norm Account and its associated problems are addressed in section 1, the Intrinsic Potential Account is discussed in section 2, while section 3 offers some remarks on the role of misfortune verdicts.

1. Species Norm Account

Why should we think that a cognitively impaired human is unfortunate for having such limited cognitive capacities, and thus a restricted capacity for well-being, while a similarly endowed animal is not? One obvious way to measure fortune, which seems *prima facie* to adhere with our common-sense intuitions, is to compare how well an individual is faring relative to the levels of well-being that are possible for members of the species to which she belongs. Thus, given that the cognitively impaired human's level of well-being falls short of the levels accessible to those with cognitive and emotional capacities which are the norm for the members

of her kind, we can conclude that she is unfortunate. An animal with a similar endowment of cognitive capacities would not, however, fall short of the norm for its species and thus should not be considered to be equally unfortunate. Thus, we can think of how well a human's life is faring *for a human*.

The Species Norm Account, then, preserves our strongly held intuitions, namely *Fortunate Non-Humans* and *Unfortunate Humans*. So far, so good. McMahan, however, takes the Species Norm Account to be undermined by the following two counterexamples.

1.1. Superchimp #1

A newborn chimp is treated to a form of gene therapy which causes its brain to develop the cognitive and emotional capacities of the human brain so that by adulthood, it has the cognitive capacities of a 10-year-old human. After some years of enjoying these enhanced capacities, this Superchimp suffers brain damage which reduces him to the cognitive levels of a regular chimp, after which he enjoys a contented life among other chimps.²⁹

As it stands, the Species Norm Account cannot view Superchimp as any more unfortunate than regular chimps given the fact that, after his reversal, his level of well-being is close to the levels enjoyed by those with the cognitive levels characteristic of members of his species. McMahan, however, thinks that the Species Norm Account miscalibrates misfortune. A human 10-year-old who suffers the same loss, leaving her cognitively impaired, is regarded as “a tragic misfortune that leaves the victim in a pitifully unfortunate condition” (2002, p. 148). Superchimp's decline, however, is treated on this account as a misfortune, but one that leaves him in a good state overall.³⁰ For McMahan, then, it is arbitrary “to suppose that the mere

²⁹ See Tooley (1972) for a similar case involving a kitten.

³⁰ It is possible that there may still be room within the Species Norm Account to say that Superchimp has been unfortunate: he has suffered a reversal of good fortune. Thus, the Species Norm Account is not precluded from acknowledging that Superchimp has suffered a *personal* misfortune. Here it is important to note that there are different jobs for verdicts of misfortune to do: to organize standards for compensation, and to offer

difference in species should make the human being's state an unfortunate one while the same state in the Superchimp is not" (2002, p. 148).

McMahan argues that the Species Norm Account is put under further pressure when this example is extended in the following case.

1.2. Superchimp #2

Suppose that a number of cognitively enhanced chimps are created at a time when regular chimps have become endangered and that the new enhanced chimps become more numerous than regular chimps. The species norm for the levels of well-being possible for chimps will now be significantly higher and, as such, the regular chimps will have become abnormal.

On the Species Norm Account, the regular chimps will have been rendered unfortunate simply by the existence of another group of individuals and this, McMahan argues, is absurd (2002, p. 148). How fortunate or otherwise an individual is should not, he maintains, be subject to change depending on whether there exists another group with access to the higher levels of well-being which accompany greater cognitive capacities.

Therefore, in *Superchimp #1*, the problem is that Superchimp is *not* considered unfortunate when he suffers a change in his circumstances for the worse, while in *Superchimp #2*, the problem is that the normal chimps *become* unfortunate even though there is no change in their circumstances. As such, while the Species Norm Account passes the second test of a successful account of misfortune in preserving our widely-held intuitions about the relative levels of

commentary on how well an individual is doing. See Lang (2012) for a discussion of this. What McMahan seems concerned with and, indeed, what is central to justice, is the former. As such, despite suffering a personal misfortune, it is clear that this sort of misfortune is not the type that triggers any demand for compensation. I discuss the role of notions of fortune and misfortune in more detail in section 3 below.

fortune of cognitively impaired humans and those of animals, it is unable to avoid being undermined by McMahan's counterexamples and thus fails the first test. We should note here that for these judgements to be considered a problem for the Species Norm Account, we need to assume that there are right answers to these cases; specifically, we need to assume that in the first case, Superchimp's change in circumstances renders him unfortunate, while in the second case, the normal chimps do not suffer a misfortune.

Can the Species Norm Account be rescued? To do so, we need an argument which shows that McMahan's counterexamples do not undermine the account as persuasively as McMahan takes them to.

1.3. Superchimp #1 revisited:

Eva Kittay (2005) takes exception to the first Superchimp case, arguing that there is nothing arbitrary in supposing that differences in species membership can make a difference to an individual's level of misfortune.³¹ She highlights this with the following case:

Flash

A drug is administered to Barry as an infant which allows him to develop the speed of a cougar as an adult. Given his enhanced speed, Barry is forbidden from participating in competitive races. Barry then suffers an accident which causes him to lose his enhanced speed reducing him to the speed of an above-average human sprinter. Compare this to a cougar who suffers the same loss and is reduced to a similar speed.

Kittay argues that while this loss will actually be good for Barry, as he can now compete in races, the cougar will have suffered a tragic loss as it can no longer hunt and survive in the

³¹ Grau (2015) makes a similar point.

wild.³² She suggests that as both individuals suffer the same loss of speed, species membership is required to make sense of the difference of the scale of loss in each case. If this is right, Kittay concludes, then “why should we assume that species membership is arbitrary in the area of cognitive capacities and not in the area of mobility? That surely does seem arbitrary” (2005, p. 113).

Moreover, Kittay maintains that as well as providing a norm against which to measure well-being, there are other ways in which species membership can be relevant in assessing well-being, such as the following. McMahan, she argues, is mistaken in treating the losses of Superchimp and the human as comparably bad. The human would have suffered a greater misfortune than Superchimp because the latter is able to go back to living the normal life of a regular chimp and to enjoy the levels of well-being that come with this. “Poor Superchimp”, as Kittay refers to him, would have been quite unfortunate prior to his reversal back to regular chimpood, owing to his outsider status in both the human and chimp communities resulting in his missing out on many of the social needs required for a good life. As such, the Species Norm Account does not give us the wrong answer when it tells us that Superchimp is not unfortunate when he suffers his cognitive loss.

Addressing these remarks in turn, it is not clear that species membership is playing the role that Kittay assumes it is in the *Flash* case. Our knowledge of the fact that the cougar needs speed to hunt to survive, and has thus suffered a tragic misfortune when it has its speed reduced, does not seem to be contingent on knowledge of what species it is a member of. It’s enough

³² Here, I think, Kittay overcomplicates the issue by framing Barry’s loss as a positive one and, in doing so, makes her case more difficult to make than it needs to be. It occurs to me that all Kittay requires for her argument is that both Barry and the cougar suffer a similar loss, but that the cougar’s loss is much more detrimental to it than Barry’s loss is to him. As such, I suggest that we think of Barry’s loss as a genuine loss, despite any coincidental benefits which might arise from it to avoid comparisons of the sort which suggest that it might be good for an individual athlete who loses a leg because they could then participate in the Paralympics.

that we take an inventory of the cougar's intrinsic properties, and its environment, to know that it cannot survive in the wild without speed. Here it might be suggested that when we tell informative stories about both the environment and the individual's properties, we are ending up with a story about typical species activity in all but name. But consider a cougar in a zoo, whose survival does not depend on hunting, and who suffers the same loss of speed. Here it seems that the loss of speed is not the tragic misfortune that it would be for the cougar in the wild. As such, it cannot be the case that it is the wild cougar's species membership which tells us how unfortunate it is when it loses its speed, otherwise it would tell us that the captive cougar is equally unfortunate. But this is clearly not the case. The captive cougar is perhaps unfortunate (to the extent that any individual who loses speed which is not required to survive would be), but not nearly to the same degree as the wild cougar who will surely starve without its speed. Thus, we do not make sense of the difference in degree of loss by pointing to the fact that they both belong to the same species. While species membership might provide us with a rough guide, the lesson from the above example is that looking to species membership is apt to give us the wrong answers.

It is also not clear that highlighting Superchimp's sense of isolation in the first Superchimp case offers any help to the Species Norm Account. Firstly, we should note that Kittay has gerrymandered the example a little when she highlights the pitiable life that Superchimp would lead prior to his reversal, when in McMahan's original example it is assumed that the loss in cognitive capacities will be accompanied by a decline in his overall levels of well-being. Indeed, we might think it necessary for the example to be framed in terms of a loss of well-being so as to allow us to ask the right questions. More importantly, though, the point that McMahan is making with this counterexample is that both the human and Superchimp suffer a similar loss in their *access* to higher levels of well-being, irrespective of whether either individual was realising those levels before their respective accidents. While it is true that

Superchimp can now enjoy the levels of well-being associated with regular chimphood, the point is that these levels are now severely curtailed by his new cognitive capacities. Furthermore, it is also true that the now cognitively impaired human can also enjoy the limited levels of well-being open to cognitively impaired humans. As such, whether an individual is unfortunate or not cannot be entirely tied to contingent matters such as the society you happen to find yourself in. For consider Shelly Kagan's plague case where an incurable plague sweeps across the Earth and causes almost all humans to never develop into persons (2015, p. 14).³³ It seems plausible to think that a human person in this world, surrounded only by non-persons, would still be deeply unfortunate were she to suffer an accident and to have her cognitive capacities reduced to the level of those around her. It is true that both this human and Superchimp are unfortunate to find themselves living in a society where they are the only member of their kind to have access to the high levels of well-being that they enjoy, but this type of misfortune does not negate the more important, non-contingent, misfortune which accrues with the loss of their cognitive capacities.³⁴ The fact that both this human and Superchimp would now be able to lead 'normal' lives, like those other members of their species, does not diminish the fact that they are both acutely unfortunate when they suffer their respective cognitive losses, even if it does perhaps cushion the blow a little. Furthermore, in describing the misfortune in this way, it is unclear that Kittay is actually appealing to species here, and thus it is difficult to see how this helps to demonstrate that species membership can be relevant in assessing well-being.

1.4. Superchimp#2 revisited:

³³ I discuss this case in more detail in the following chapter.

³⁴ This still holds true even when we consider that the human in question is likely not living and functioning as a 'normal' human person before her accident in such an example. See the next chapter for an argument that 'normal' here should not be determined on a statistical basis.

Recently, Doran Smolkin (2015) has argued that the unintuitive implication of *Superchimp#2* – that regular chimps could be rendered unfortunate without any change to their circumstances – should not prevent us from embracing the Species Norm Account. Smolkin maintains that such a result is not as absurd as McMahan takes it to be and, to make it a little more palatable, he suggests adopting a recent “promising option” from Shelley Kagan (2015, p. 19). He proposes that we might hold that an individual who is a non-person is unfortunate if they have the possibility of becoming a person but don’t become a person, and that how unfortunate they are depends on how realistic that possibility is, or on whether enough members of one’s kind are persons so that the species norm includes personhood (2016, p. 11).³⁵ Adopting this approach allows Smolkin to “explain why in some *very minor sense* one’s dog is unfortunate for not being a person, as one’s dog could in some possible world be a person” (2016, p.11).

Perhaps Smolkin has a point here. There does seem to be some intuitive support for the thought that an individual can be unfortunate for not having the higher capacities and potentials that others have. I find it difficult to accept, however, that how fortunate an individual is can be subject to change overnight, depending on the discovery of another group of individuals. Consider an updated version of Kagan’s Martian dogs case (2015, 13). Suppose that we discover a population of dogs living on Mars. These Martian dogs are like Earth dogs in all ways bar one; they are persons. Suppose further that this population greatly outnumbers the population of dogs living on Earth so that the average cognitive capacities for dogs is much higher than those possessed by Earth dogs. On Smolkin’s proposed account, Earth dogs who were hitherto considered to be fortunate have now become unfortunate overnight in virtue of the discovery of another, larger, group of dogs living on another planet. Perhaps, though, it

³⁵ I discuss such modal properties in more detail in chapter 4.

might be argued that the Earth dogs were already unfortunate and that it simply took the discovery of the Martian dogs for us to realise this. Perhaps, then, the discovery sheds light on something else: namely, on how likely it is that dogs on Earth could have been in the same situation as those on Mars. However, it is not clear that this is any less counterintuitive. For consider the following. Suppose that a week after this discovery, a pathogen introduced by the human explorers who make the discovery wipes out the entire population of Martian dogs. On one hand, if we think that the Earth dogs were already unfortunate, then they have suddenly become fortunate overnight without having experienced a change in their levels of well-being. On the other hand, if we think that the Earth dogs were fortunate before the discovery then, with the extinction of the Martian dogs, they will have suddenly become fortunate once again. Therefore, within the space of a week, all dogs on Earth have either went from being fortunate, to unfortunate, and back again to fortunate, *or* from being unfortunate to fortunate, all without ever experiencing any change within their own lives. Either way, the notion that whether a group of individuals, whose levels of well-being remain unchanged, are considered badly off or not hinges on the extinction of another entirely distinct and unrelated group of individuals has a whiff of the absurd about it. As such, one outcome strikes me as counterintuitive as the other.

2. *The Intrinsic Potential Account*

The problem with the Species Norm Account, which McMahan thinks the Superchimp counterexamples highlight, is that it judges an individual as unfortunate because she fails to realise the levels of well-being enjoyed by others, even though she may lack the cognitive and emotional capacities for well-being that these others possess. It is this extrinsic comparative element which is problematic. Instead, a successful account of misfortune should look at an

individual's intrinsic "nature or constitution" (2002, p. 149). McMahan offers the following as his favoured account of misfortune:

Whether a being is well or badly off depends on how its level of well-being compares to the range of well-being made possible by the highest cognitive and emotional capacities that it has actually achieved or that it natively had the potential to achieve (1996, p. 19).

It is clear that McMahan has a particular understanding of an individual's 'intrinsic potential' here. For him, the neural hardware of an individual's brain "sets limits on its cognitive and emotional capacities and potentials...[that] in turn define or delimit the individual's capacities for well-being" (2002, p. 150). He maintains that, "if the incapacity to have...a certain good is...part of an individual's congenital or native psychological constitution, the absence of that good is not so much a misfortune as simply a fact about that individual's essential nature" (2002, p. 150). Thus, for McMahan, an individual's intrinsic potential is determined by their congenital makeup irrespective of what species they belong to.³⁶ He ties an individual's intrinsic potential to their congenital makeup in an effort to rule out including animals within the scope of those who are considered unfortunate for not having access to the higher levels of well-being made possible by the cognitive levels characteristic of humans. The idea is that should enhancement technology become widely available and reliable, the potential that this would provide some animals with would be merely *extrinsic* potential which, McMahan maintains, has less normative importance than intrinsic potential. Before moving on to look at whether a clear distinction between intrinsic and extrinsic potential can in fact be comfortably drawn and, if so, whether one should be awarded with more

³⁶ McMahan's understanding of a 'congenital' condition is one where the "cause is physically present and operative at, or perhaps shortly after, conception" (1996, p. 5).

significance than the other, we should note some of the implications which arise from such an understanding.

Despite endorsing the Intrinsic Potential Account, McMahan recognises the following strange result. Consider a congenitally cognitively impaired infant and an infant who suffers brain damage after birth. Even though both infants have the same cognitive levels, and intuitively seem equally unfortunate, the Intrinsic Potential Account views only the brain-damaged infant as unfortunate, “whereas the congenitally impaired infant simply is what it is” (1996, p. 19). Thus, the fact that the latter infant’s psychological condition is congenital restricts the levels of well-being she had the potential to achieve, while because the former infant’s condition is not congenital, she possessed, and then lost, the potential to achieve the higher levels of well-being that accompany greater cognitive capacities. The brain-damaged infant is thus unfortunate for losing this potential, while the congenitally impaired infant didn’t have any potential to lose.

McMahan also makes a distinction between the two types of comparison in the Intrinsic Potential Account: A comparison between an individual’s well-being and the levels made possible by the highest cognitive capacities that she has *actually achieved*, and another between an individual’s well-being and the levels made possible by the highest capacities she had the *potential to achieve*. For ease of reading, I will refer to the first comparison class as ‘Actual-Comparison’ and the second as ‘Potential-Comparison’. He holds that “it seems worse to lose certain capacities than to lose only the potential to have those capacities” (1996, p. 20). Therefore, an individual with a divergence between the second comparison (Potential-Comparison) would be considered unfortunate, but she would be even more unfortunate if she had a divergence with respect to the first comparison (Actual-Comparison). Thus, if we could enhance a normal adult human who has her cognitive capacities severely curtailed following a

stroke, or the brain-damaged infant, but could not help both, we should restore the adult as she is more unfortunate.

Finally, despite McMahan's criticisms of the Species Norm Account, he recognises that the Intrinsic Potential Account must include some extrinsic comparative element to escape the following counterexample. Suppose that an adult human with above average cognitive capacities fulfils much less of her potential than others with similar capacities, but still has much higher levels of well-being than a cognitively impaired human who fulfils most of his potential. As it stands, the Intrinsic Potential Account will regard the cognitively impaired human as more fortunate than the above-average human, and this seems implausible. As such, McMahan suggests the addition of the following comparative dimension to his account. This dimension looks at an individual's levels of well-being and compares them to other individuals with similar cognitive capacities. McMahan seems to have two main comparison classes in mind for humans here: in one class are all persons and those who are not persons but who have the intrinsic potential to be persons, and in the other are all non-persons who lack the intrinsic potential for personhood (1996, p. 23-4). He says,

while it is a misfortune to fare poorly relative to the range of well-being that defines one's category, it is not a misfortune to belong to a lower rather than a higher category. Thus, we do not regard animals as unfortunate because they are incapable of achieving levels of well-being as high as ours. Nor do we regard ourselves as unfortunate because we are not gods (1996, p. 23-4).

The Intrinsic Potential Account has two dimensions, then, a non-comparative one and a comparative one. The non-comparative element compares an individual's levels of well-being with the levels made possible by her cognitive capacities and potentials.³⁷ This is the deeper,

³⁷ Although these dimensions of comparison might be more accurately labelled 'intrinsic' and 'extrinsic' comparison classes, here I keep with McMahan's terms despite the fact that it may be a little misleading to refer to an essentially comparative dimension of comparison as non-comparative.

more objective, measure of fortune. While the comparative element compares an individual's levels of well-being with other individuals who share similar cognitive capacities and potentials and determines whether an individual is unfortunate in a less deep sense.

2.1. Problems with the Intrinsic Potential Account

2.1a. Intrinsic/Extrinsic Potential

Given that the Intrinsic Potential Account holds that an individual can be badly off when her level of well-being falls below the levels made possible by the highest cognitive capacities she has actually achieved, the account suggests that Superchimp is unfortunate for having his cognitive capacities reduced to the levels of normal chimps. The Intrinsic Potential Account, then, gives us the right answer when we consider *Superchimp#1*. So far so good. However, two problems present themselves: first, it's difficult to come up with a descriptive specification of the difference between intrinsic and extrinsic potential; and second, it's difficult to justify the relevance of that distinction, even assuming that the first problem can be solved.

To see this, suppose that several months after Superchimp has been treated with the gene therapy, while still an infant, he suffers brain damage which causes his brain to revert to developing in ways that parallel the brain of a regular chimp. On McMahan's understanding of 'intrinsic potential', it seems that in this case, Superchimp will be no more unfortunate than ordinary chimps even though a human infant who suffers a comparable decline through brain damage is regarded as unfortunate. But if the goal is to have an account that recognises that when two individuals "have both fallen from the same height to the same lower state, it seems that either both are unfortunate or neither is" (2002, p. 148), then it looks as though McMahan should reject the Intrinsic Potential Account for the same reason that he rejects the Species Norm Account.

McMahan might respond here that the fact that Superchimp needs to suffer brain damage to deflect him from becoming a person suggests that the gene therapy has endowed Superchimp, in the relevant sense, with intrinsic potential. If so, then the Intrinsic Potential Account does in fact suggest that Superchimp has suffered a misfortune. This raises the question of what exactly counts as *intrinsic* potential. If the potential that Superchimp loses is “intrinsic” in the relevant sense then this suggests that intrinsic potential is the kind of thing that can be acquired rather than something which is necessarily part of an individual’s native endowment.³⁸ Thus, if this potential, which is extrinsic before Superchimp is administered the therapy, becomes intrinsic after the therapy, then this seems to blur the line between intrinsic and extrinsic potential. But if McMahan is to maintain the line where he has drawn it and to exclude the congenitally impaired as well as animals from being considered unfortunate for not being persons, then he is reliant on the possibility of a clear and firm delineation of intrinsic potential from extrinsic potential. Although McMahan is aware of the difficulty of drawing such a line, he notes that “we do not normally consider everything that can be externally induced in an individual is therefore inherent in the individual’s potentiality” (1996, p. 22). He thinks that although difficult, it is possible “to defend a firm distinction between genuinely native potential – i.e., potential that is grounded in the physical constitution of the individual – and a broader notion of potential that includes all that a being could become...by being externally augmented” (1996, p. 22).

It is not clear, however, that this distinction can in fact be comfortably drawn in the ways that the Intrinsic Potential Account requires. This is characteristic of a more general trend within the literature in which the interaction between genes and environment is downplayed, and more weight is given to intrinsic influences over extrinsic ones. For, it seems plausible

³⁸ This then raises another troublesome question: can it be a misfortune to not have access to the *opportunity to gain* intrinsic potential?

that how we turn out is an interaction between our genes and environment. If so, then it is unclear how we can determine an individual's potential simply by looking at her congenital makeup and ignoring external influences. Consider the following. Intuitively, an acorn has the intrinsic potential to become an oak tree before it is planted and fed water and sunlight. But what makes this process different from Superchimp receiving gene therapy treatment? In both cases we are adding something new to encourage growth, and in neither case can the target of growth (the development of the brain and the development of the plant) occur without this necessary external factor. In both cases, the addition uses what is already there and works within the pre-existing framework to encourage the required development. It is not obvious that we should consider water and sunlight to be relevantly dissimilar to Superchimp's gene therapy treatment in this context. For, they are clearly not internal factors. It is not the case that the gene therapy creates a new brain in a brainless object such as a rock. Rather, the gene therapy merely encourages the already existing brain to develop in different ways than it would have without the therapy. While it is true that Superchimp's brain would not develop in ways that parallel the development of a human person's brain without the therapy, it is also true that the acorn would not develop into an oak tree without the addition of sunlight and water. Perhaps there is a line to be drawn here between nourishment and enhancement, but this is not obvious, and simply maintaining that such a distinction exists does not suggest that moral significance ought to be placed on it.

Furthermore, McMahan's inclusion of the "highest cognitive capacities that one has the potential to achieve" as a comparison class, seems to suggest that in at least one sense, we are surely all unfortunate to at least some degree. Perhaps if I had spent less time as a child playing video games and less time as a teenager playing sports and drinking alcohol, instead spending my weekends in the library reading my way through the Russian classics of literature while snacking on superfoods, my capacities and potentials would be much higher than they presently

are. It seems clear that my growth and development has been stunted to at least some degree by a less than optimum diet and education which would have given me access to higher levels of cognitive and emotional capacities. It appears, then, that an individual's potential will, to a large extent, be determined by the external environment around her and, as such, it is not clear that McMahan can plausibly rule out the influence played by external factors.

Moreover, consider an individual's potential lifespan. Clearly an individual who leads a life with high levels of well-being and dies at 20, is more unfortunate than an individual who leads a similarly good life but who dies at 80. But, how long a life an individual has the potential to lead changes radically with time, location, and both existence of, and access to, technology. Given this, it seems difficult for McMahan to claim that we are somehow internally programmed to live to a certain age when this is subject to change depending on where and when you live, as well as your means and access to certain technologies and lifestyles. Yet all these things are extrinsic facts. Recent research, for instance, suggests the first person to live to 150 years old has already been born. But, it is implausible to hold that a normal human born two centuries ago had the potential to live to 150 years old. Nor can McMahan simply maintain that if we were to remove these external factors from the situation and to place the individual in a different environment then the individual would have the potential to become something different. For if that is the case then we can similarly contend that the congenitally impaired also have the potential to be created differently if the technology for doing so is available. Such cases push the door wide open for what we understand as an individual's intrinsic potential.

A further worry is that McMahan owes us an explanation for why we should not think that the difference between the congenitally impaired infant and the brain-damaged infant merely lies with *when* the potential is blocked. The congenitally impaired infant came into being that

way, while the brain-damaged infant had its potential blocked through an accident of birth. Surely, however, the misfortune lies in the fact *that* the access to higher capacities and potentials is blocked, rather than *when* it is blocked. McMahan, then, is working with a metaphysical assumption that a congenitally impaired infant could not have been born otherwise (1996, p. 15). That is, if the lack of capacity is congenital in nature then the individual could not have come into existence without it in an identity-preserving way. It is not clear, however, that McMahan can fall back on such metaphysical claims here. For, if an individual can undergo extreme cognitive deviation in the direction of decline in an identity-preserving manner, as McMahan suggests brain-damaged infants can, then it is unclear why extreme deviation in the direction of improvement cannot also happen in way that is identity-preserving. Indeed, this is precisely what McMahan suggests when Superchimp has his cognitive capacities both enhanced and then diminished.³⁹ Furthermore, often whether a particular type of external augmentation is unproblematically considered to be part of an individual's intrinsic potential depends on how common and reliable its usage is. For example, it seems unproblematic to consider an individual born deaf who receives cochlear implants shortly after birth, to have the intrinsic potential to hear, even though the individual has a congenital defect which blocked the development of normal human hearing apparatus. Should the technology to enhance an individual's cognitive capacities exist, therefore, it is not obvious that McMahan can rely on a distinction between intrinsic and extrinsic potentiality to exclude from the scope of those who have the potential for personhood both the congenitally impaired and similarly cognitively endowed animals.

In summary, despite McMahan's acknowledgment of the difficulty involved, the Intrinsic Potential Account relies on being able to draw a clear distinction between intrinsic and extrinsic

³⁹ Smolkin makes a similar point (2016, p. 11).

potentiality so that it can tether misfortune solely to intrinsic potential. As such, it is clear that a lot of work is being done by this distinction. And yet, if, as I have argued, it is unclear whether such a distinction can be comfortably drawn at all, then it seems premature at best, and reckless at worst, to rely on it to determine an individual's level of misfortune. The point is not simply that a clear distinction cannot be drawn, and thus McMahan's account should include those with extrinsic potentiality. I agree with him that it is absurd to think that my dog has the potential for personhood and is thus unfortunate for not being a person. The point is perhaps a deeper one: that given the difficulty in attempting to delimit the notion of potential, and the propensity of any delimitation to fluctuate with time and both the existence of technology and how common and reliable its usage is, then clearly potential is not the kind of thing we should be using to determine whether an individual is unfortunate or not.⁴⁰ Potential is simply not the right tool for the job of line-drawing.

2.1b. Connor & Abigail

Another challenge to the Intrinsic Potential Account is the following counterexample from Doran Smolkin. Connor, a congenitally impaired human, has the potential for cognitive capacities below the level of a person. He fulfils most of his limited potential, and does better than most humans with comparable cognitive endowments. Thus, the Intrinsic Potential Account holds that, "Connor is fortunate in *all* respects. For he has fulfilled his (limited) potential and he has done so to a greater extent than others in his comparison group" (2016, p. 9). For Smolkin, however, it seems much more plausible to think that "there is at least some respects in which Connor is unfortunate, given...the limits on his well-being" (2016, p. 9). This is further highlighted, he thinks, when we consider Abigail, a person with above-average cognitive capacities who realises less of her potential than most others with similar cognitive

⁴⁰ See Harman (2003) for more thoughts on specifying this distinction and the relevance of it.

endowments but, whose overall levels of well-being are significantly higher than Connor's. While the Intrinsic Potential Account holds Connor to be fortunate in all respects, it views Abigail as unfortunate in all respects. If we think that morality is sensitive to misfortune, then this has the counterintuitive implication that "Abigail has stronger claims to resources than Connor" (2016, p. 9).

One thought here is that this counterexample only seems to be a problem for McMahan's account if we *compare* Connor's level of misfortune with Abigail's level of misfortune. That is, if we were to claim that *because* Connor is fortunate and Abigail is unfortunate, Connor is therefore *more* fortunate than Abigail. But this is precisely what McMahan's account urges that we should not do. Rather, the comparative dimension of the Intrinsic Potential Account tells us that comparisons are restricted to the category the individuals are members of. As such, it is not the case that Connor is fortunate simpliciter, while Abigail is not, but rather we should think of Connor as fortunate *for an individual with his capacities and potentials* and of Abigail as unfortunate *for a person*. Alternatively, a second-order comparison seems immediately available here. That is, we might judge Connor to be more fortunate than Abigail on the grounds that Connor is more fortunate-relative-to-his-cognitive-level than Abigail is relative-to-her-cognitive-level. We can make that comparison without having to say that it is better, or worse, to have one set of cognitive levels than the other. What the Intrinsic Potential Account tells us then, is that neither non-human animals nor congenitally impaired humans are to be considered unfortunate for not having access to the levels of well-being typically enjoyed by human persons.

Furthermore, Connor's level of well-being is as high as it could possibly be given his (admittedly limited) cognitive levels. Thus, there is perhaps an argument to be made in favour of holding him to be fortunate, even if we simultaneously hold that Abigail, who is faring better

than Connor in terms of well-being, is nevertheless unfortunate. Therefore, although it does seem counterintuitive to suppose that if morality is sensitive to levels of fortune then we should give priority to Abigail, this implausibility appears to dissipate when we take the details of the case into consideration. For, recall that Connor is faring as well as he possibly could be relative to the levels of well-being he has the potential for, while Abigail is not. Thus, what would priority amount to in this case given that Connor is faring as well as he could be? The fact that how well he is capable of faring is limited does not alter the fact that he is maximizing his levels of well-being while Abigail is not.

But perhaps this misses the point. Perhaps the point is simply that, given the limits set on Connor's potential level of well-being by the cognitive capacities that he possesses, he is unfortunate for not having *access* to the levels of well-being enjoyed by those with higher cognitive capacities. And as Abigail has access to these higher levels of well-being it seems strange to view her as unfortunate when we view Connor as fortunate. So perhaps an option that McMahan might opt for is the addition of another comparative dimension to his account of yet lesser importance again. This dimension would broaden the comparison class to include both persons and non-persons. In the deepest sense, then, Connor is fortunate because he fulfils most of his limited potential, and in a less important sense he is fortunate because he fares better than most individuals in his comparison class, but in another respect, we can consider him to be unfortunate because his cognitive capacities prevent him from having access to the levels of well-being that would be possible if he had higher cognitive capacities. This new additional comparative dimension would command less normative or evaluative weight than McMahan's other two dimensions. This does, however, leave the Intrinsic Potential Account open to the suggestion that animals are unfortunate for not having access to the higher cognitive capacities that persons have. However, bearing in mind that the respect in which they are unfortunate is ranked below McMahan's other two respects in terms of importance, it's not so

implausible to think that there is a very minor sense in which my dog is unfortunate for not having access to the higher levels of well-being that I have, even though he is fortunate in a deeper sense (as he fulfils most of his potential). Similarly, if we were to discover a species of intelligent Martians which vastly outstripped our cognitive capacities and thus had access to much higher levels of well-being, intuitively I think that there is a sense in which I might consider myself unfortunate for not being created with these higher cognitive levels which would have allowed me to enjoy higher levels of well-being. Of course, this does not preclude me from being considered, in the deepest sense, to be fortunate relative both to the levels of well-being I had the potential for and compared to other members of my comparison class.

3. *The Role of Misfortune Verdicts*

We might, however, wonder what purpose this amendment serves when there is so little normative weight being placed on it. Indeed, this measure of misfortune is like the misfortune I suffer for not being born with the talent of Chuck Berry. While I might consider myself unfortunate for not having access to the levels of well-being that such talent would allow, it is unclear what, if anything, should follow from this. Furthermore, the intuitive force behind this extension is not simply that there is some minor sense in which Connor is unfortunate for not being a person but rather that Connor is *genuinely* more unfortunate than Abigail and, moreover, that he is more unfortunate than a dog is for lacking personhood. As such, the addition of a minor dimension of misfortune which bears so little normative weight does little to satisfy such an intuition.

Indeed, we might wonder if McMahan's comparative measure of fortune does in fact solve the problems suggested by cases like Connor and Abigail. For, McMahan attempts to avoid the counterintuitive result that Connor is more fortunate than Abigail, by holding that we cannot make cross-comparisons such as this. That is, we cannot compare the fortune of congenitally

impaired individuals with that of persons. But something has gone wrong here. For, it is plausible to suppose that the verdict that we want a successful account of misfortune to provide us with is that the congenitally impaired individual *is in fact* more unfortunate than the (underachieving) human person. Indeed, it seems likely that it is this thought which lies behind the discomfort we encounter at the original verdict: that the congenitally impaired individual is more fortunate than the underachieving human person. If so, then McMahan has not solved the problem by suggesting that comparisons between these two individuals are illicit. Rather, he has been engaged in a piece of misdirection, distracting us from the original problem, by suggesting that the problem never existed.

Consider the following. In a state with a privately funded healthcare system, we create an organisation to determine which individuals in our society are badly off so that we can allocate the limited resources allocated to healthcare benefits, and we want to determine this by looking at an individual's salary. I might propose an account which measures how well an individual is doing by comparing her salary with others in her postcode. Raquel, who makes £9000 a year, earns more than most individuals in her postcode and, as such, is considered to be well-off on this account. Derek, on the other hand, who makes £90,000 a year, earns significantly less than most people in his postcode and, thus, is badly off on this account. It might then be argued that this account is problematic because it seems to suggest that Derek is worse off than Raquel, and if our healthcare system is to be sensitive to how badly off an individual is in terms of her salary, then this system suggests that we ought to allocate a greater proportion of healthcare benefits to Derek than to Raquel. Intuitively, though, it's plausible to think that we should be arriving at the opposite result; that Raquel is worse off than Derek and should receive more healthcare benefits accordingly. Now, I might respond that this is not actually a problem for my account, as it does not allow such comparisons between postcodes. Rather, an individual can only be worse or better off than an individual within her own postcode. As such,

Raquel is well-off-relative-to-her-postcode, and Derek is badly-off-relative-to-his-postcode. But this is hardly the point. The problem is not simply that the account gives us an unintuitive result when we compare the two individuals, but rather that it doesn't offer us such a comparison *at all*. The entire purpose of the account is to determine who is worse off overall and, indeed, to tell us that Raquel is in fact worse off than Derek. For what we want *is* such a comparison, which would then enable us to provide healthcare benefits to Raquel rather than Derek, or at least, a greater proportion of healthcare benefits to Raquel. The proposed system, therefore, is a flawed one.

Similarly, while there might be other roles or functions for misfortune verdicts to perform, organising standards of compensation is central to justice, and it seems that the Intrinsic Potential Account is simply not going to be up to scratch for this purpose. An account of misfortune should help us do this by telling us who is badly off and, intuitively, it seems that a congenitally impaired individual is worse off than a normal adult human. The Intrinsic Potential Account, because of the criteria it employs, gives us the result that Connor is fortunate and Abigail is unfortunate. This looks like a problem because it seems to suggest that Connor is *more* fortunate than Abigail. McMahan cannot solve this problem by simply claiming that the criteria used by his account of misfortune does not permit such comparisons. That might well be true. But the problem is a deeper one. The problem is that an account of misfortune *should* make such comparisons. Indeed, that is the purpose we desire from a successful account of misfortune. And, when it does make such comparisons, it should give us answers which correlate with our deeply held intuitions: namely, that Connor is more unfortunate than Abigail. Therefore, McMahan's addition of this comparative element does not neatly sidestep the problem. Rather, it merely highlights a failure of his account: that it neither offers the comparison we want it to, nor provides the right judgement if it were to allow the comparison.

4. *Conclusion*

In conclusion, while the Intrinsic Potential Account gives us the right answers to the Superchimp cases, it fails to preserve our common-sense intuitions. In particular, it fails to uphold the following intuition:

2. Kenneth, a cognitively impaired human, *is* unfortunate for not being a person.

While we can preserve these intuitions by adopting the Species Norm Account, we can do so only at the cost of failing to clear the first constraint on a successful account of misfortune, namely, it fails to provide us with the right answers to the Superchimp counterexamples. While one of these results might be more palatable than the other, I suggest that either failure is enough to undermine an account of misfortune.

Works Cited

- Grau, C., 2015. McMahan on Speciesism and Deprivation. *The Southern Journal of Philosophy*, 53(2), pp. 216-226.
- Harman, E., 2003. The Potentiality Problem. *Philosophical Studies*, 114, pp. 173-98.
- Kagan, S., 2015. What's Wrong with Specieism?. *Journal of Applied Philosophy*, Volume (Early Online Edition), pp. 1-21.
- Kittay, E. F., 2005. At the Margins of Moral Personhood. *Ethics*, Volume 116, pp. 100-131.
- Lang, G., 2012a. Discrimination, Partial Concern, and Arbitrariness. In: U. Heuer & G. Lang, eds. *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*. Oxford: Oxford University Press, pp. 293-332.
- McMahan, J., 2002. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press.
- McMahan, J., 1996. Cognitive Disability, Misfortune and Justice. *Philosophy and Public Affairs*, 25(1), pp. 3-35.
- Singer, P., 1993. *Practical Ethics*. 2nd ed. New York: Cambridge University Press.
- Smolkin, D., 2016. Overall Lifelong Fortune: A Critique of the Intrinsic Potential Account. *Journal of Applied Philosophy*, Volume (Early Online Edition), pp. 1-13.
- Tooley, M., 1972. Abortion and Infanticide. *Philosophy and Public Affairs*, 2(1), pp. 32-65.

Chapter 4

How to Get Away with Speciesism

Introduction

This chapter presents a set of problems to be overcome if we are to justify the ‘Common Sense Intuition’: that the interests of humans count for more than those of non-human animals. While I do not offer solutions to the problems raised, and in fact argue that proposed solutions to these problems have so far been unsuccessful, I think that overcoming these problems will take us most of the way toward a successful defence of the Common Sense Intuition.

The chapter begins by defending Peter Singer’s claim that those who defend speciesism leave themselves open to parallel defences of other forms of prejudice such as racism and sexism. As such, those who wish to defend some preference for all humans over all animals must provide a justification for doing so which does not leave them open to such troublesome analogies. This will be the first constraint on a successful defence of the intuition that humans matter more than animals.

The chapter will then consider potentiality arguments which attempt to provide such a justification. Often such arguments simply take the form of asserting that ‘this is because of the *kind* of thing that he is’ where the ‘kind’ in question is explained in reference to ‘normal’ (Stone, 1994), ‘paradigmatic’ (Oderberg, 2000a), ‘generic’ (Kagan, 2015), ‘typical’, or ‘characteristic’ members of that kind. What is needed, though, is an explanation of what these terms mean, why we should accept this particular meaning rather than another, why we should think that being a member of a kind is imbued with any moral significance, and how exactly we are to delineate which ‘kind’ is the relevant kind to be taken into consideration? Such

questions are of interest to more than just those interested in speciesism as similar arguments are often found within the abortion debate.

I separate such accounts into two main categories: those attempting to provide a defence of speciesism and those attempting to justify the Common Sense Intuition while avoiding the charge of speciesism. Considering these in turn, I maintain that both types of arguments are hamstrung by similar problems.

1. Is Speciesism a Mere Prejudice?

Before we consider attempts to justify or avoid speciesism we should ask: what is so bad about speciesism, and is it something that actually needs to be avoided?

Peter Singer argues that those who attempt to defend the intuition that we should give more weight to the interests of humans than to the interests of non-human animals - in virtue of the fact that they are *human* interests - without pointing to a morally relevant property, are guilty of speciesism, a prejudice objectionable for the same reasons that racism and sexism are. Singer says:

Racists violate the principle of equality by giving greater weight to the interests of members of their own race when there is a clash between their interests and the interests of those of another race...Similarly, those I would call 'speciesists' give greater weight to the interests of members of their own species when there is a clash between their interests and the interests of those of other species (1993, p. 58).

Of course, not every preference for members of a kind is malign. There are many types of discrimination which we consider to either be innocuous or upstanding: for example, we favour, for our legal representation, members of the class of qualified lawyers, rather than non-members of that class and we favour members of the class of trained medical doctors for our

medical treatment, not those who fall outside the class. For Singer, though, the type of creature you are makes no contribution to the inventory of morally significant properties you have when deciding whether to count your interests, and how much weight to give to them. Species membership is morally arbitrary, or uninteresting, as is sex and race. For Singer, the properties of species, race, and sex conceal morally significant properties rather than present us with morally significant properties.

In a recent paper, however, Shelley Kagan has argued that speciesism is not necessarily a mere prejudice and thus not objectionable in the way that other forms of prejudice, like racism and sexism, are. He maintains that anti-speciesism arguments like Singer's, while *prima facie* convincing, do not actually contain "much by way of philosophical argument" and, as such, speciesism may be a more philosophically defensible view than first thought (2015, p. 2). Singer argues that we should treat like interests alike and maintains that "pain is pain" no matter who is experiencing it (2009, p. 20). But Kagan asks, what are the morally relevant ways that two interests can be like one another or differ, and what are the morally relevant features of pains? He argues that Singer has not provided an argument about which differences we should consider to be relevant. For Singer, when it comes to pain, only intensity and duration matter. Kagan, though, thinks that it is also relevant to ask "*who* the pain belongs to – that ownership of the pain is in fact a morally relevant difference, even among pains that are otherwise alike (in terms of duration and intensity)" (2015, p. 5). He maintains that Singer has not provided an argument which establishes that it is illegitimate for a speciesist to give more weight to human pain than to animal pain, simply because it is experienced *by a human*. As such, Kagan holds that Singer has given no reason to think that the speciesist's views are the product of mere prejudice rather than being sensitive to a morally relevant feature – namely, the fact that the pain is had by a human. As such, Kagan argues that Singer is missing an argument to get him from 'pain is pain' to 'speciesism is a mere prejudice'.

However, it is not obvious that the burden is in fact on the anti-speciesist to demonstrate that what species an individual is a member of is irrelevant when it comes to deciding whether to count one individual's pain more than another's. Singer maintains that we cannot decide the question of whether or not an individual can feel pain simply by looking at their species membership any more than we can by looking at their skin colour or sex. Kagan might want to hold that species membership can provide us with a good indication of whether or not an individual can feel pain - that is, whether or not typical members of that species are sentient – but that cannot get him the conclusion that he is seeking: that we should give more weight to human pain than animal pain. Whether the typical members of a species are sentient or not can only tell us whether to register their pain, but it cannot tell us *how much* to count a given individual's pain compared with another's. As such, it would seem that the burden rests with the speciesist to offer an argument suggesting that two like pains should be counted differently simply in virtue of the fact that one of them is had by a human while the other is had by an animal.

By way of such an argument, Kagan asks us to suppose the following case:

Prisoners: Two prisoners, Andy and Red, are serving a life sentence and are both suffering an equal amount in terms of intensity and duration. Red, however, is innocent of any crime, while Andy has been justly imprisoned.

Kagan argues that the fact that Red does not *deserve* to be punished provides us with a reason to think that his *suffering* should be given greater weight than Andy's suffering. As such, if you could free one of the prisoners on the basis of their suffering, then you should count Red's suffering for more than Andy's, despite the fact that they are "equally miserable, and for an equally long time" (2015, p. 6). If this is right, and it is possible for these two similar pains to differ in this morally relevant way, then it is plausible to think that pains can differ in

other morally relevant ways as well even when it doesn't affect how the pains feel. In particular, "just as differences in intensity and duration and desert all affect whether two pains count as relevantly similar, perhaps it is also true that it matters who has the pain".

One thing to note here is that, at most, all Kagan has argued for is that desert is a morally relevant difference and that Singer is mistaken in holding that the only morally relevant factors when it comes to pain are intensity and duration. Kagan is missing a further step here to get him from the claim that 'there are other morally relevant factors besides intensity and duration' to the claim that 'species membership is a morally relevant factor'. As such, he has not provided an argument suggesting that we should count the pain of a human more than the pain of an animal, when both pains are similar.

More importantly, though, the Prisoners case itself seems to be problematic. It appears that desert is doing all of the work in this case. For, if the suffering had by the two prisoners is the same, as Kagan states that it is, then it cannot be the case that we are focussing *solely* on the suffering had by the two prisoners when we decide who should be released. For, given that they are suffering equally, and we are only considering their suffering, how then would we choose between them? Rather, what we seem to be doing is taking a separate consideration into account: the fact that Red's punishment is undeserved.⁴¹ Indeed, we might think that even if Red was suffering only a little (or not at all – perhaps he finds life in prison to be quite pleasant), while Andy was suffering a great deal, we should still choose to release Red given that he is innocent. Red's detention is unjust in a further sense: he's innocent of the charge that placed him there in the first place. Therefore, it seems that desert is doing all the heavy lifting in this case and, as such, we are not *only* taking the suffering of the two prisoners into account

⁴¹ We might be tempted to say that Red's suffering really is worse than Andy's *because* it is undeserved. That is, Red's suffering is more *intense* because he is innocent of any crime. Of course, this would offer little help to Kagan who wants to hold that the two prisoners are suffering *equally*.

when we decide who to release. To achieve his conclusion that the suffering differs in a morally relevant way other than intensity and duration, Kagan must insist that we only consider the prisoners' suffering and nothing else. But it is difficult to see how desert can be a morally relevant feature of the suffering while also holding that both prisoners' suffering is exactly alike. Therefore, Kagan has not shown that it matters *who* the suffering is had by. Rather, all he has shown is that desert is a morally relevant factor in prisoner release scenarios. But this is not something that the anti-speciesist need deny.

Without an independent argument for the relevance of species membership, then, it appears that the speciesist is left open to a charge of prejudice. If so, then it seems that, just as Singer has argued, speciesist arguments are vulnerable to parallel arguments defending other forms of prejudice such as racism and sexism (2015, p.2). To avoid such a comparison and to distinguish speciesism from these other prejudiced views, Kagan offers an account of prejudice, claiming that "you are prejudiced if you hold a view on the basis of evidence that you wouldn't otherwise consider adequate" (2015, p. 8). Kagan thinks that racists and sexists attempt to defend their views by appealing to empirical claims about differences in intelligence or moral character and that it is not merely that such claims are false, but that "they stick to their beliefs *despite* the evidence" (2015, p. 8). Similarly, then, Kagan argues that if the speciesist were to attempt to defend her views by appealing to false empirical claims about differences in intelligence between animals and cognitively impaired humans, despite evidence to the contrary, then speciesism, too, should be considered to be a mere prejudice. However, "if one's speciesism is based instead upon a direct appeal to moral intuition...and if one is prepared to give presumptive weight to moral intuitions in other matters as well, then that...is not prejudice" (2015, p. 8). The implication for Kagan, then, is that if an intuition isn't prejudiced in this sense - if it isn't responsive to faulty evidence or would endure despite shifts in evidence - then it simply can't be criticized. That is, if an intuition is invariant to changes of evidence, then

there is some evidence that it wasn't based on evidence anyway. Thus, for Kagan, it would be no longer be open to criticism because it was held regardless of evidence.

Singer, however, argues that Kagan's account of prejudice appears to have unintuitive implications about the badness of prejudice. Singer holds that it seems implausible that what we should find morally objectionable in the "conduct of those who carried on the slave trade, or brutally administered the lash to their slaves of African descent...is that they did not apply the appropriate evidentiary standards for evaluating their beliefs about the capacities of Africans" (2015, p. 3). It seems much more plausible to think that what we should find most objectionable in such conduct is that these individuals simply did not care about the "interests of Africans" (2015, p. 3). As such, we don't want it to be the case that the special badness involved in racism is simply that racists are epistemically irresponsible. Rather, we want to say that the badness lies in the fact that they fail to give the appropriate amount of consideration to the interests of those whom they are prejudiced against.⁴²

Perhaps, though, while a full account of prejudice would both identify prejudice and tell us the wrong involved in it, Kagan is offering only a partial account of prejudice concerned only with identifying prejudice. If so, then there does not seem to be any reason that Kagan cannot hold, along with Singer, that the wrong involved in the treatment of the slaves in the above

⁴² One worry here is that this view supposes the belief in those who participated in the slave trade that Africans *had* interests. As such, this view fails to distinguish between those who ignored those interests (while not denying that African slaves had them) and those who actually thought that African slaves lacked those interests. And yet we don't, intuitively, think that those, if any, who were acting on the belief that African slaves lacked these interests are then off the hook, morally speaking. The mistakes in each case do seem different: there's an epistemic error at the bottom of one type of mistreatment, and a shortfall in humanity at the bottom of the other type of mistreatment. My thanks to Gerald Lang for suggesting this worry. Although I am somewhat sceptical about the existence of the first class of slave traders, I have two thoughts here: First, while it does seem that something must explain why the interests of slaves weren't heeded - either the slaves were supposed to lack them, or they were discounted (though their existence not denied) due to surface differences or some other source of enmity - it is not obvious to me that these two explanations are distinct. That is, it seems likely that both stem from a similar source and it is this source that our definition of prejudice should account for. Second, I think that Singer's argument still applies: when considering what we find morally objectionable in this conduct there is a strong sense that these differences are unimportant: what matters is the mistreatment.

case is that their interests are not given appropriate consideration. However, this in itself does not tell us how to identify prejudice given that there seem to be many non-prejudiced treatments of individuals in which we do not give their interests appropriate consideration. It seems, then, that along with knowing what the wrong involved in the treatment of the slaves is, we will still need a definition of prejudice to identify whether such treatment stems from prejudiced views or not. Indeed, Singer's (2002, p. 106) own, more inclusive account of prejudice, which holds that you are prejudiced if you take into consideration properties that are arbitrary or irrelevant to the matter at hand, also does not give us a full account of the wrong involved in the slavery case. For we could hold that it is implausible that what we should find morally objectionable in the conduct of those who carried on the slave trade is that they took arbitrary or irrelevant properties, such as the colour of their slaves' skin, into account when deciding how to treat them. Rather, it seems much more plausible to think that what we should find most objectionable in such conduct is that these individuals simply did not care about the "interests of Africans". Thus, Singer's objection to Kagan's account of prejudice can be turned towards his own account.

That said, if we are to accept Kagan's definition of prejudice, then it had better be the case that it accounts for all the types of prejudice that we would want a definition of prejudice to account for. For instance, Kagan's account seems to give us the wrong answers when we consider those with prejudices based on *true* empirical beliefs. To see this, suppose that it turns out that, on average, men are typically more intelligent than women. Would we want to say that those sexists who appeal to such a claim to justify giving more weight to the interests of men and less to those of women are not in fact prejudiced? Such a consequence surely provides us with a reason in favour of adopting a more inclusive account of prejudice such as Singer's, mentioned above.

Furthermore, it seems that there are instances of prejudice that do not rely on false empirical beliefs but which are no less objectionable for all that. Consider racism once more. It looks as though a certain type of racist can simply take advantage of Kagan's account while maintaining that he is not in fact prejudiced because he does not attempt to provide any kind of justification whatsoever for his racist beliefs, but instead relies on a "direct appeal to moral intuition". Consider the following mantra from what Lang refers to as "Pure Racism":

We Whites don't care about Black's possession or non-possession of the capacities possessed by Whites. Blacks' skin colour alone settles the case for their exclusion from the White community and the patterns of moral concern the White community bestows upon all its members (2012a, p. 318).

While Kagan accepts that such people would not in fact be racists on his account, he argues that racists and sexists don't *actually* make such appeals to intuition, relying instead on appeals to false empirical beliefs. As such, he thinks that actual racists and sexists are in fact prejudiced (2015, p. 8). But, as Lang argues, regardless of whether such a form of racism would enjoy success or not, "we would surely think worse of any condemnation of racism that failed to find room for the condemnation of Pure Racism" (2012a, p. 318-9).

Moreover, it is not clear why Kagan thinks that the speciesist is more likely to appeal directly to intuition than the racist or sexist.⁴³ It appears that we can adopt Kagan's arguments and maintain that while those speciesists who support their views by appealing to intuition are

⁴³ Perhaps Kagan will reply that he is simply trying to excavate a form of speciesism that would be exempt from criticism, however rarely it was actually instantiated. It is unclear how much help this line of argument provides him with though. For one thing, we can simply turn Kagan's remarks above back on him and hold that while we can accept that Pure Speciesists would not in fact be speciesists on this account, speciesists don't actually make such appeals to intuition, relying instead on appeals to false empirical beliefs. As such, *actual* speciesists are in fact prejudiced. For another, Lang's criticism of Pure Racism would still hold here: that something has gone wrong with any condemnation of speciesism which is unable to find room for the condemnation of Pure Speciesism.

not prejudiced, almost all speciesists actually appeal to arguments based on empirically false beliefs - such as the cognitive superiority of even severely cognitively impaired humans over all animals, say. Indeed, this seems to be the case with many who hold that we should give more weight to human interests over animal ones. Such people often begin with arguments which are empirically false, such as the claim that humans should count for more because they are rational, have the capacity for language, are persons, and so on. However, often when such people are shown the difficulty of pointing to a morally relevant property that *all* humans have and no animals have, only then do they resort to the claim that they simply have the intuition that humans matter more. Kagan has given us no reason to think that the same thing does not also happen in other forms of prejudice. For instance, it would not be surprising to discover that a similar thing held true of those racists who appeal to empirically false beliefs. That is, it would not be surprising to discover that after being shown that their beliefs about differences in intelligence or moral character are false, that such racists would then reply that they simply have the intuition that white interests count for more than black interests, say. Perhaps, there is an argument to be made that the racist is less likely to claim that they simply have this intuition if it is instead put in terms of the moral significance of skin colour. Rather, it is what the colour of someone's skin indicates, that the individual is someone with low intelligence, low moral character, etc. But, then, it seems that a parallel argument can be used for speciesism. It is not simply that the individual is a member of a different species, it is what this indicates about the individual – that the individual has low intelligence, and so on. But it seems plausible to think that what underlies both views are similar intuitions that both skin colour and species membership are morally significant.

Indeed, we might wonder whether this isn't implicitly built into Kagan's account of prejudice. Kagan says that it is not merely that such claims are false, but that "they stick to their [empirical] beliefs *despite* the evidence" (2015, p.8). We might then ask, what is it that

makes such people stick to these beliefs despite having evidence to the contrary? It seems that what is feeding this stubborn refusal to adhere to the evidence presented to them is the intuition that skin colour is a morally relevant difference. If this is right, then Kagan's account of prejudice, far from dispelling the analogy between racism and speciesism, will instead have the consequence of either allowing cases of racism to avoid the charge of prejudice or of showing that speciesism is a prejudice after all.

Having said this, perhaps Kagan's point is simply that, given that we are appealing to intuitions, there is a sense in which the intuition that human interests count for more than animal interests is more intuitively plausible than the intuition that white interests count for more than black interests. That is, even if the same type of intuition lies at the bottom of both racism and speciesism, there is still something more counter-intuitive about 'pure racism' than there is about an equivalent 'pure speciesism'. Thus, perhaps Kagan's point is simply that even racists find 'pure racism' counter-intuitive, while most anti-speciesists find *something* intuitively plausible about the notion that human interests count for more than animal ones, at least in certain circumstances. But, we might think that this asymmetry can be explained by social factors. For, while it is true that we have a long history of racism, it is also true that we have an (admittedly more recent) history of social condemnation of racism. The same cannot be said for our speciesist practices. Thus, it makes sense to think that even the racist will be influenced by widespread societal condemnation of racism and, as such, will seek to justify his views with false empirical beliefs rather than appealing to intuition alone. In contrast, the speciesist, unencumbered by widespread condemnation of speciesism, will be much more comfortable in appealing purely to intuition to justify her speciesist views. But surely this fact should not determine whether we condemn one view as prejudice but not the other.

In summary, even if we defer to Kagan's specification of what prejudice is, and that speciesist views based only on intuition are not a mere prejudice, it is clear that such views steer uncomfortably close to prejudice and, moreover, that it isn't all that plausible to suppose that 'pure' intuitions are all guaranteed to be unimpeachable. Therefore, if possible, we should seek ways of justifying our speciesist accounts which would not also justify those prejudices that we do find objectionable. This, therefore, suggests a constraint on accounts which offer a defence of speciesism: they must not, in the course of defending speciesism, leave themselves open to parallel defences of objectionable prejudices such as racism and sexism. Call this 'Singer's Prejudice Constraint'. This is the first hurdle that a successful defence of speciesism must overcome.

2. *Embracing Speciesism: Potentiality*

How do we to justify the Common Sense Intuition without falling foul of the well-known argument from marginal cases? This argument holds that whichever morally relevant property we decide upon to demonstrate that humans are more worthy of consideration than animals, it will either be the case that this property is not possessed by *all* humans or that the property is held by some animals.⁴⁴

One way in which speciesists have attempted to defend their position, and avoid the argument from marginal cases, is by putting forward potentiality, or kind-membership, arguments. In order to defend the Common-Sense Intuition, they argue that the emphasis should be placed on kind-membership rather than a moral individualism that determines moral status by looking only at the intrinsic capacities of individuals. Cohen, for instance, has argued the following:

⁴⁴ For one example of this argument see Singer 1975, p. 265.

Persons who are unable, because of some disability to perform the full moral functions *natural* to human beings are certainly not for that reason ejected from the moral community. The issue is one of *kind*. Humans are of such a kind that they may be the subject of experiments only with their voluntary consent...Animals are of such a kind that it is impossible for them, in principle, to give or withhold voluntary consent or to make a moral choice. What humans retain when disabled, animals never had (Cohen, 1986, p. 866 *Italics mine*).

What this property is that all members of the kind ‘human’ have and what exactly ‘kind’ means here, differs from account to account, but the thought is roughly that those humans who do not *actually* instantiate the morally relevant properties – for whatever reason - which entitle humans to special treatment, are nevertheless entitled to such treatment *in virtue of* the fact that they are members of a kind whose ‘typical’ or ‘normal’ members do instantiate the relevant properties. Such terms as ‘typical’ and ‘normal’ are ambiguous and will require unpacking. It will be helpful, then, to look at the different ways that such terms can be understood in the hope of finding the most robust and defensible forms of them.

What is it about the kind ‘human’ that makes it the case that it is invested with moral significance? Those who adopt such positions claim that the kind ‘human’ gets its moral significance from the fact that its typical members are persons, and persons instantiate morally important properties, such as rationality, self-consciousness, future-orientated plan-making, and so on. But why should we then think that those ‘atypical’ non-person members of the kind who either no longer, cannot, or do not yet, instantiate such properties should derive a higher moral status simply in virtue of sharing a kind-membership with those who do?

The answer appears to be bound up with some notion of potentiality. Again, though, the language of potentiality can be ambiguous. For instance, does potential refer only to those

properties that will in fact be realised, those properties that have a high probability of being realised, or simply those properties that *could* be realised? This distinction will be important for answering the objection levelled at potentiality accounts by Nobis, who holds that while such accounts seem to confer moral status on those humans who have yet to become persons, they cannot provide the speciesist with what he requires, as they fail to cover those human non-persons whose impairment is so severe that they lack even the potential to become persons (Nobis, 2004, p. 49).

What the speciesist needs, then, is a broader conception of potentiality which covers all humans, according to which it can be said that human non-persons could have become mature human persons (had circumstances not intervened). Oderberg puts forward an Aristotelian account of potential holding that,

When Aristotle said that ‘man is a rational animal’, he was not making a statement about only those mature, normally functioning members of humankind...he was defining the *essence* of humankind, in other words, he was telling us what human nature is, and hence what *every* human being is, simply by being a member of humankind (2000a, p. 182-3).

Further, he claims that the “kind of thing an entity is determines its potentialities” (2000a, p. 178) and that while “some of us might only be potential *persons*, we are still actual persons, because we are human beings” (2000a, p. 182). For Oderberg, ‘*persons*’ refers to Singer’s sense of personhood, that is, those individuals who actualise the property of being a person, while ‘persons’ denotes *all* humans, including those who fail to instantiate the property. Thus, for Oderberg, simply being a member of humankind entitles you to a higher moral status regardless of whether your capacity to become a person (in Singer’s sense) is damaged.

Expanding on this Aristotelian notion of potentiality, Megone (2002) maintains that this potential to become a person, which all humans share as their essential nature, is what defines an individual as a member of the kind ‘human’. This is because ‘good’ or successful members of the kind actualise the potential to become a person, where success refers to survival of the kind. Only those potentialities that good members actualise will constitute the nature of every member of the kind. Thus, if a human individual fails to actualise the potential to become a person, she fails to “behave as a good member of [the] kind does (bearing in mind that the behaviour of good members of the kind is definitive of any member of the kind)” (Megone, 2002, p. 161). Thus, the potential actualised by a ‘bad’ or unsuccessful member of a kind, which Megone refers to as an individual’s “state” potential, is not part of its nature and cannot be explained by reference to the kind of thing it is (2002, p.173). Rather, only the potentialities that a flourishing member of a kind actualises will constitute the nature of every member of that kind. This is because it is only these “essential potentialities” that “figure in a teleological explanation of its development” (2002, p. 162). Thus, what makes an individual the thing it is “– a member of that kind – is its instantiation of this set of potentialities”. Therefore, a severely cognitively impaired infant and a normal adult human “instantiate the same essential potentialities in the sense that there is a set of potentialities, which a perfect human actualises, which each is in the process of actualising, with greater or lesser success” (2002, p. 162).

While such accounts seem able to justify attributing a higher moral status to all humans than all animals, they must address problems, recently put forward by Lawlor (2012), which arise from with relying on the species boundaries drawn by taxonomists. Lawlor argues that the criterion used by taxonomists to sort animals into groups according to species membership is in fact quite poorly suited to that task (2012, p. 152). He thinks that this becomes a problem for the speciesist when we consider the fact that moral status should be transitive. That is, if A has the same moral status as B, and B has the same moral status as C, then A and C should also

have the same moral status. But, building on work from Darwin and Dawkins, Lawlor highlights the fact that species membership does not seem to be similarly transitive. Both Darwin and Dawkins have argued that it is only possible for us to identify distinct species because all of the “intermediates” have become extinct. Had this not be the case, separating animals into separate species would not be possible as we would simply have “a smooth continuum of animals” (2005, p. 317). Holding interbreeding as the recognised criterion of whether two groups of animals should be classified as distinct species, Lawlor stresses that this is a problem for the speciesist because we can imagine a case where we discover an intermediate species between humans and chimpanzees. This species would be capable of interbreeding with both humans and chimpanzees, even though humans and chimpanzees are incapable of interbreeding. Thus, while humans and the intermediate group would be considered to belong to the same species, and chimpanzees belong to the same species as the intermediate group, humans and chimpanzees do not belong to the same species. Thus, species membership is not transitive. But if we think that moral status should be transitive, Lawlor continues, this will have the absurd result that a given animal’s moral status will be dependent on “contingent and seemingly irrelevant facts about which other animals have become extinct” (2002, p. 153).

One reply might be that while an individual’s species may depend on these contingencies, once species is nailed down, we can say other things about the nature of that creature, as Oderberg and Megone do. Thus, Oderberg and Megone can say that, had we been constituted in some other way, other claims about us would have had to be made. But we are constituted this way, and this permits us to do some metaphysics. The problem, though, is not simply that we could have been constituted in some other way, but rather that because of these contingencies, it is difficult to even nail down an individual’s species in a concrete way. Thus, it is problematic to place moral significance on the ‘nature’ of an individual when, for the neo-

Aristotelian, one's nature is determined by species membership and species membership is, in turn, determined by criteria which are poorly suited to the task and which could be subject to change.

One way to avoid this problem is to supply an account which does not place any direct emphasis on the importance of species membership, yet still has the result of attaching a higher moral status to all humans over all animals.⁴⁵ Liao puts forward such an account, which he refers to as “the genetic basis for moral agency account” (2010). He argues that a sufficient condition for having moral status is having the genetic basis for moral agency or personhood. He maintains that the set of physical codes which generate personhood are located in the human genome. Thus, all humans possess the genetic basis for personhood. Liao argues that if this capacity for personhood did not have a genetic basis then its development would not be so regular.⁴⁶ If this is right, then we can point to an identifiable, physical property that all members of the human species possess rather than relying on the fact that the severely cognitively impaired are just members of a group whose typical members possess a particular property.

For this account to escape the transitivity problem, though, it will need to be the case that it does not simply collapse into another potentiality account. Liao claims that his account differs from, and is preferable to, potentiality accounts and, as such, escapes any problems associated with such accounts.⁴⁷ One such problem, Liao suggests, is the following case. Ash

⁴⁵ Stone (1994) also seems to hint at such an account.

⁴⁶ I leave the science involved here aside, but see Burgess (2010, p. 150) for some reasons to doubt that this type of genetic determinism is defensible.

⁴⁷ Liao is keen to stress the differences between his account and potentiality accounts but it is not clear that the genetic basis account is in fact distinct from such accounts. On Liao's view, full undisturbed development from the genetic basis will eventually deliver the typical mature person, and this is what matters, even in cases where that developmental path is disturbed. But this is more or less what neo-Aristotelians think. There are some differences in what is taken to be doing the deep explanatory work - either the genetic basis or the more Aristotelian metaphysics - but structurally, they seem rather similar.

is a human with two hands. On a potentiality account we want to claim that Ash actually has two hands and that he has the potential for having two hands. But suppose that Ash's hand becomes badly infected, forcing him to cut it off. Liao argues that Ash no longer actually has two hands or has the potential to have two hands. On the genetic account, though, we can still claim that Ash retains the genetic basis for having two hands (2010, p.173). However, Liao seems to be working with an unduly narrow concept of potentiality here, one which proponents of potentiality arguments would not accept. Rather, the Aristotelian accounts discussed above would instead claim that having two hands is part of a set of potentialities that a perfect human actualises. As such, Ash has simply "failed to actualise some of the more significant ones that would be actualised by a good member of the kind" (Megone, 2002, p. 162). In this sense of the term, Ash retains the potential to have two hands.

It is not clear, however, that the genetic basis account is in fact superior to potentiality accounts. For instance, Liao recognises the following problem for his account. If there is such a thing as the genetic basis for personhood, which is contained in the human genome, then it seems that even human cells can be said to possess this since it is theoretically possible to use a human cell to produce a human being (Liao, 2010, p. 178). Here, the proponent of the potentiality account can rely on a distinction between intrinsic and extrinsic potentiality (Lang, 2012b, 142-3), or as Burgess refers to it "active potency" and "interactive potency" (2010, p. 141). Burgess holds that something has an active potential to become x just in case it will, if not prevented by external factors, become an x . By contrast, something has an interactive potential to become x just in case it has the potential to become an x given the right kind of interaction with something else. The proponent of potentiality can maintain that only intrinsic or active potentiality counts towards an individual's moral status. As such, it is not the case that the human cell has moral status because it only has extrinsic potential to become a person.

But given that Liao holds that his account is not a “potentiality account in disguise” (2010, p. 169), it is clear that he cannot help himself to such a reply. For, if what gives both human persons and non-persons a higher moral status than non-humans is the fact that they share a genetic basis for personhood, then it’s unclear how he can hold that those other entities which possess this same genetic basis are not also moral status holders. This seems to leave Liao facing a dilemma here. He can either stick to his guns and hold that his account is not a potentiality account, leaving himself open to the implausible implications about the moral status of human cells. Or he can accept that his account is in fact a type of potentiality account where having the genetic basis for personhood is just having the potential for it, thus allowing him to make use of the potentiality theorist’s distinction between intrinsic and extrinsic potentiality. But if he takes the latter horn then it is not clear that he doesn’t also inherit the problems associated with such accounts and, more specifically, Lawlor’s transitivity problem.

Moreover, this highlights another problematic issue, which Liao also acknowledges, for his account. It appears as though it will be difficult for him to hold that the moral status that one enjoys in virtue of having this genetic basis comes in degrees. For it is not as though one can be closer than another to realising this genetic basis – an individual either has this genetic basis for personhood or does not. And indeed, Liao claims that while an actual person might matter more than an embryo with only the genetic basis for personhood, he thinks that this does not show that they have different moral statuses (2010, p. 175). He gives the example of your partner and a stranger drowning in a lake and argues that even though your partner matters more to you, this does not show that she has a higher moral status than the stranger. Liao is right that it doesn’t show this. However, surely we do not want to say of an embryo and your partner, that while they have the same moral status, your partner matters more in virtue of her relationship to you. Rather, we want to say that your partner matters more *simpliciter*, for agent-neutral reasons – because she is a person while the embryo is not.

3. *Avoiding Speciesism: Personism*

Can we defend the Common Sense Intuition and avoid the problems facing the speciesist mentioned above, by attempting to avoid speciesism rather than attempting to justify it? Kagan argues that what those who want to defend this intuition actually believe is not speciesism but what he refers to as “personism” (2015, p. 12). This is the view that the interests of persons count for more than the interests of non-persons. However, this won’t quite do as a full account of what we believe when we consider how those humans who are non-persons should be treated, and specifically, those who have never been persons and never will be (that is, the severely cognitively impaired). Kagan asks why we think that it is worse to experiment on such a human than on an animal? He thinks that what we are concerned with is not that the impaired human is a *human* but rather that she “is a member of a species whose typical adult members are *persons*” (2015, p. 12). Membership in a person species is important because it tells us that “even if the given individual is *not*, in fact, a person, nonetheless the individual in question *could have been* a person” (2015, p. 15). It is this ‘modal’ property that the impaired human possesses which is at play in our intuition that the interests of the human in question count for more than the comparable interests of an animal. Kagan argues that this is not speciesism because while being a member of the human species is a sufficient condition, it is not a necessary one, as he thinks that we would also give such special consideration to impaired Martians as well (provided ‘normal’ Martians turn out to be persons). So, while species membership is a morally significant fact on Kagan’s view, it is not that there is something special about being a human per se, but rather that there is something special about being a member of a “person species” (2015, p. 12).

Is ‘modal personism’ a morally acceptable view? Kagan recognises the following two problems for his account which need to be addressed. First, in what sense should we understand

‘typical’ when looking at a kind whose typical members are persons and, second, why should we think that of all the kinds one can belong to the ‘species’ that an individual is a member of is the relevant grouping to look at?

3.1. *‘Typical’ in what sense?*

To see how ‘typical’ can be ambiguous, consider the following case from Kagan. He asks us to imagine that an incurable plague sweeps across the Earth and causes almost all humans to never develop into persons with the result that humankind, the objection continues, “would cease to be a person species” (2015, p. 14). This would mean that how much consideration we give to the interests of an impaired human will depend on whether such plagues ever occur, and this seems absurd. Kagan, however, replies that this objection simply misunderstands the relevant sense of ‘typical’. He argues that the relevant question to ask is what the *generic* members of the species are like, and not the statistical question of what the majority of members are like. He says, “[T]he generic lion has hair, even if some disease leaves most lions bald.⁴⁸ Similarly, the generic human is a person – even if some disease leaves few of us that way” (2015, p. 14).

Here Kagan seems to be relying on the same understanding of ‘generic’ as those proponents of potentiality arguments discussed earlier. However, what reason is there to think that the most intuitive interpretation of such terms is the Aristotelian one? Indeed, the Aristotelian interpretation seems to have implausible implications if we extend Kagan’s plague case further and imagine something similar to McMahan’s Superchimp case where the plague has the

⁴⁸ There is a question to be asked about the truth conditions of generics in general as it seems that there is a sense in which generics such as ‘ducks lay eggs’ are true, despite the fact that less than 50% of ducks lay eggs. Yet, we simultaneously reject the truth of ‘ducks are female’ despite the fact that all ducks which lay eggs are obviously female. Similarly, we reject the truth of ‘books are paperbacks’ even though 80% of books are in fact paperbacks. I leave this aside for now but see Chapter 5 for a more detailed discussion of this.

opposite effect on a non-person species, such as chimpanzees (1996, p. 13).⁴⁹ Suppose, then, that the plague causes all chimpanzees to develop into superchimps (who are persons), so that what we would refer to as ‘normal’ chimpanzees would now be considered to be cognitively impaired and abnormal. Would Kagan claim that despite the fact that almost all chimpanzees are now persons, they are not members of a person species because the ‘generic’ chimpanzee is not a person? Presumably, Kagan will contend that even though superchimps are not members of a person species, they will still have their interests counted for more because they are nonetheless persons. But what about the interests of a superchimp born cognitively impaired? Such an individual would be what Kagan would refer to as a *generic* chimpanzee, and what the Aristotelian would refer to as a *successful* member of its kind. As this cognitively impaired superchimp is neither a person nor a member of a person species, despite the fact that most members of the chimpanzee species are now in fact persons, it can have no cause for complaint on Kagan’s account if its interests are given less consideration than cognitively similar humans who *are* members of a person species.

Suppose, now, that the plague has both of the above effects. That is, it causes most humans to never become persons, and it causes most chimpanzees to become persons. We now have a situation where the majority of members of the human species are non-persons, while the majority of members of the chimpanzee species are persons. And yet, it remains the case, on Kagan’s account, that a cognitively impaired superchimp - who has the same cognitive level as most humans – will have its interests counted for less than both the majority of members of the human species (because they belong to a person species) and the majority of the members of its own species (because they are persons). How can Kagan explain this asymmetry in moral status between the cognitively impaired human and the cognitively impaired superchimp?

⁴⁹ See Tooley (1972) for a similar case involving a kitten.

Furthermore, how will the Aristotelian deal with such a case? Here, the proponent of the potentiality account might reply that the *reason* that most humans don't become persons is significant. If it is due to some external factor blocking or impeding their potential to become persons, then the generic human is still a person. But why should that matter if the result is that most reproduction leads to the creation of non-persons, who can never become persons? At what point does the individual's set of essential potentialities no longer include becoming a person? Take the superchimps, for instance. On the Aristotelian view, which member of the kind 'chimpanzee' should we consider to constitute a 'good' or 'successful' member – the superchimp or the 'normal' chimpanzee? On one hand, if this account holds that the 'normal' chimpanzee actualises the "essential potentialities" of a good member of the kind, then it seems implausible to think that the superchimp has somehow "failed to behave as a (good) member of a kind does" (Megone, 2002, p. 161). On the other hand, if the 'essence' of a chimpanzee has now changed so that the superchimps actualise the essential potentialities of their kind, then the Aristotelian will be forced to hold that the set of essential potentialities that are actualised by a good member of a kind are subject to change. Thus, in Kagan's plague case, it would no longer be the case that 'being a person' is a part of a human's essential potentialities.

We might think that this raises an issue about the arbitrariness of the significance placed on genetics over environment by the Aristotelian. But how we turn out seems to be a product of an interaction with our genetics and our environment. We can highlight this with a science fiction example. Megone holds that while an acorn might actualise many different potentialities, such as being eaten by pigs, becoming manure, and so on, only an acorn that grows into an oak tree actualises the essential potentialities of its kind and flourishes (2002, p. 159). In some sense, though, this particular example is an easy case for the Aristotelian as it makes sense to think that only those acorns which become oak trees are good members of the kind 'acorn', because only this is associated with survival of the kind. Suppose, however, that

while an acorn is capable of growing into an oak tree, given the right set of conditions when planted on Earth, if we were to instead take an acorn to Mars and plant it there, the acorn would instead be capable of growing into a tortoise, say, rather than an oak.⁵⁰ Or imagine that due to the atmosphere on Mars, all chimpanzees born there turn out to be superchimps. How does the Aristotelian deal with such cases? Presumably, both the acorn which turns into an oak and the acorn which turns into a tortoise would have an equal claim to be instantiating the potentialities that a good member of the kind actualises, given that they can both be said to be flourishing in ways “that are conducive to reproduction and thus the survival of the species” (Megone, 2002, p. 173).⁵¹

Similarly, there appear to be cases closer to home which present a problem for the Aristotelian. Consider eusocial organisms like ants and bees. In such populations, workers and soldiers are born infertile while reproduction is carried out by a single fertile queen and her male consorts. Will the Aristotelian claim that – just like Ash and his one hand - those infertile drone members of the kind instantiate the crucial essential potentialities but fail to actualise some of the more significant ones that would be actualised by a good member of the kind? Despite the fact that such members of the kind are necessary for the survival of the kind?

3.2. *The relevant sense of ‘kind’:*

Let’s assume that there is some significance to being a member of a kind whose typical members are persons. Of all the groups that an individual belongs to, why should we think that

⁵⁰ I am grateful to Rob Lawlor for suggesting this case.

⁵¹ The neo-Aristotelian might suggest that despite the element of arbitrariness to the standards they endorse, such standards are at least rooted in the actual world, even if they start misfiring in unusual possible worlds like these. Admittedly, how much of a problem you think such cases present for the Aristotelian will depend on your tolerance for science-fiction examples and the intuitions which they pull on. I admit my tolerance is rather high, but in any case, this is not a route that most neo-Aristotelians could plausibly take given that the literature is quite rich with such examples designed to pull our intuitions in the opposite direction.

‘species’ is the important one, rather than, say, genus, class, or order, etc.? Indeed, why even think that we should be looking at natural groups rather than artificial groups like what football team one supports? Nobis argues that without some independent motivation, whichever kind we choose is “going to require an arbitrary and indefensible pick” (2004, p. 51). Levy, though, thinks that designating a kind ‘natural’ shows that the category is not arbitrary as “there is good reason to think that animals, including human beings, can be classified into kinds that ‘cut nature at its joints’” (2004, p. 215). This is because, he continues, natural kinds have “explanatory value” (2004, p. 215). Kagan, too, holds that knowing what natural group an individual is a member of usually tells us something about her *nature*, while knowledge about an individual’s artificial groups does not. Specifically, Kagan says, “what membership in a person species normally tells us is this: even if the given individual is *not*, in fact, a person, nonetheless the individual in question *could have been* a person” (2015, p. 15).

But assuming that there is something morally significant about an individual’s nature, why should we think that ‘species’ is the relevant natural group to look at in order to discover it, rather than, say, genus? Why draw a line at species rather than looking at a higher group? For Levy, the reason that we should focus on species is because it “is indisputably a natural kind if there are any in biology” (2004, p. 215). Furthermore, Kagan contends that we do not gain any extra morally relevant knowledge by looking at any of the higher biological groupings than we do by looking at an individual’s species, as the features that are characteristic of a genus, say, will also be characteristic of a species. Thus, when we look at the species an individual is a member of, we will have already taken into account any features that might be morally relevant at the level of genus or a still higher grouping. And, we should not draw the line at a lower grouping than species because if we do we might miss some of the morally relevant features which seem only to arise at the level of species (2015, p. 16). The thought seems to be that if we don’t select a wider group, like species, there’s the danger that certain important moral

information goes missing: we might fail to see that some creatures have experienced misfortune as a result of their non-possession of properties that other members of their group or community possess.

Are Levy and Kagan right to think that species is the relevant kind to tell us whether an individual possesses the property of ‘could have been a person’? That species is the relevant grouping, Kagan maintains, is reinforced with an example where generic members of a particular genus have the morally relevant property of being a person, and yet there is a particular species within this genus where typical members lack this feature. Now, consider an animal from that species, Kagan says:

Should we conclude that because this animal is a member of a ‘person genus’ it too should have its interests counted for more, in the same way that we count the interests of impaired humans for more? Intuitively, that seems like the wrong answer. The fact that members of the *genus* are typically persons seems *irrelevant*. What seems relevant, rather, is the fact that members of the *species* in question are *not* typically persons (2015, p. 16).

This example seems problematic for the following reason. It seems that we can use Kagan’s logic here and apply it to those natural groupings lower than species with troubling results. For example, imagine that for a particular species, the generic member of the species is a person, and yet there is a particular sex within this species where typical members of that sex are not persons. For example, in A. M. Lightner’s (1970) *The Day of the Drones*, the protagonist encounters a matriarchal society modelled on bee society where the typical member of the female sex is a person, while the typical member of the male sex is a non-person. Indeed, those few males who become persons are executed as ‘defective’. Now, using Kagan’s above argument, consider a male from that species:

Should we conclude that because this individual is a member of a ‘person species’ he too should have his interests counted for more, in the same way that we count the interests of impaired humans for more? Intuitively, that seems like the wrong answer. The fact that members of the *species* are typically persons seems *irrelevant*. What seems relevant, rather, is the fact that members of the *sex* in question are *not* typically persons.

This seems like an objectionable outcome. Surely we would want to hold that the males in this society should have their interests counted for as much as their female counterparts. Further, it seems that we can also imagine similar science fiction cases where we can apply the same argument to other groupings, such as race.

However, we might wonder whether, despite seeming *prima facie* unintuitive, there is anything morally objectionable about giving less weight to the interests of the non-person sex over the interests of the female person sex. After all, if most members of that sex never become persons then is there any reason to think that they should have their interests counted for as much as those of persons? Here it is important to keep in mind the position that Kagan is attempting to defend. Like Oderberg, Kagan is attempting to justify the Common Sense Intuition, that the interests of humans count for more than the interests of animals. However, unlike Oderberg, Kagan is not arguing that all human interests outweigh *all* animal interests. Rather, Kagan simply wants to defend *some* preference for human interests over those of animals. This preference could take a number of different forms but Kagan is clear that he thinks that, even if successful, his account would not justify many of our current treatments of animals (2015, p. 2). So, if Kagan has only a minimal preference in mind, is there anything wrong with counting the interests of the male non-person sex less than those of the female person sex? One reason for thinking that there *is* something objectionable here becomes clear when we take the same strategy from the previous section and imagine both a cognitively impaired female and a supermale. Kagan’s account would have it that they are both entitled to

have their interests counted for more than those normal non-person males. But, again, it seems implausible to think that we should give greater weight to the interests of the cognitively impaired female than those of a normal male when the only seemingly morally relevant difference between the two is that the female belongs to a person-sex while the male does not. Furthermore, we might think that if there is nothing problematic with giving less consideration to the interests of the non-person sex then we should also hold that it is unproblematic to give less consideration to the interests of human non-persons, a position Kagan clearly wants to avoid.

4. Conclusion

This chapter presented a set of problems to be overcome if we are to justify the ‘Common Sense Intuition’: that the interests of humans count for more than those of non-human animals. I began this chapter by stating what I take to be a relatively uncontroversial claim: that one constraint on accounts which seek to offer a successful defence of the intuition that humans matter more than animals must offer a justification which does not leave themselves exposed to parallel defences of malicious forms of prejudice such as racism and sexism. Such accounts, I have argued, can be separated into two main camps: those who attempt to offer a defence of speciesism and those who attempt to justify the Common Sense Intuition while avoiding the charge of speciesism. Considering these in turn, I maintain that both types of arguments face the same variety of problems: they both lead to holding unintuitive conclusions when we consider cases which deviate from the norm. While I have argued that the proposed solutions to these problems have so far been unsuccessful, and have not volunteered any solutions myself, identifying the problems involved in a successful defence of the Common Sense Intuition, and what is required to defeat such problems, is a worthwhile exercise as it provides us with a useful roadmap towards a successful defence of some form of speciesism.

Works Cited

- Burgess, J. A., 2010. Potential and Foetal Value. *Journal of Applied Philosophy*, 27(2), pp. 140-153.
- Cohen, C., 1986. The Case for the Use of Animals in Biomedical Research. *New England Journal of Medicine*, Volume 315, pp. 865-869.
- Kagan, S., 2015. What's Wrong with Specieism?. *Journal of Applied Philosophy*, Volume (Early Online Edition), pp. 1-21.
- Lang, G., 2012a. Discrimination, Partial Concern, and Arbitrariness. In: U. Heuer & G. Lang, eds. *Luck, Value, and Commitment: Themes from the Ethics of Bernard Williams*. Oxford: Oxford University Press, pp. 293-332.
- Lang, G., 2012b. Is There Potential in Potential. *Philosophical Papers*, 41(1), pp. 129-147.
- Lawlor, R., 2012. The Ethical Treatment of Animals: The Moral Significance of Darwin's Theory. In: M. Brinkworth & F. Weinert, eds. *Evolution 2.0: Implications of Darwinism in Philosophy and the Social and Natural Sciences*. New York: Springer, pp. 147-163.
- Leslie, S. J., 2012. Generics. In: G. Russell & D. G. Fara, eds. *The Routledge Companion to Philosophy of Language*. New York: Routledge, pp. 355-367.
- Levy, N., 2004. Cohen and Kinds: A Response to Nathan Nobis. *Journal of Applied Philosophy*, 21(2), pp. 213-217.
- Liao, S., 2012. The Genetic Account of Moral Status: A Defence. *Journal of Moral Philosophy*, Volume 9, pp. 265-277.
- Lightner, A. M., 1970. *The Day of the Drones*. New York: Bantam Books.
- McMahan, J., 1996. Cognitive Disability, Misfortune and Justice. *Philosophy and Public Affairs*, 25(1), pp. 3-35.
- Megone, C., 2002. Potentiality and Persons: An Aristotelian Perspective. In: M. G. Kuczewski & R. Polansky, eds. *Bioethics: Ancient Themes in Contemporary Issues*. Cambridge: MIT Press, pp. 155-177.
- Nobis, N., 2004. Carl Cohen's 'Kind' Arguments For Animal Rights and Against Human Rights. *Journal of Applied Philosophy*, 21(1), pp. 43-49.
- Oderberg, D. S., 2000a. *Moral Theory: A Non-Consequentialist Approach*. Oxford: Blackwell Publishing.
- Oderberg, D. S., 2000b. *Applied Ethics: A Non-Consequentialist Approach*. Oxford: Blackwell Publishing.

- Singer, P., 1993. *Practical Ethics*. 2nd ed. New York: Cambridge University Press.
- Singer, P., 2002. *Unsanctifying Human Life*. Oxford: Blackwell.
- Singer, P., 2009. *Animal Liberation*. Updated ed. New York: Harper Collins.
- Singer, P., 2015. Why Speciesism is Wrong: A Response to Kagan. *Journal of Applied Philosophy*, Volume (Early Online Version), pp. 1-5.
- Stone, J., 1994. Why Potentiality Still Matters. *Canadian Journal of Philosophy*, 24(2), pp. 281-293.
- Tooley, M., 1972. Abortion and Infanticide. *Philosophy and Public Affairs*, 2(1), pp. 32-65.

Chapter 5

Humans are Persons or: How I Learned to Stop Worrying and Love the Generic

“I would’ve gotten away with it if it hadn’t been for those meddling [intuitions]”

-- Various Scooby Doo villains

Introduction

Intuitively, statements such as ‘Humans are persons’ and ‘The generic human is a person’ seem to be true. Indeed, the truth of these statements appears to be quite robust. That is, such statements appear to be true even in imagined cases where the majority of humans are not in fact persons. For example, consider the following scenario presented by Shelly Kagan (also discussed in chapter 4). He asks us to imagine that an incurable plague sweeps across the Earth and causes almost all humans to never develop into persons (2015, p. 14). Even here, despite the fact that most humans will never become persons, ‘humans are persons’ still seems to be true. As Kagan says, “the generic lion has hair, even if some disease leaves most lions bald. Similarly, the generic human is a person – even if some disease leaves few of us that way” (2015, p. 14). But what explains the truth of such statements? Statements like ‘humans are persons’ are what linguists refer to as ‘generics’ – sentences which express a generalisation. Much space in the philosophy of language literature has been devoted to the debate over what way the world has to be in order for a generic to be considered true - the generic’s truth conditions. This literature has been enriched by Sarah Jane Leslie’s (2008; 2012) psychological treatment of generics and their truth conditions.

In what follows, I set out the components of this account of generics in order to address the following worry: if the argumentative work presented in earlier chapters is as persuasive as I suggest, why then do I still find it difficult to shake off the common-sense intuition that there is something morally significant about humankind and that, for example, performing medical experiments on severely cognitively disabled humans is worse than performing similar experiments on non-human animals? For part of this thesis I have attempted to break what appears to be an entrenched stalemate between two opposing positions within the animal ethics literature on the moral significance of species membership. On one side of the debate, is an Aristotelian position, defended by David Oderberg (2000) and Chris Megone (2002), which holds that there is something morally significant about humankind. On the other side, is the position most often attributed to Peter Singer (1993), who puts forward compelling arguments suggesting that what species an individual belongs to holds no moral significance. For Singer, what actually matters are individual capacities.

Like many others, I encounter a strong clash of intuitions here. While I feel the intuitive force of the arguments put forward by Singer and others, and have attempted to defend them, I find that the common-sense intuition - that there is something morally significant about humankind - persists, even after accepting the force of the arguments presented in this thesis. This is a troubling outcome. Why should each new counterexample, or thought experiment, presented by those who resist the arguments of Singer and others, cause so much trouble by pulling on our intuitions despite powerful arguments discrediting the examples? Indeed, such nagging intuitions appear to retain a foot in the door even after acceptance of the persuasive arguments against these examples. If the arguments put forward in this thesis are as powerful as I and others have claimed, shouldn't I be able to sit back and put my feet up, guilt-free, content in the knowledge that the important legwork has been done? Clearly, the problem is not the counter-examples presented (new or old) - many of which appear to appeal to the same

intuition. Rather the problem is the common-sense intuition which refuses to die no matter how many times we throw it into the wood-chipper. The obvious explanation is that it refuses to disappear because the intuition is pointing to some deeper truth; that there really is something morally significant about humankind. The neo-Aristotelian has an answer for the persistence of this intuition: that animal kinds have an ‘essence’ and, as a part of their essential nature, humans have the property of personhood. Thus, even if some humans aren’t persons and will never become persons, they are still ‘potential’ persons and it is this fact which explains why the common-sense intuition and others like it persist in the face of persuasive arguments discrediting them. Given that the arguments presented in this thesis, among others, have failed to deal the requisite killer-blow to this pesky intuition, I can either continue to live in fear of the next clever thought-experiment which leans on it, accept the Aristotelian’s explanation and their conclusion, or offer an alternative explanation for why these nagging intuitions persist despite acceptance of arguments which discredit the thought-experiments. The chapter takes the last of these positions, aiming to explain why the common-sense intuition might remain despite the intuitive force of arguments by Singer and others. The generics explanation offered here, then, should be seen as an alternative explanation for the persistence of the intuitions that neo-Aristotelians frequently appeal to.⁵²

The chapter will proceed as follows. The first section will set up the argument by detailing the role played by generics within the neo-Aristotelian’s account. The second section will outline Leslie’s account of generics, while the third section will show how this account offers an alternative explanation for the intuitive force of examples employed by the neo-Aristotelian

⁵² It is important to note that the Aristotelian relies on many other arguments for their position, some of which have been addressed in earlier chapters. My purpose in this chapter is not to engage directly with any one of the Aristotelian’s arguments, or to attempt to defeat them, but rather to simply attempt to explain why this nagging intuition persists even when such arguments are engaged with directly.

which is consistent with the arguments of Singer and those presented in earlier chapters and, therefore, consistent with a rejection of the common-sense intuition.

1. *Essence and Generics*

To understand the relevance of Leslie's account to this debate, it is important to be clear on the role that generics play within the neo-Aristotelian's argument. Those who defend an Aristotelian account of the thought that there is something special about humankind, hold that animal kinds have an "essence" (Oderberg 2007, p.49), or a set of "essential potentialities" (Megone 2002, p. 163). These essences explain the changes that 'successful' members of such kinds go through. Humans have, they contend, as a part of their essential nature, the property of personhood. Thus, even if some humans aren't persons and will never become persons, they are still 'potential' persons or, in Aristotelian terms, 'persons'.⁵³ That is, all humans are persons because they share the same essential nature. For example, Oderberg claims that the "kind of thing an entity is determines its potentialities" (2000, p. 178) and that while "some of us might only be potential *persons*, we are still actual persons, because we are human beings" (2000a, p. 182). For Oderberg, '*persons*' refers to Singer's sense of personhood, that is, those individuals who actualise the property of being a person, while '*persons*' denotes *all* humans, including those who fail to instantiate the property. Megone, like Oderberg, holds that all humans share, as their essential nature, a "set of essential potentialities" (2002, p. 163) which may or may not be actualised. As such, a severely cognitively impaired infant and a normal adult human "instantiate the same essential potentialities in the sense that there is a set of potentialities, which a perfect human actualises, which each is in the process of actualising,

⁵³ Here, as discussed in chapter 4, the Aristotelian is employing a broader notion of 'potentiality' than the everyday sense of the term.

with greater or lesser success” (Megone, 2002, p. 162). If this is right and those humans who are not in fact actual persons are nevertheless to be considered persons regardless of degree of maturity, we can’t do to young members of this type what we can’t do to mature members of this type. What matters isn’t the degree of maturity, but the fact that, whether young or mature, it’s a person we’re dealing with.⁵⁴

Why, though, should we think that personhood is a part of the essential nature of humans given that many humans do not in fact actualise this property? Here the Aristotelian often appeals to intuitive examples which make use of the truth of generics to support her case. Phillipa Foot and Michael Thomson explicitly appeal to our acceptance of the truth of certain generic statements, or what they refer to as “Aristotelian categoricals” (Thomson 1995, p. 267; Foot 2001, p. 29). For example, Foot holds that we can derive from the truth of the generic statement ‘deer are swift of foot’, that an individual deer which does not possess the property predicated in the statement is somehow “defective” (Foot 2001, p. 34). Others implicitly appeal to our intuitive acceptance of generics through the use of intuitive examples. For instance, Megone holds that even if most acorns do not grow into oak trees, it remains the case that acorns become oak trees (1998, p. 193; 2002, p. 162). Similarly, Oderberg claims that “a tadpole is a potential frog” (2000, p. 178) and, thus, even if most tadpoles get eaten by fish, it is still the case that tadpoles become frogs. Such examples are also employed by non-Aristotelian thinkers who wish to defend the significance of species-membership. For example, Kagan holds that, as members of a “person species” – a species whose typical adult members are persons - the interests of humans count for more than the interests of members of non-person species (2015, p. 12). It is not the fact that most humans are in fact persons, he argues, which tells us that humans belong to a person-species. The plague case highlights that

⁵⁴ Of course, this relies on the assumption that personhood is a morally relevant property, but I take this to be relatively uncontroversial.

it cannot be a matter of statistics. Rather, it is the fact that “the generic human is a person” (Kagan, 2015, p. 14) that tells us that humans are members of a person species. Given, then, that the truth of these generic statements seem to be doing some of the heavy lifting in both the Aristotelian’s and Kagan’s arguments, it is important to ask: how do we judge the truth of such statements? That is, what are the truth conditions of the statements, ‘Acorns become oak trees’, ‘Tadpoles become frogs’, and ‘Humans are persons’? Why should we think that such statements are true even when most of the members of the species in question do not in fact go on to actualise the specified property? Furthermore, if such statements are true, is there any reason to think them morally significant?

If such examples are used to support the Aristotelian’s position, then it matters *why* we judge them, rather than the converse statements, to be true. For the Aristotelian, it is the metaphysical fact that acorns have, as their essential nature, the property of becoming an oak tree, that explains why the claim ‘Acorns become oaks’ is accepted as true.⁵⁵ In what follows, I will argue that Leslie’s account of the truth conditions of generics can provide us with a convincing alternative explanation of the truth of ‘Humans are persons’ and ‘Acorns become oaks’, without having to posit any metaphysical ‘essences’. Instead, this explanation relies on evolutionary mechanisms within human cognition and is therefore a more parsimonious explanation. Leslie’s account of generic statements suggests that generics express an innate “basic generalization mechanism” which stems from inductive learning and has evolved to be as efficient as possible (2008, p. 28-9). I argue that if the truth conditions of such statements are merely *psychological* in nature, as Leslie’s account suggests - rather than *metaphysical*, as the Aristotelian claims - then this casts significant doubt on the role played by the truth of

⁵⁵ Whether the metaphysical fact should be imbued with the moral significance placed on it by the Aristotelian is a separate question which I lack the space to deal with here effectively. For some reasons to be sceptical, see Tooley (1972); Singer (1993); Nobis (2004).

statements like ‘humans are persons’ in the Aristotelian’s argument. This, then, in turn, casts doubt on the moral significance of these claims. In what follows, I steer away from presenting the generics account offered in this chapter as an argument to the best explanation and thus a rejection of the neo-Aristotelian’s explanation that what best accounts for the truth of ‘humans are persons’ is that all humans have, as a part of their essential nature, the property of personhood. This is a stronger claim than I wish to make here. Rather, my purpose here is to offer an alternative explanation for the intuitive force of the examples that rely on the intuition that humans are persons which is in turn consistent with arguments that reject attributing a moral significance of humankind. Thus, while I have provided some reasons in earlier chapters to doubt that personhood is part of human essence, I will not be engaging directly with any one of the neo-Aristotelian’s arguments and will thus provide little reason, in this chapter, to think the neo-Aristotelian account is mistaken. That said, while not decisive in the debate, neither does the generics account have a negligible impact. What the account does perhaps show is that it is not simply enough, for the plausibility of the neo-Aristotelian’s account, that we have these intuitions when considering various thought-experiments. The fact that an alternative, and perhaps more plausible, explanation for these intuitions exists shows that a mere reliance on such intuitions cannot provide enough support for their arguments. While this does not remove the opportunity to appeal to thought-experiments which pull on this intuition, such as the plague case, it calls into question the purpose of doing so. Such examples are used because of their troublesome nature; that is, the fact that they are still considered to be true in the face of many counterinstances, is employed by the neo-Aristotelian to provide support for the claim that natural kinds have, as their intrinsic nature, an essential set of potentialities which may or may not be actualised (Megone, 2002, p. 162). The generics account, by positing a psychological explanation, rather than a metaphysical one, for why we judge ‘Humans are

persons' to be true, questions the effectiveness of using these intuitive examples as support for the Aristotelian's argument.

That said, the generics account does seem to provide some measure of support for the advocates of Singer's position in that it allows us to explain away these nagging intuitions which remain even after accepting the force of his arguments in a way that is consistent with those arguments. The extent to which the generics argument shifts the balance in the stalemate away from the neo-Aristotelian position and towards Singer's position on the moral significance, or lack thereof, of species membership, is open to debate. However, it is possible that one benefit of this account is that it fills a gap left by those, like Rob Lawlor (2012), who put forward compelling rejections of essentialism but who, unlike Singer, still recognise the force of common-sense intuitions that, for instance, performing experiments on severely cognitively impaired humans is worse than performing similar experiments on animals. The generics argument fills this gap by offering an account which is able to square our common-sense intuitions with a rejection of essentialism.

In summary, then, I have presented some reasons in the previous chapters to doubt both that species membership is significant and that personhood is part of human essence (essentialism). However, it is clear that despite such arguments, the intuitive force of the examples used by neo-Aristotelians stubbornly persists. Thus, without an explanation of this persistence which adheres to Singer's position, it seems as though the stalemate between his and the neo-Aristotelian's account also remains. But the significance of this chapter is to show that the advocates of Singer's arguments are at an advantage: in addition to the arguments presented throughout this thesis, the generics account *can* explain the intuitive force of the examples in a way that is consistent with a rejection of both essentialism and the significance of species membership.

2. *Generics and their Truth Conditions*

The truth conditions of many generics might seem to be quite straightforward. For instance, ‘Dogs have four legs’ does not at first appear to present us with any difficulty - even if some dogs only have three legs, most dogs have four. The fact that a majority of the members of a given kind satisfy the property predicated in the generic seems, in such cases, to be enough for the truth of that generic. However, many generics are what Leslie refers to as “troublesome generics”. Why is it, for example, that the generic ‘Ducks lay eggs’ is true, but ‘Ducks are female’ is not, despite the fact that all those ducks that lay eggs are obviously female? Why is the claim ‘Mosquitoes carry the West Nile Virus’ true when only 1 percent of mosquitoes actually carry this virus, while ‘Books are paperbacks’ is false when more than 80 percent of books are paperbacks (Leslie, 2008, p. 2; 2012, p. 355)? These examples suggest that it is not necessary, or even sufficient, for the truth of a generic that most of the kind under consideration possess the given property. They also indicate that other classical quantifiers such as ‘all’ or ‘some’ will either be too strong or too weak and, as such, cannot account for the meaning of generics. Similarly, Leslie argues, adverbs of quantification such as ‘typically’ and ‘generally’ cannot account for the truth of generics either. For example, while ‘Books are paperbacks’ is false, ‘Typically, books are paperbacks’ is true (2008, p. 7).

Troublesome generics, like those mentioned above, are either generally considered to be true⁵⁶, despite the majority of the kind in question failing to possess the given property, or false, despite the majority of the kind possessing it (Leslie, 2008, p. 13). According to Leslie, such

⁵⁶ A quick note on terminology. For cases of troublesome generics, we might be tempted to employ descriptions of their truth conditions which hedge our bets. That is, we might prefer descriptions which include ‘intuitively true’, ‘seems/appears to be true’, rather than ‘is true’. However, as the truth ascriptions made here are merely psychological in nature, such descriptions appear to be interchangeable; the generic statements remain true even when we can point to counterinstances of, say, male ducks. It may therefore be superfluous, for the purposes of this discussion to use such language when ascribing truth to generic statements. As such, for the remainder of the chapter I will continue to use the simple ‘is true’.

generics reflect our psychology, with two factors in particular influencing our judgements about generics. The first is the background expectations that we associate with kinds, and the second is the nature of the counterinstances to the generalisation (2008, p. 37). The next section will explain these in turn.

First, though, the following bears mentioning. While most troublesome generics are explained by these two factors, there is another category of troublesome generics that are not explained by this, which I will not discuss in detail as it is not relevant to this debate. According to Leslie (2008; Forthcoming), this category includes sentences like ‘Mosquitoes carry the West Nile virus’ and ‘Pit bulls maul children’. Here, the kind in question is attributed a harmful, dangerous, or striking property; the type of property that, in evolutionary terms, it would be in our best interests to know about, even if our chance of encountering it is minimal. Leslie maintains that this phenomenon also appears in our other judgements: “compare the number and regularity of times one must worry to be a worrier versus the number of murders one must commit to be a murderer” (2008, p. 15). This category of troublesome generics is remarkably tolerant to exceptions, needing only a few members of a kind to express the dangerous property for us to form the generalisation. Moreover, Leslie proposes that the same mechanism is at work in both our judgements of statements like ‘Mosquitoes carry the West Nile Virus’ and prejudiced statements like ‘Muslims are terrorists’ (Forthcoming). However, she argues that while the former statement is true, the latter is false because there is a mistake involved when we judge that Muslims are terrorists, whereas there is none when we judge that mosquitoes carry the virus. The mistake, Leslie suggests, lies in the fact that generics which predicate striking properties are true “if and only if some members of the kind in question possess the relevant property, and the others are typically disposed to possess it” (Forthcoming). Thus, the truth of ‘Mosquitoes carry the West Nile virus’ depends on whether or not those mosquitoes that do not carry the virus would carry it if circumstances allow.

Leslie's account is more complicated than I have suggested here, but I leave such cases aside in this chapter, turning instead to those troublesome generics which are explained by the following two psychological factors: the background expectations that we associate with kinds and the nature of the counterinstances to the generalisation.⁵⁷

3. Characteristic Dimensions of a Kind

Leslie holds that troublesome generics like 'ducks lay eggs' and 'chickens are females'⁵⁸ are true and false respectively because of the background knowledge that we have which causes us to expect certain things of them. For example, we expect that if an animal kind is to survive, it must exhibit certain characteristics. Among these "kind relative expectations", Leslie suggests, are:

There must be both male and female members of the species

There must be a manner in which this reproduction and subsequent gestation occurs (2008, p. 14).

Generics which express these kind-relative expectations, which Leslie refers to as "characteristic dimensions of a kind", are true in virtue of our background assumptions about the kind in question being a successful biological one (2008, p. 16). That is, we assume that if an animal kind is to be a successful one and survive, it must display certain characteristics, such as a method of reproduction, mode of locomotion, diet, etc. Generics which express these characteristic dimensions are still judged true even when the number of exceptions is large. Thus, generics such as 'ducks lay eggs' are true in spite of the large number of exceptions.

⁵⁷ See Leslie (Forthcoming) for an interesting, informed, and detailed discussion on how generics play a role in forming such prejudices, as well as suggestions for how we might work to combat prejudice formation. See also Haslanger (2011) for a discussion of how Leslie's account of generics can be used to suggest pragmatic reasons for rejecting the truth of prejudiced generics.

⁵⁸ Approximately 80 percent of chickens are hens (Leslie, 2008, p. 10).

Leslie contends that the number of exceptions required for us to give up on the truth of these generics needs to be very great, almost 100 percent in fact. This makes such statements remarkably tolerant of exceptions. This explains why, she maintains, those generics which conflict with our background assumptions appear false, such as ‘bees are sterile’ (2008, p. 14).⁵⁹ This is because the background expectations that we associate with animal kinds are often related to reproduction, as reproductive features lie along characteristic dimensions for animal kinds (2008, p. 31). However, when almost 100 percent of the kind in question display the given property, such as ‘mules are sterile’,⁶⁰ then the generic statement is considered true (Leslie, 2008, p. 15). Thus, Leslie thinks that the “relation between objective statistical frequencies and the truth-values of generics is in some way mediated by psychological considerations” (2008, p. 16).

According to Leslie, this is because our cognitive system’s innate default method of generalizing – how a property held by a member of a kind is generalised to the entire kind - is expressed by generics. This is a basic information gathering mechanism which needs to be as efficient as possible. One way that this mechanism increases its efficiency, Leslie suggests, is by taking advantage of the many regularities in our environment by “drawing inferences from particular instances of a category to novel and unobserved ones” (2008, p. 29).⁶¹ Thus, we can think of this mechanism as sorting the world into groups of things which seem to us to behave in similar ways. This provides young children with information as fast as possible. For

⁵⁹ Only a small minority of bees are in fact capable of reproducing (the queen and her consorts), and only the queen is capable of laying eggs.

⁶⁰ Mules have 63 chromosomes, as a result of being the offspring of a male donkey (62 chromosomes) and a female horse (64 chromosomes). This odd number of chromosomes prevents them from dividing into chromosome pairs thus rendering most mules infertile (although there have been extremely rare documented cases of female mules producing offspring when mated with male donkeys or horses: <http://www.denverpost.com/2007/07/25/mules-foal-fools-genetics-with-impossible-birth/>).

⁶¹ Evidence from experimental tasks suggests that children do infer regularities about novel categories by making use of information from similar categories (Macario, et al., 1990).

instance, many animal kinds are alike in a variety of ways; they all have a “characteristic diet, characteristic noises, modes of locomotion”, etc. (Leslie 2008, p. 29). An information gathering system evolved to be as efficient as possible is able to prevent the learner from having to constantly relearn basic facts about animal kinds by exploiting such regularities. This powerful mechanism draws robust inferences from only a few instances, enabling inductive learning by allowing us to gather information from minimal experience. Thus, a learner is able to ‘filter out’ less important pieces of information by having “a prior notion of what the characteristic dimensions of a kind are”, thus allowing her to concentrate “on finding values for *those* dimensions” (2008, p. 31).

What is the source of these characteristic dimensions? In the psychological literature, such generalisation of kinds is referred to as ‘essentialism’ (Gelman, 2003). A learner essentialises a kind if she forms the implicit belief that there is “some hidden, non-obvious, and persistent...underlying nature shared by members for that kind, which causally grounds their common properties and dispositions” (Leslie, Forthcoming). For instance, a learner might form the belief that there is something about dogs which causes them to bark, to have four legs, etc., and that these are intrinsic features of dogs, grounded in the very nature of doghood. As such, a learner who essentialises this kind, tacitly believes that even a dog born with three legs possesses this ‘essence’, even if it does not actualise it. We believe (implicitly) that this ‘essential’ nature of dogs “causally grounds these dispositions, though does not guarantee their manifestation, since adventitious factors may intervene” (Leslie, Forthcoming). It should be stressed here that this is merely a psychological claim about our folk theoretical notions of biological kinds rather than a metaphysical claim. As such, the notion of ‘essence’ used here is distinct from the Aristotelian’s metaphysical notion mentioned above. The psychological claim is that we (implicitly) ‘essentialise’ characteristic dimensions rather than there truly *being* an essence as the Aristotelian thinks of it. We might be tempted here to view this account

as an error theory. However, Leslie is not claiming that we make a mistake when we essentialise these properties (at least not in the cases under consideration in this chapter). Indeed, this process aids us in inductive learning. Even in troublesome cases, we do not want to claim that the generics are false. In fact, we are beginning with the fact that they are true and are then trying to explain why they are true, rather than trying to show that they are false, and we only think that they are true. That said, while the account is not an error theory for the psychological claim, we might perhaps view it as an error theory for the metaphysical claim made by the Aristotelian. That is, the Aristotelian is making a mistake when they impute moral and metaphysical significance to what may be a purely psychological claim.

Leslie proposes that, from a young age, we are “innately disposed to seek out characteristic dimensions for kinds” as these dimensions supply learners with an “informational template” upon encountering a new kind (2008, p. 32). Given the importance of characteristic dimensions in our basic generalizations then, it is not surprising that they are also highly influential in how we judge the truth of generics. Leslie continues:

When a value is found for a characteristic dimension of a kind, it is thereby generalised to the kind by the basic generalisation mechanism, and so the generic that predicates that property of the kind is accepted. (Leslie 2008, p. 32).

Thus, when we consider ducks, we seek a value for their method of reproduction. We know that ducks reproduce by laying eggs and so we generalise that property to the entire kind and this leads us to accept ‘Ducks lay eggs’ as true (Leslie 2008, p. 33).

4. Nature of Counterinstances

Why is it that ‘Ducks lay eggs’ tolerates male ducks as an exception but ‘Ducks are females’ does not, and why does the latter generic seem to require that *all* ducks be female for its truth, while it is not required that *all* ducks lay eggs for the former generic to be considered true?

Leslie suggests that it matters whether the exceptions are positive rather than negative. A generic has a positive exception, she holds, when a member of the kind in question “has a concrete alternative property”, that is, when it has a *salient alternative* property to the one predicated in the generic. A generic has a negative counterinstance, on the other hand, when a member of the kind in question simply fails to possess the given property (Leslie 2008, p. 33-4). Take ‘Ducks are female’. The counterinstances to this generic are positive ones as those ducks which are not female possess the alternative positive property of ‘being male’. Whereas, those ducks which do not lay eggs simply fail to possess the property rather than possessing a salient alternative property such as, for example, ‘bearing live young’ (Leslie 2008, p. 34).⁶² Thus, to count as a positive counterinstance and to be considered salient, the property in question needs to be an alternative to the one predicated in the generic as well as being *either* equally striking *or* lying along characteristic dimensions for that kind.

The relevance of this distinction, Leslie continues, is that those counterinstances which are negative are more likely than positive ones to be tolerated by generics. We can see this with the puzzle of the truth of ‘Ducks lay eggs’ and the falsity of ‘Ducks are female’. The latter generic requires that we overlook positive counterinstances while the former does not. Thus, Leslie suggests that “there is an intuitive difference between simply lacking a feature and lacking it in virtue of having another, equally memorable, feature instead” (2008, p. 35). Therefore, for those members of the kind who lack the property in question, it matters *how* they lack it. Leslie demonstrates this with a simple thought experiment. ‘Peacocks have bright blue tails’ seems to be true. One reason for this, Leslie contends, is that the female peacocks which fail to possess this property are often drab and unremarkable compared with the males.

⁶² It is worth noting that, again, this distinction between positive and negative counterinstances is meant as a psychological distinction, rather than a metaphysical one. What is important here is whether we *perceive* the exceptions to be negative or positive (Leslie 2008, p. 34).

However, imagine that female peacocks have bright pink tails instead of the brown stumps that they do in fact have. Leslie claims that it is plausible that we would no longer judge ‘peacocks have bright blue tails’ as true, because it is more difficult to overlook such a salient and striking feature (Leslie, 2008, p. 35). She thus proposes that we are unlikely to judge a generic to be true if those members of the kind in question who fail to have the property predicated in the generic do so “in virtue of having an equally salient, memorable, and striking feature” (2008, p. 36). By contrast, we are much more likely to accept the generic as true if they fail to possess the property in an unremarkable way, such as in the peacock case. This helps to explain why generics such as ‘Books are paperbacks’ are considered false despite the majority of the kind possessing the relevant property, while generics like ‘Ducks lay eggs’ are considered true even when many ducks do not possess the property of laying eggs. Thus, if positive counterinstances exist, there must be a very high majority of the kind that possess the relevant property in order for the generic to be considered true. Indeed, “we would need approximately universal satisfaction of the predicate for those to be true” (Leslie, 2008, p. 36).

Using both of these two factors which influence our judgements about the truth of generics, Leslie suggests the following hypothesis:

The inductive learning mechanism is strongly biased towards filling in only positive values for characteristic dimensions, and once it has done so, it readily retains the value in the face of negative counterinstances (2008, p. 38).

Consider ‘Bees lay eggs’. In spite of the fact that almost all bees do not possess this property, being an animal kind, we seek a value for their reproductive mechanism. But as there is no alternative salient property possessed by those bees that do not lay eggs, “we seize on the only positive instances available and conclude that bees lay eggs” (Leslie, 2008, p. 38). Because all of the exceptions are negative, the generic is both formed and retained despite the high number of counterinstances.

4.1. *The Generics Debate*

Before moving on to the relevance of generics to the Aristotelian's argument, we need to note the following. While Leslie's psychological account is intuitively plausible, it is not entirely uncontroversial and there are competing accounts of the truth conditions of generics. One account in particular would seem to provide some support for the Aristotelian position. This account suggests a possible worlds approach, where generics tell us what is 'normal' for a kind, where 'normal for a kind' here represents the changes that a member of that kind would go through should nothing external hinder its development (Pelletier & Asher, 1997).⁶³ On this approach, while the actual world might be quite 'abnormal' for the kind in question, we should consider a possible world where life goes as normally as possible for that kind and then ask whether a member of that kind would have the property predicated in the generic under these ideal circumstances. Thus, while a dog might have a birth defect causing it to only have three legs in the actual world, a dog born under the circumstances 'most normal' for dogs will have four legs. While it is far beyond the scope of this chapter to fully address the concerns with competing accounts of generics, there seem to be *prima facie* reasons for thinking that this account is implausible.⁶⁴ For instance, gender-specific generics, such as 'Ducks lay eggs', seem to present a significant problem for this account as there does not seem to be anything abnormal about being a male duck (Leslie, 2008, p. 8). Furthermore, generics such as 'Mosquitoes carry the West Nile virus' appear to be even more problematic as it is difficult to see how it could be *normal* for mosquitoes to carry this virus (Leslie, 2012, p. 360). Leslie's

⁶³ Another account suggests that the truth conditions of generics are explained by *probabilistic* factors, where the property predicated in the generic is either more likely to hold of members of that kind than of other kinds, or, for any member of that kind, it is more likely than not that they will possess the property in question (Cohen, 1996). In addition to various problems with this account (see Leslie 2008; 2012), it is difficult to see what support it could provide the Aristotelian.

⁶⁴ See Leslie (2008; 2012) for a more in-depth discussion of this.

account, though, does not have any difficulty accounting for, and explaining the truth of, the troublesome generics just mentioned, nor does it appear to face any counterexamples of its own (Liebesman, 2011, p. 8). Moreover, as I will argue in the final section, the fact that Leslie's account seems to adhere so well with our competing intuitions in the speciesism debate, suggests further support for her account. In any case, my intention in this chapter is not to argue that Leslie's account of generics is the correct one, but simply to lay out the components of one convincing account of generics in the hope of shedding some new light on a debate within the animal ethics literature.

5. Essence and Personhood

Having presented Leslie's psychological account of generics, it should now be clear how this account explains why we consider the statement 'Humans are persons', and other examples put forward by the Aristotelian, such as 'Acorns become oak trees', to be true. Leslie proposes that for a generic to be considered true, the following circumstances must hold: *either* the property predicated in the generic is striking or lies along a characteristic dimension for that kind, and the counterinstances to the generic are negative ones - in which case it is sufficient for the truth of the generic that only "*some* members of the kind satisfy the predicate" (2008, p. 16); *or*, it is the case that almost 100% of members of the kind satisfy the predicate (2008, p. 43). It is plausible to think that one of the kind-relative expectations that we have for humankind is *personhood*. This feature lies along characteristic dimensions for humankind and, thus, we essentialise it as an intrinsic feature grounded in the nature of humankind. (Again, the claim that we essentialise personhood to the kind 'human' describes a psychological mechanism and is thus distinct from the Aristotelian's metaphysical claim that there truly *is* such an essence). If so, then it is not surprising that the claim 'humans are persons' is so robust. Indeed, if what Leslie describes about those features which lie along a kind's characteristic

dimensions is true, we would expect that the number of exceptions to the generic for it instead to be judged false would need to approach almost 100%. Therefore, because personhood lies along the characteristic dimensions for humankind, ‘humans are persons’ is still considered true even in the plague case where almost all humans no longer become persons.

Furthermore, even in the plague case, all of the counterinstances to ‘Humans are persons’ are negative ones. That is, those humans who do not possess the property of being a person, “simply fail to do so” (Leslie, 2008, p. 34) without possessing a salient alternative property. More specifically, ‘Humans are persons’ is tolerant of exceptions, even in the cases where there are a large number of exceptions, because non-persons fail to be persons without possessing an alternative property which lies along the perceived characteristic dimensions for humans. As Leslie writes, “generics that pertain to characteristic dimensions can tolerate any number of exceptions as long as they are negative”, as is the case with the claim ‘Bees lay eggs’ (Leslie 2008, p. 39). This also explains why the generic ‘Humans are non-persons’ is still judged to be false, even in the plague case where the majority of humans are in fact non-persons, in the same way that we judge ‘Bees are sterile’ to be false. This is because the property of ‘being a person’ is a positive counterinstance to the generic, that is, it is a “salient, memorable, and striking feature” which lies along the characteristic dimensions of the kind (Leslie, 2008, p. 36). When even a small number of positive counterinstances exist, that is, those exceptions which are salient alternatives to the property predicated in the generic, then the generic is judged to be false, such as ‘Chickens are female’. Generics such as these which do not predicate a property that is either striking, or that lies along the kind’s characteristic dimensions, require for their truth a large percentage of the kind to possess the property in question.

But what of some of the Aristotelian's other favourite examples, such as 'Acorns become oaks' (Megone, 2002, p. 159)? One could argue that those acorns which do not possess the property of 'becoming an oak tree' do not simply fail to possess it, they possess alternative properties such as 'becoming manure' or 'becoming pig food' and so on. Why, then, do we not judge 'Acorns become oaks' as false, given that so few acorns actually become oak trees? Again, if the psychological account of generics is right, this can be explained without having to posit an Aristotelian 'essence'. Leslie's alternative explanation involves the fact that we are innately disposed to seek out characteristic dimensions for kinds. Given that acorns are a 'natural' kind, we seek to find a value for their method of reproduction and so, even when the majority of acorns do not become oak trees, "we seize on the only positive instances [of reproduction] available and conclude that" acorns become oak trees (Leslie 2008, p. 36). This is because the alternatives do not lie along the characteristic dimensions for the kind and are thus not perceived by us to be salient, whereas 'becoming an oak tree' does, as this expresses how they grow and reproduce – properties that we expect a kind to have if it is to be a successful biological one. Again, and at the risk of labouring the point, it is important to stress that the focus here is on what we *take* to count as a 'characteristic dimension' rather than some Aristotelian notion of 'intrinsic nature'. It is a psychological fact about how our innate primitive generalization mechanism actually works in that we perceive certain things as characteristic for kinds and not others, rather than the Aristotelian's metaphysical claim that such things actually are intrinsic to the kind. As such, it is not the case the neo-Aristotelian account is in a uniquely privileged position because it, and it alone, is able to explain our assent to these generic statements involved and the intuitive force of the troublesome examples used by them. Leslie's account of generics is able to explain our acceptance of these statements in a way that is consistent with a rejection of essentialism.

However, should there be another positive alternative property possessed by acorns, then it seems that we would no longer judge the generic to be true. To see this, consider the following test. Suppose that while an acorn is capable of growing into an oak tree, given the right set of conditions when planted on Earth, if we were to take an acorn to Mars and plant it there, the acorn would instead be capable of growing into a tortoise, say, rather than an oak.⁶⁵ It is plausible to think that knowing about this alternative potential of ‘becoming a tortoise’, which also lies along characteristic dimensions, would change how we judge the truth of ‘Acorns become oak trees’. It seems likely that we would instead consider it to be false, in the same way that we do with Leslie’s peacock example. Indeed, I think that it would require very few acorns to actually become tortoises for us to judge ‘Acorns become oaks’ as false. The virtue of the generics account is that it can account for this context-sensitivity in such cases, something that the Aristotelian account plainly struggles with. For, if acorns have, as their essential nature, the property of ‘becoming an oak tree’, then the truth of ‘Acorns become oak trees’ should not be context-sensitive.

5.1. Yet Another Plague

It might seem that the generics account is open to the following objection based on a revised version of Kagan’s plague case. Suppose that, just like in the original case, an incurable plague sweeps the Earth, but in this scenario it causes 100% of humans to never become persons. The generics account suggests that in this case we should no longer consider the claim ‘humans are persons’ to be true. What then if we find that the intuition that ‘Humans are persons’ is true, remains firm even in imagined scenarios where *no* humans ever become persons? I suggest that this can also be explained by Leslie’s psychological account. Given that personhood lies along

⁶⁵ I am grateful to Rob Lawlor for this example.

characteristic dimensions for humankind, it is plausible that we have the belief, implicit or explicit, that it is central to the survival of humankind that humans are persons, as being a person is conducive for human reproduction, and that this belief persists even when we imagine cases like this where humans exist but none are persons. Indeed, the reason that neo-Aristotelians consider personhood to be part of humankind's essential nature is because it is "conducive to reproduction and thus the survival of the species" (Megone, 2002, p. 173). Thus, while other animals manage to persist and survive without personhood, it seems difficult to think of humans as continuing to survive without a sufficient number of them being persons. Experience tells us that those humans who are non-persons lack the capability to reproduce without the intervention of persons. As such, in a world without persons it is plausible that humankind would fail to persist. As McMahan puts it, "unlike animals, who are equipped with instincts that make independent survival possible despite their otherwise rudimentary cognitive capacities, the cognitively impaired are utterly helpless and dependent" (1996, p. 27). It seems likely then, that when we consider humans within a world where no humans are persons, the fact that we are considering an animal kind leads us to seek a value for their mode of reproduction, and given that the kind in question is humankind, we assume that this involves personhood. We see a similar example of this with 'Bees are sterile'. Despite the majority of bees being sterile, we still judge 'Bees are sterile' to be false.

Why, then, we might ask, do we consider 'Mules are sterile' to be true? This, I suggest, is because to understand *this* particular generic is to know relevant facts about mules, including what they are, how they are created, and the fact that they continue to exist – or exist in the first place – because humans purposely create them by mating male donkeys with female horses. As such, once the number of instances of sterile mules reaches 100% (or close to) we judge the generic to be true knowing that mules do not need to reproduce with one another to continue to exist. Indeed, it seems likely that an individual who confuses 'mule' as another

term for ‘donkey’, would judge the claim ‘Mules are sterile’ to be false.⁶⁶ Why is the claim ‘Humans are persons’ like ‘Bees lay eggs’, rather than ‘Mules are sterile’? The reason for this, I think, is that we find it difficult to think of humans as existing at all without being persons. Consider the following scenario. After the plague - which causes *no* humans to ever become persons - sweeps across the Earth, intelligent Martians decide to continue to create human non-persons for their zoos by growing them from stem cells in laboratories. In such a world, would we still have the intuition that the statement ‘Humans are persons’ is true? I am inclined to think that this would no longer be the case.

Having used Leslie’s account of generics to offer an alternative explanation for the truth of ‘Humans are persons’, it now seems clear that the explanation for why this statement is true suggests no reason for investing it with any moral significance. For, if our acceptance of its truth can be explained in a straightforward and plausible manner by appealing to primitive psychological mechanisms – which can further be accounted for in terms of the evolutionary benefits of such mechanisms - and not to metaphysical essences, then it seems difficult to see why we should think of it as having any moral importance.

6. As For Those Pesky Intuitions...

What then of those stubborn intuitions which refuse to completely dissipate and rear their heads at the slightest mention of an intuitive example relating to humans and personhood? I have presented what I take to be a convincing explanation for the intuitive force of troublesome examples used by neo-Aristotelian accounts which accounts for this intuitive force without investing in any of the Aristotelian’s metaphysics and which is, therefore, consistent with a rejection of both essentialism and the moral significance of species membership. My purpose

⁶⁶ As a somewhat embarrassing autobiographical note, I had long thought that this was the case.

in this chapter, then, has been to show is that it is not simply enough, for the plausibility of the Aristotelian's account, that we have these intuitions when considering various thought-experiments. The fact that an alternative, and more plausible, explanation for these intuitions exists shows that a mere reliance on such intuitions cannot provide enough support for their arguments. The simple existence of an alternative plausible explanation for the persistence of this intuition suggests that the intuition itself is not proof of the Aristotelian's position and, as such, casts doubt on the benefit of appealing to the intuitive force of examples and the intuitions they pull on. Such examples are used *because* of their troublesome nature, that is, the fact that they are still considered to be true in the face of many counterinstances, is employed by the Aristotelian to provide support for the claim that natural kinds have, as their intrinsic nature, an essential set of potentialities which may or may not be actualised (Megone, 2002, p. 162). The generics account, by positing a psychological explanation, rather than a metaphysical one, for why we judge 'Humans are persons' to be true, questions the effectiveness of using these intuitive examples as support for the Aristotelian's argument. The thought, then, is that the generics story is metaphysically uninvested, and does all the work we need it to do without having to call in the neo-Aristotelians.

7. The Generics Debate Revisited

While a full account of the debate about which theory of the truth of generics is the correct one is far beyond the scope of this chapter, I noted earlier that there appear to be *prima facie* reasons for thinking Leslie's psychological account plausible. In addition to these, the fact that Leslie's account adheres so well with, and fully explains, many of our conflicting intuitions in the speciesism debate, offers some further independent support for her account. The psychological account convincingly explains why we consider the claim 'Humans are persons' to be true even in the plague case where most humans do not become persons, as well as why

we consider the converse claim ‘Humans are non-persons’ to be false in such a scenario. It also explains why we judge the claim ‘Acorns become oaks’ to be true, as well as being able to account for the context-sensitivity of such judgements. For example, when we consider the science-fiction case discussed earlier where acorns are capable of growing into a tortoise rather than an oak tree if planted on Mars, Leslie’s account is able to explain why, in such a scenario, we no longer consider ‘Acorns become oaks’ to be true, instead assenting only to ‘Acorns become oaks or tortoises’.

8. Conclusion

My purpose in this chapter, then, has been to dispel the feeling of an unanswered question which results from accepting the force of persuasive arguments from Singer and others while not quite being able to shake off a nagging intuition that there is something morally significant about humankind. I began this chapter by noting that, despite the arguments presented in the previous chapters to doubt both that species membership is significant and that personhood is part of human essence (essentialism), it is clear that the intuitive force of the examples used by neo-Aristotelians stubbornly persists. As such, it seems that without an explanation which is consistent with these arguments, the stalemate between Singer’s position and that of the neo-Aristotelian’s account also remains. But the significance of this chapter is to show that the advocates of Singer’s arguments are at an advantage: in addition to the arguments presented throughout this thesis, the generics account can explain the intuitive force of the examples in a way that is consistent with a rejection of both essentialism and the significance of species membership. Like Kagan, I would be more comfortable with an argument that better adheres with all of my common-sense intuitions about the moral significance of species membership. Indeed, much of this thesis has been spent in search of such an argument. Perhaps one is forthcoming, but in its absence, I find myself content to settle for what seems like the next best

thing: an account which offers an explanation as to why the common-sense intuitions stubbornly persist even in the face of compelling Singer-type arguments, and is thus consistent with those arguments.

Works Cited

- Cohen, A., 1996. *Think Generic: The Meaning and Use of Generic Sentences*. PhD Dissertation: Carnegie Mellon University.
- Cohen, C., 1986. The Case for the Use of Animals in Biomedical Research. *New England Journal of Medicine*, Volume 315, pp. 865-869.
- Foot, P., 2001. *Natural Goodness*. Oxford: Clarendon Press.
- Gelman, S., 2003. *The Essential Child*. New York: Oxford University Press.
- Haslanger, S., 2011. Ideology, Generics, and Common Ground. In: C. Witt, ed. *Feminist Metaphysics: Explorations in the Ontology of Sex, Gender and the Self*. London: Springer, pp. 179-207.
- Hollander, M. A., Gelman, S. & Star, J., 2002. Children's Interpretation of Generic Noun Phrases. *Developmental Psychology*, Volume 36, pp. 883-94.
- Kagan, S., 2015. What's Wrong with Specieism?. *Journal of Applied Philosophy*, Volume (Early Online Edition), pp. 1-21.
- Lawlor, R., 2012. The Ethical Treatment of Animals: The Moral Significance of Darwin's Theory. In: M. Brinkworth & F. Weinert, eds. *Evolution 2.0: Implications of Darwinism in Philosophy and the Social and Natural Sciences*. New York: Springer, pp. 147-163.
- Leslie, S. J., Forthcoming. The Original Sin of Cognition: Fear, Prejudice and Generalization. *The Journal of Philosophy*.
- Leslie, S. J., 2012. Generics. In: G. Russell & D. G. Fara, eds. *The Routledge Companion to Philosophy of Language*. New York: Routledge, pp. 355-367.
- Leslie, S. J., 2008. Generics: Cognition and Acquisition. *Philosophical Review*, 117(1), pp. 1-48.
- Liebman, D., 2011. Simple Generics. *Nous*, 45(3), pp. 409-442.
- Macario, J. F., Shipley, E. & Billman, D., 1990. Induction from a Single Instance: Formation of a Novel Category. *Journal of Experimental Child Psychology*, Volume 50, pp. 179-99.
- Megone, C., 1998. Aristotle's Function Argument and the Concept of Mental Illness. *Philosophy, Psychiatry, & Psychology*, 5(3), pp. 187-201.
- Megone, C., 2002. Potentiality and Persons: An Aristotelian Perspective. In: M. G. Kuczewski & R. Polansky, eds. *Bioethics: Ancient Themes in Contemporary Issues*. Cambridge: MIT Press, pp. 155-177.

- Nobis, N., 2004. Carl Cohen's 'Kind' Arguments For Animal Rights and Against Human Rights. *Journal of Applied Philosophy*, 21(1), pp. 43-49.
- Oderberg, D. S., 2007. *Real Essentialism*. New York: Routledge.
- Oderberg, D. S., 2000. *Moral Theory: A Non-Consequentialist Approach*. Oxford: Blackwell Publishing.
- Pelletier, F. & Asher, N., 1997. Generics and Defaults. In: J. van Benthem & A. ter Meulen, eds. *Handbook of Logic and Language*. Cambridge: MIT Press, pp. 1125-1179.
- Singer, P., 1993. *Practical Ethics*. 2nd ed. New York: Cambridge University Press.
- Sterkin, R. K., 2015. Lelsie on Generics. *Philosophical Studies*, 56(1), pp. 75-93.
- Thompson, M., 1995. The Representation of Life. In: R. Hursthouse, G. Lawrence & W. Quinn, eds. *Virtues and Reasons*. Oxford: Clarendon Press, pp. 247-296.
- Tooley, M., 1972. Abortion and Infanticide. *Philosophy and Public Affairs*, 2(1), pp. 32-65.

Conclusion

This thesis has been concerned with two main areas within animal ethics: rights and discrimination. Within these areas, I think, lie the most intractable problems and those which are most in need of being brought into the light and clearly defined. In chapters one and two, I looked at some of the problems and inconsistencies which follow from taking animal rights seriously, as Tom Regan does. In those chapters, I argued that taking animal rights seriously within an overall account of rights creates problems of consistency and I concluded that, specifically, there remains much more work to be done with regard to the permissions operating around self- and other-defence in this area. Chapter three looked at attempts to justify the Common-Sense intuition – that there is something morally significant about humankind – by holding that it seems plausible to suppose that the cognitively impaired human is unfortunate in a way that similarly endowed animals are not. I focus on two popular and promising accounts of misfortune but ultimately conclude that neither account can fully account for our intuitions. Chapter four considered whether we can ‘get away with’ speciesism by denying that it is, in fact, a prejudice. I argued that one constraint on a successful defence of speciesism is to offer a justification for it which does not also leave proponents open to parallel defences of malign forms of prejudice like racism and sexism. I conclude that the justifications considered in this chapter are unable to offer a defence of speciesism which can successfully avoid such troublesome analogies. Finally, in Chapter five, I considered the stubborn persistence of the Common-Sense Intuition despite persuasive arguments offered against it throughout the thesis. I put forward an account of the truth conditions of generics and argue that this account is able to explain this intuition in a way that is consistent with the arguments presented in earlier chapters and, thus, consistent with a rejection of the moral significance of species membership.

I have wrestled with the above problems and have attempted to make some progress with them. It will not be lost on the reader that, in many ways, the chapters in the thesis have raised

more questions than they have answered. However, while there remain problems left unsolved, I feel that by providing some clarity on what the most significant problems in these areas are and what would be required to solve them, I have left the overall state of the debate in a better position than which I found it in, and have hopefully provided a useful roadmap to achieving a successful defence of the Common-Sense Intuition.