

**The Agency Account of Moral Status:
Defending the Equal Moral Status of
Humans and
Non-Human Animals**

Marc Gareth Wilcox

Submitted in accordance with the requirements for the degree of
Philosophy PhD

University of Leeds
School of Philosophy, Religion and the History of Science
September 2017

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

Acknowledgements

I could not have completed this thesis without the help and support of some important individuals. Whilst I cannot thank everyone who has helped me by name, there are some important people who deserve a special mention.

First of all I must of course thank my supervisors, Gerald Lang and Pekka Väyrynen. Through constantly challenging my ideas both of these individuals have pushed me to do better, clarify my ideas and develop my arguments further. Without their input this thesis would not be as clear and well-developed as it is. I must also thank Pekka for helping me navigate the bureaucratic university processes surrounding completing a PhD and providing me with a book for every problem I encountered.

I would also like to thank my examiners Helen Steward and Mark Rowlands. Other than fulfilling their roles as examiners, the work of both of these individuals has played an important part in the development of my research. I would also like to thank Helen in particular for the trouble she went through to arrange my *viva voce* and for the time she has spent with me discussing animal action and animal minds over the last four years.

Next I would like to thank two very special individuals who have played an important part in the completion of this piece of work. First is my wife Penelope. She encouraged me to reapply for PhD programs after failing to secure a place the first time around. Without her encouragement and belief in me, I would not have even started my doctoral research, let alone completed it. Moreover, through the four years it has taken me to complete this thesis she has moved across the country for me, been a sounding board for my inchoate ideas and pushed me to get back to my work after the inevitable and frequent setbacks that arise in research, not to mention providing me with meals whilst I've been occupied with reading and writing. There are many times when I thought that I couldn't continue with my PhD and I have no doubt that if not for Penny's support, I would not have finished this thesis. I cannot thank her enough for everything she has done to ensure my success.

Second is Enzo, who played no small part in inspiring this project. Merely through his day to day behaviour, Enzo has made me constantly doubt the dogmatic approach of previous philosophers regarding animal action and inspired me to develop the account of animal action I defend here. Enzo sadly died last year but he will be fondly remembered whenever I think back to the ideas contained in this work.

Finally I am grateful to the School of PRHS at Leeds who provided me with a Graduate Teaching Scholarship for the first three years of my PhD allowing me to undertake this research in the first place. I must also thank my friends inside and outside this department, who have spent time discussing the ideas contained within this thesis, as well as my parents and grandparents who have helped me financially during the completion of this work.

Abstract

In this thesis, I argue that humans and sentient animals have equal moral status in the sense that they ought to have like interests equally considered. Furthermore, they are owed strong *pro tanto* duties to be free from having pain inflicted upon them, having their lives ended and having their liberty restricted. I argue for these claims by developing and defending an account of moral status grounded in agency. This account takes agency, understood as the capacity to act on motivating reasons, to be the necessary and sufficient condition for moral status. Further, I argue that agency is sufficient to have interests in liberty, continued existence and freedom from pain. As such we *pro tanto* wrong agents when we frustrate these interests. I show that sentient beings necessarily possess agency in the relevant sense, because the best account of the nature of sentience, entails that sentient beings have the psychological resources to form and act upon motivating reasons. Thus, I argue that sentient animals must possess interests in liberty, continued existence and freedom from pain, just as autonomous agents do. Therefore we should take all agents, regardless of further facts about their abilities, to possess equal moral status and be owed *pro tanto* duties to be free from having pain inflicted upon them, having their lives ended and having their liberty restricted.

Contents

Introduction	1
i. Why Think Animals Matter at All?	4
ii. Overview	12
Chapter 1 What is Moral Status and Why is it Important?	15
I. The Moral Significant of Interests	17
i. Two Accounts of Interests	18
ii. Biological Functions as Interests	19
iii. Producing Offspring, Living and Flourishing	21
II. Is the Notion of Moral Status Useful?	24
III. Does Moral Status come in Degrees?	32
i. The Unequal Consideration Model	34
ii. The Unequal Interests Model	36
IV. Relationalism and Individualism	38
Conclusion	41
Chapter 2 Why do we need an Agency Account?	43
I. The Sentience Account	44
II. The Subject-of-a-Life Account	49
III. The Kantian Account	53
IV. Constructing an Ideal Account of Moral Status	60
V. The Agency Account	61
i. The Agency Account Determines Treatment	62
ii. Human and Animal Moral Status	63
iii. The Agency Account is Analytically Reductive	64
Conclusion	66

Chapter 3 From Sentience to Desires	68
I. Evolution and Instincts	69
II. Sentience	72
III. Accounts of Sensory Pleasant Feeling	73
i. Felt Quality Accounts	76
ii. Attitudinal Accounts	78
IV. Pleasant Sensory Feelings as Desire Satisfaction	82
i. General Desires	86
ii. Desires for No Reason	87
iii. Moods	88
iv. Lost Desires	89
v. Novel Experiences	90
V. The Nature of Desire	91
i. Desires with Motivation	93
ii. Non-Actable Desires	94
Conclusion	98
Chapter 4 From Desires to Action	99
I. The Nature of Beliefs	99
i. Action-Guiding Beliefs	102
ii. The Argument from Specification	106
iii. The Argument from Concepts	110
II. Acting Intentionally	116
i. Acting for a Reason	116
ii. Reason Super Blind-Sighted	119
iii. Autonomy and Moral Agency	121
iv. Control without Autonomy	124
III. The Nature of Agency	127
i. Acting without a Body	127
ii. Agency as a Capacity	130

Conclusion	132
Chapter 5 The Agency Account of Moral Status: Liberty and Pain	133
I. Agency as a Necessary Condition for Moral Status	133
II. Agency as a Sufficient Condition for Moral Status	134
III. Pain	137
IV. Liberty	140
i. Kinds of Liberty	140
ii. The Instrumental Value of Liberty	143
iii. The Desire-Based Account	144
iv. The Teleological Account	147
v. Liberty as Choice	152
vi. The Instrumental Value of Choice	152
vii. The Self- Determination Account	156
viii. Animals and Humans	164
ix. Practical Implications	165
Conclusion	167
Chapter 6 The Agency Account of Moral Status: Continued	
Existence	169
I. Continued Existence	169
i. Future-Directed Desires	173
ii. Psychological Connectedness and Time-Relative Interests	178
iii. Categorical Desires	188
II. The Loss of Autonomy Argument	193
III. Objections to the Agency Account of Moral Status	196
Conclusion	197
Concluding Remarks	199
Bibliography	202

Introduction



Henry Salt suggested that ‘the prime cause of man’s injustice to the lower animals is the belief that they are mere automata, devoid alike of spirit, character, and individuality...’¹ Salt wrote these words over 100 years ago and in that time our perceptions of animals have changed; most people do not think of animals as mere automata. Nonetheless, Salt’s point still has bite; it seems we do not recognise animals as true individuals with unique characters. Where we think of ourselves as existing ‘in the world’ we often think of them as part of it, like rocks and trees.² They are seen as something fundamentally different from humans. Often thought to be controlled by their biological nature, they are alive but not really *living a life*. In this thesis I will argue against this position. I will argue that those animals that are capable of sentience are also agents, capable of determining their actions minute to minute, just like humans.³ Further, I will argue that because of this, sentient animals possess a ‘moral status’, that is to say, they are individuals that matter, and as such they ought not to be treated in certain ways, just as is the case with humans.

More generally, the aim of this thesis is to give an account of what the conditions are for the possession of moral status, and in doing so plausibly explain why both humans and sentient non-humans matter, taking into account the variety and depth of duties we owe to these individuals. I will argue that though previous accounts of moral status have made us aware that there is something intuitively wrong about our current treatment of animals, they have failed to provide an account that robustly and plausibly defends this claim. It is not that animals matter because they are capable of feeling pain, or because they are living beings. I will argue that sentient animals (and humans) matter because they possess agency.

The claim that animals are agents is paradoxically both deeply intuitive and highly controversial. Animals manage to navigate the world, avoid dangers and seek out food,

¹ Henry Salt, *Animal’s Rights Considered in Relation to Social Progress* (Clarks Summit: Society for Animal Rights, Inc., 1980), 90.

² Sue Donaldson and Will Kymlicka, *Zoopolis: A Political Theory of Animal Rights* (Oxford: Oxford University Press, 2013), 37.

³ Throughout the course of this thesis one should take the term ‘animal’ and ‘animals’ to mean sentient non-human animal and animals respectively, unless otherwise specified. Which non-human animals possess sentience is a tricky question that I will not discuss in this thesis; however I take it that vertebrates, some cephalopods and some crustaceans (at least some crabs and lobsters) are sentient.

comfort and companionship. What is more, if animals are individuals with desires and subjective mental states, as it seems they are, it is simply implausible to suggest that they are not agents under some description. Yet at the same time almost no theorists recognise the moral significance of this capacity to act in animals (far fewer argue that this is central to the explanation of their moral status).⁴ I suspect that at least one reason for this is that philosophy has a deeply entrenched view of an agent as a kind of being that is capable of hyper conscious reasoning about their actions, a being that can recognise, deliberate over and modify their reasons for action. In short, an agent is usually thought to be a being that has the psychological architecture of a human. In this respect, I will argue, philosophy is wrong.

There are many senses in which one can be an agent. Flowers open, the tides rise, planets orbit and engines turn. One might describe some of these movements as 'agency'. This kind of agency is not what I am interested in here, however. Animals are obviously capable of agency in this basic sense but I will argue that they are agents in a more interesting sense. I will argue that animals are agents in the same sense that humans are agents. We think of humans as intentional beings; individuals capable of purposive action and acting upon reasons. This is the kind of agency that grounds moral status, and this is the kind of agency that I will argue animals possess.

Of course animals do not exercise their agency through the exact same processes as humans. Most humans are capable of conscious deliberation over their reasons among other things. These humans are *autonomous* agents. However I will argue that agency does not require such metacognitive reasoning or conscious reflection upon reasons. I contend that agency is a more basic capacity than this. This is not to say that agency is a simple cognitive capacity; it involves motivating mental states, beliefs, desires, reasons as well as a phenomenal consciousness. Nonetheless I will argue that animals possess these mental states and thus possess agency.

What is so special about agency then that makes it a plausible condition upon which to build an account of moral status? After all, humans and animals share lots of attributes. We are all oxygen breathers for example. Why shouldn't we base an account of moral status on

⁴ Korsgaard appears to use something like agency to explain the moral status of animals but, as I will argue in Chapter 2, she in fact has something more like natural goals or goods in mind. Christine M. Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals," *The Tanner Lectures on Human Values* 25 (2004): 26. Since starting this thesis two other theorists have suggested agency accounts of moral status independently of myself: Jeff Sebo, "Agency and Moral Status," *Journal of Moral Philosophy* 14, no. 1 (2017): 1-22; Natalie Thomas, *Animal Ethics and the Autonomous Animal Self* (London: Palgrave Macmillan, 2016).

this shared capacity? Importantly, we consider agency to have substantial moral significance for humans. Being an agent is what makes one an individual with the power to control the kind of life one lives. Agency makes one not just a passive receiver of goods and harms, but a being capable of bringing about goods and harms oneself. As agents we have preferences, likes, desires, dislikes and aversions. If one possesses agency then, one is not just a something, but *a someone*; an individual with a unique subjective psychological life.

To paraphrase Isaiah Berlin, being an agent makes one different from the rest of the world in that one should be treated as an individual, have their uniqueness recognised and not be classed as a featureless amalgam, a statistical unit without identifiable features or purposes of their own.⁵ Thus, by understanding animals as agents, we recognise their individuality; we recognise that different individual animals have their own preferences, desires and lives. What is more, we take agency to be a plausible explanation for why various kinds of treatment are wrong. Ending our lives, stealing our possessions, using our bodies, force-feeding and imprisoning us are all actions we would consider objectionable because they undermine our agency. Such actions disregard, or attempt to undermine our desires, actions or plans we have put into place, stop us from engaging in any further actions or take away our opportunity to choose. Plausibly then agents seem to be individuals that matter and warrant moral consideration.

Thus, through this thesis, I will argue for three major claims. I will argue that the capacity for sentience logically entails the capacity of agency and therefore all sentient beings are agents. Further I will argue that agency is the necessary and sufficient condition for the possession of moral status. Finally, I will contend that we owe moral status holders three central *pro tanto* duties: a duty not to inflict pain upon them, a duty not to end their lives and a duty not to restrict their liberty.⁶ Before I start the argument proper however, it is worth saying a little bit about why one might think animals matter at all.

i. Why Think Animals Matter at All?

⁵ Berlin was arguing for the value of political liberty to humans but his words seem appropriate here. Berlin, *Two Concepts, Four Essays on Liberty*, 155.

⁶ These duties are *pro tanto* in the sense that they are not absolute. So, though we owe these duties to moral status holders, we are justified in not fulfilling these duties if we are obligated to fulfil some other stronger, conflicting duties.

It is a plausible assumption that all individuals are owed the same treatment unless there is some relevant difference between them that can justify a difference in treatment.⁷ This is what Rachels and Rowlands call, *the Principle of Equality*. We happily apply this principle to our everyday treatment of humans. For instance, Rachels argues, if a university must make an admissions decision between two applicants, a relevant difference between the applicants, is their grades. That one applicant has higher grades than the other, is a relevant difference that can justify the decision of the university's admission board to accept the more highly qualified applicant. However the fact that one candidate has a broken arm, whereas the other does not, would clearly be an irrelevant difference upon which to base a university admissions decision. Such a difference could not justify a difference in treatment in this case.⁸

If we are motivated by this basic principle of equality it seems that animals as well as humans must be an object of moral concern. If we owe all individuals the same treatment unless there is some relevant difference, then how can we justify treating animals so radically different from how we treat humans? Many animals used for food and research for instance have lives where they are kept in overcrowded conditions, physically confined, have lives significantly shorter than their natural life span and routinely have severe pain and long-term suffering inflicted upon them. In short, they have lives that we would consider to be torturous for humans. Even those animals that are not within the food or research industries, still live lives over which humans have an overwhelming amount of control. Surely the mere fact that animals are not humans, is not a relevant difference that can justify the different ways in which we treat humans and animals?

Merely belonging to a species does not in itself seem to grant one moral status. In fact, the Principle of Equality appears directly opposed to such an argument. We use this principle (or principles like this)⁹ to argue against racists and sexists and to show that claims that some gender or race is due greater consideration than some other, is groundless. Differences such as gender and race are not differences that are morally significant. Belonging to one gender or race does not make one matter any more or less than anybody else. Those that claim otherwise

⁷ James Rachels, *Created from Animals: Moral Implications of Darwinism* (Oxford: Oxford University Press, 1999), 176; Mark Rowlands, *Animals Like Us* (London: Verso, 2002), 35.

⁸ Rachels, *Created from Animals*, 177.

⁹ Singer suggests a similar formulation of the basic principle of equality: *the Equal Consideration of Interests Principle*, according to which similar interests should be treated similarly regardless of other facts about the possessor of the interest i.e. their race, gender, species. Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1993), 21.

are merely prejudiced. Similarly, many have argued, by failing to recognise that animals are individuals who matter, we are guilty of speciesism.¹⁰

Speciesism is a prejudice against other species. We act in a speciesist way when we unfairly discriminate against an individual because they belong to some species. Of course not all discrimination based upon species is unjustified, just as not all discrimination based upon gender is unjustified. What makes some action speciesist is the using of species to discriminate, when species is not a relevant property. Thus it would not be speciesist to refuse a dog admission onto a university course just as it would not be sexist to refuse a man the legal right to an abortion. In both cases gender and species count as a relevant difference and therefore the treatment we give to these individuals, despite being different from how we would treat other humans and women respectively, is justified. However by refusing to recognise the plight of animals *because* they are not humans, we seem to be engaging in speciesism. Being human is not a necessary requirement for being able to experience suffering. Thus species cannot justify the current difference in treatment between humans and animals. Seemingly animals must matter, at least to some extent.

A common objection to this argument is that although biological properties like species do not *themselves* justify the difference in treatment between humans and animals, being human means having some unique property, and it is possession of this property that can explain why humans alone possess moral status. This is what Rachels calls 'Qualified Speciesism'.¹¹ The possession of this property amounts to a morally relevant difference between humans and animals which can justify our different treatment of humans and animals. Historically, potential candidates for this property were often religious in nature, for example the possession of souls, being made in the image of God, or being given dominion over all other forms of life.¹² Presently however, psychological capacities are often believed to be what separates us from other animals. Among the most frequently suggested are possessing rationality, self-consciousness, language or autonomy.

Contrary to this argument, many animal advocates have argued that none of these properties can justify the current difference in treatment between humans and animals. Two

¹⁰ Oscar Horta, "The Scope of the Argument from Species Overlap," *Journal of Applied Philosophy* 31, no. 2 (2014): 142-54; Peter Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals* (London: Granada Publishing Ltd, 1977).

¹¹ Rachels, *Created from Animals*, 184.

¹² I won't discuss the moral significance of these religious based properties further since I take it they are no longer widely accepted and in any case require significant metaphysical commitments.

arguments are commonly given in defence of this claim. Firstly, even assuming that animals lack all of the psychological properties that humans uniquely possess, it is not clear why possession of these properties counts as a morally relevant difference. That an individual possesses self-consciousness for example, may well entitle them to treatment different from that which is owed to an individual that lacks self-consciousness. For instance, often it is argued that only self-conscious beings can be harmed by death (because death is only a harm when it frustrates future-directed desires and only self-conscious beings possess future-directed desires).¹³ Even assuming that this is true however, self-consciousness cannot plausibly justify the immense difference in current treatment between humans and animals.¹⁴ For instance, if I could inflict pain on two individuals, one self-conscious and one not, the fact that one possesses self-consciousness does not appear to be a relevant difference that could justify inflicting pain upon one but not on the other. The only difference that could be relevant here is the ability to suffer. Since both individuals are capable of suffering, both ought to be treated alike (assuming there is no other relevant difference).¹⁵ Neither it seems should have pain inflicted upon them.¹⁶

An argument of this same form can be given for any of the properties which supposedly justify the considerable disparity between the treatment of humans and animals. Whilst these properties may justify a difference in treatment in some circumstances, they do not justify the dismissal of the moral status of animals. Even if one lacks, self-consciousness, rationality, language and autonomy, the fact that one can still suffer or be harmed, obligates moral agents to recognise one as an individual that has moral status.

Secondly, even assuming that these properties could justify the dismissal of the moral status of animals, many theorists argue that none of these properties are uniquely possessed by all and only humans. Moreover they contend that there is no property that all humans and only humans uniquely possess, and as such there is nothing unique about being human save for the mere fact that one belongs to the species *homo sapiens*. This is *the Argument from Species Overlap*.¹⁷ The argument suggests that for any property supposed to be uniquely possessed by

¹³ Singer and Tooley famously argue for this claim though it is not particularly uncommon. Singer, *Practical Ethics*; Michael Tooley, "Abortion and Infanticide," *Philosophy & Public Affairs*, 1972, 37–65.

¹⁴ I will argue that self-consciousness is not required to be harmed by death in Chapter 5.

¹⁵ Self-consciousness may make a difference to how much one is harmed by some given amount of pain but this is beside the point here.

¹⁶ Though he is not the first to suggest this argument, Rachels makes this argument very clearly. Rachels, *Created from Animals*, 178–79.

¹⁷ Also known as the Argument for Marginal Cases, Horta argues that we should instead refer to this as the Argument from Species Overlap due to the negative connotations associated with referring to

humans, it is the case that either a) some human lacks this property b) some non-human animal possesses this property or c) some human lacks this property *and* some non-human animal possesses this property. To illustrate the point of this argument consider the property of self-consciousness.

It cannot be the case that what grants humans moral status is the possession of self-consciousness since some humans do not possess this property. The very young, the senile and severely intellectually disabled humans lack this property. As such, if those beings that matter are those that possess self-consciousness, then some humans must not matter. This is deeply problematic; humans that lack self-consciousness still intuitively matter and, not only that, but they appear to matter as much as paradigmatic self-conscious humans. What is more, not only do some humans lack this capacity for self-consciousness but some animals possess this capacity.

So called 'Mirror Tests' test for self-consciousness by marking an individual with removable dots on their face/head without their knowledge and providing them with a mirror.¹⁸ If one can use the mirror to locate and investigate the dots on one's own face and head, then one must be able to recognise that one is looking at a reflection of oneself in the mirror, and thus possess some form of self-consciousness. Apes and dolphins, among other animals, are capable of passing dot tests, whereas some humans are not. Thus apes and dolphins are self-conscious in a way that some humans are not. If self-consciousness is a sufficient condition for moral status, it seems dolphins and apes must possess moral status and ought to be treated similarly to humans. An argument of this form can be given for autonomy, language or rationality too. Moreover, it is the case that for any property which has been suggested to justify the significant difference in our treatment of humans and animals, at least one human will lack this property or some animal will possess this property (or both).¹⁹ Hence, there can be no defence of the claim that humans uniquely possess the property that grounds moral status.

Through these two arguments, defenders of animals' moral status have argued that one cannot defend speciesism by arguing for the uniqueness of humans. Humans do not

intellectually disabled humans, infants and the senile as 'marginal'. He is, I think, correct. See: Horta, "The Scope of the Argument from Species Overlap"; Singer, *Practical Ethics*, 59–60; Donaldson and Kymlicka, *Zoopolis*, 26–27; Evelyn B. Pluhar, *Beyond Prejudice: The Moral Significance of Human and Nonhuman Animals* (Duke: Duke University Press, 2012), chap. 1; Rowlands, *Animals Like Us*, 44–45.

¹⁸ Gordon G. Gallup, "Chimpanzees: Self-Recognition," *Science* 167, no. 3914 (1970): 86–87.

¹⁹ This argument of course does not apply to the aforementioned religious properties, these are generally rejected for independent reasons.

uniquely possess any properties that can justify the current difference in the way we treat humans and animals. Whilst the possession of some properties may justify some difference in treatment, for instance a lack of self-consciousness may justify moral agents ending one's life, there is no property that is possessed by humans that can justify the claim that humans uniquely possess moral status. Further, if one endorses a qualified speciesist account, one would have to give animals better treatment than humans in some cases at least. For instance, if we compare the treatment owed to a severely intellectually disabled human and a typical adult gorilla, assuming self-consciousness is a relevant difference, and assuming that the gorilla possesses self-consciousness but the human does not, one would be justified in ending the life of the human but not the gorilla.

Many will find this conclusion unpalatable and several counter-arguments have been given in an attempt to defend the qualified speciesist position. Some such arguments claim that the criterion for moral status is a property possessed by only normal adult humans but that the humans that lack this property (infants, the senile and the intellectually disabled) fulfil some other, closely linked, sufficient condition for moral status. For instance, potentiality arguments suggest that the very young have the potential to possess the property that grounds moral status.²⁰ Other arguments state that the senile have previously possessed the property that grounds moral status. Since these groups of individuals possess the property for moral status at some point in their lifetime, these arguments maintain, infants and the senile possess moral status too. However, even assuming these arguments are correct, they cannot defend the moral status of intellectually disabled humans. And, since intellectually disabled humans intuitively possess the same moral status as normal adult humans, these arguments alone are insufficient to show the qualified speciesist position to be plausible.

There is another argument however, that attempts to establish the moral status of infants, the senile, the intellectually disabled and any other atypical humans that might lack the morally significant property that normal adult humans possess. McMahan calls this the *Nature-of-the-Kind Argument*.²¹ The Nature-of-the-Kind Argument contends that though some humans do not possess morally significant property *x*, humans are the kind of beings that *typically* possess this property and *this* is sufficient to possess moral status. Thus one need not actually possess morally significant property *x* in order to possess moral status. It is sufficient that one belong to a group of which the typical members possess morally significant property

²⁰ Tooley, "Abortion and Infanticide."

²¹ Jeff McMahan, "'Our Fellow Creatures,'" *The Journal of Ethics* 9, no. 3-4 (2005): 353-80.

x. Hence, because intellectually disabled humans belong to the species *homo sapiens*, a species of which the typical member possesses morally significant property *x*, intellectually disabled humans possess moral status, though they lack this property themselves. If this argument was successful it could secure equal moral status for all humans. However I think it has significant problems.

There are problems with viewing species as distinct natural kinds. Though the different species we see in the world around us seem to be entirely unrelated groups of individuals, the theory of evolution tells us that this isn't the case. If we were able to see every being that had ever existed, arranged offspring next to parent, we would see a slow and gradual change from individuals we designate as belonging to one species, to those we designate as belonging to another. The gaps that appear to differentiate separate species in the present day are merely contingent; they are present because the individual animals that fall between the categories of the extant species are no longer in existence. In truth there are no sharp breaks or gaps between species, there is a continuum of beings all closely related to their ancestors and their descendants.²² This doesn't just show that species boundaries are vague however, but that the notion of species as entirely separate natural kinds is wrong-footed.

Species are more plausibly artificial categories that we impose on nature to help us make sense of different lifeforms. Darwin himself stated that the application of the notion of 'species' is arbitrary and used merely for convenience.²³ Of course we can pick out groups of individuals and recognise them as members of some species. We don't want to deny that species exists as a concept or that individuals genuinely belong to specific species. However, we shouldn't think that these individuals are separated from others in virtue of belonging to different, metaphysically deep, natural kinds. All individuals differ from others only in varying degrees. There is nothing fundamentally different about humans that separates us from our nearest non-human biological relatives. To claim otherwise is not only at odds with evolutionary theory but seems to suggest an implausible understanding of how the first humans began to exist. Thus, since species are not metaphysically distinct kinds, it cannot be

²² Rachels, *Created from Animals*, 174; Richard Dawkins, *The Ancestor's Tale* (London: Orion Publishing Group, Limited, 2005), 317; Bernard E Rollin, *The Unheeded Cry Animal Consciousness, Animal Pain and Science*. (Oxford: Oxford University Press, 1990), 32.

²³ Charles Darwin, *The Origin of Species and the Voyage of the Beagle*, ed. Ruth Padel (London: Vintage Classics, 2009), 579.

the case that all and only humans possess moral status in virtue of belonging to the species *homo sapiens*.

A further argument against species as metaphysically distinct kinds draws on the fact that co-membership of a species is a non-transitive relation.²⁴ A reasonably uncontroversial biological criterion for belonging to a given species is the ability to interbreed with other members of that species. According to this understanding of species, being a member of the same species as another individual is a non-transitive relation. This means that co-membership of a species is the same kind of relation as being next to another person for example. If I am next to Hannah, and Hannah is next to Matt, I will not be next to Matt. Though I share the same relation to Hannah, as Hannah shares to Matt, I do not share this relation to Matt. This is in contrast to transitive relations like being taller than someone else. If I am taller than Hannah and Hannah is taller than Matt then I must be taller than Matt too. The claim that species-membership is non-transitive means that an individual could share membership of their species with two other individuals and yet these two individuals need not be members of a single species. To see how this is possible imagine all of one's ancestors, from now back to the earliest forms of life, existing in the present time and arranged offspring next to parents, in family tree-like shape.

As one would expect, I would be able to breed with my recent ancestors at the top reaches of the tree. Further I would be able to breed with my ancestors many generations back (let's say up to 9000 generations back). Let's call my ancestors of between 1-9000 generations back 'my intermediate ancestors'. Because I can breed with all of these individuals, we can say that all my intermediate ancestors must belong to the same species as myself. However there will come a point at which I will not be able to breed with my ancestors (9001 generations back). Let's call individuals past this point 'my distant ancestors'. Since I cannot breed with my distant ancestors it seems that we cannot be members of the same species.

Interestingly, though my distant ancestors cannot breed with me, there will be some overlap in the individuals with which my distant ancestors and I can breed. My distant ancestors will be able to breed with many of their own ancestors, but will also be able to breed with some of their decedents; some of my intermediate ancestors. Let's say that they can breed with generations 7000-9000 of my intermediate ancestors. This being the case there are 2000

²⁴ Rob Lawlor, "The Ethical Treatment of Animals: An Example of the Moral Significance of Darwin's Theory," in *Evolution 2.0: Implications of Darwinism in Philosophy and the Social and Natural Sciences*, ed. Martin Brinkworth and Friedel Weinert (New York: Springer, 2011), 150-55.

generations of individuals with whom both I and my distant ancestors can breed. Thus I and my intermediate ancestors are members of the same species, and my distant ancestors and my intermediate ancestors are members of the same species, but I and my distant ancestors are *not* members of the same species. Thus, at least on this understanding of species membership, species membership is non-transitive.

This is problematic because moral status, unlike species membership, *is* transitive. If A and B have the same moral status and B and C have the same moral status then A and C *must have the same moral status*. Having moral status then is like being taller than someone else. However in the example above, A and B are the same species, and B and C are the same species but A and C are not the same species. Therefore because species membership is non-transitive, but moral status is, species membership cannot ground moral status.²⁵ If all members of one species are moral status holders, then possessing membership of a species must be a transitive relation, which it is not, or having the same moral status as another must be a non-transitive relation, which it is not.

Finally, one should also note that even if the Nature-of-the-Kind argument is successful in explaining why all humans matter, it still does not plausibly explain why non-humans do not matter. Assuming suffering is plausibly bad for whoever experiences it, the suffering of non-human sentient beings still seems like it should be an object of concern for moral agents. Therefore, at best, if the Nature-of-a-Kind Argument is successful, it can show only that humans matter equally but not that other animals do not matter *at all*.

It seems that animals as well as humans must have some kind of moral status. Further, it is doubtful that individuals possess moral status merely in virtue of being a member of the species *homo sapiens*. Thus far, though my argument in this introduction is far from watertight, I have hopefully said a sufficient amount to show that the orthodox position that humans matter and animals do not, is far from well-defended. It is not my aim here to dispel all possible doubt that humans matter any more than animals, but merely to show that there is a case to be made in favour of the moral status of sentient animals. The rest of this thesis will be dedicated to developing and defending such a case, in which I will argue that the necessary and sufficient condition for the possession of moral status is agency and that humans and animals, in virtue of possessing this property, possess equal moral status.

²⁵ *Ibid.*, 152.

ii. Overview

I will start in Chapter 1 by elucidating the notion of moral status and explaining why it is a necessary and useful concept for moral philosophy. I argue that moral status does not plausibly permit degrees and thus there cannot be a moral hierarchy. I contend that all beings that possess moral status must possess it equally. I argue that moral status must be grounded in one's intrinsic capacities, not relational properties, and I also argue that in order for the notion of moral status to be applicable to an individual, they must possess a subjective mental life and subjective interests.

In Chapter 2 I assess some popular accounts of moral status that recognise the moral status of animals. I consider the Sentience Account, the Subject-of-a-Life Account and the Kantian Account. I argue that despite the intuitive appeal of these accounts, they all have shortcomings. Overall none of them can explain why both humans and animals matter in a plausible and convincing way, nor can they adequately explain what duties moral status holders are owed. An account that takes agency as its necessary and sufficient condition I argue, is closer to the ideal in these respects, as well as having other advantages. As such we should adopt the Agency Account of Moral Status.

In Chapters 3 and 4 I argue that sentience entails agency and thus that sentient animals are agents. I start in Chapter 3 by assessing the notion of sentience and arguing that it is best understood as the capacity to experience affective mental states i.e. mental states that have a positive or negative 'feel' to them. I argue that we can understand sensory pleasant experiences to be the result of desire satisfaction. In any case where we experience pleasant sensory feelings, we are having final *de re* desires for sensory experiences satisfied. As such I suggest, being sentient must mean that one possesses desires. Further, I argue that in almost all cases (and at least in the relevant cases here) desires are necessarily motivating. Endorsing the Humean Theory of Motivation I argue that these desires, when combined with appropriate beliefs, form motivating reasons and that these reasons are sufficient to motivate action.

In Chapter 4 I argue that sentient beings must possess beliefs and thus have the cognitive resources to form motivating reasons. Further I argue that acting on these reasons is a sufficient condition for intentional action and thus being an agent. I suggest that contrary to what many claim, acting for a reason does not require the ability to consciously reflect on

or recognise one's reason for action. Acting for a reason merely requires that reasons can rationally explain one's action in the right way. I give a sketch of what acting on a reason in this way might be like, drawing on the example of blind-sighted individuals. I then demarcate autonomy and the relevant notion of agency and the differences between autonomous and non-autonomous agents. Finally I draw the chapter to a close by clarifying some problems regarding agency over time and mental agency.

In Chapters 5 and 6 I argue that agency is a necessary and sufficient condition for full moral status. I do this by arguing that agency is necessary for full moral status since, as I argued through previous chapters, one cannot possess morally significant interests without a subjective mental life and if one has a subjective mental life then one has agency. I then argue that agency is also sufficient for full moral status. I argue that we understand full moral status holders to possess three central interests and that these interests ground three central *pro tanto* duties we owe to them. These interests are: an interest in being free from pain, an interest in continued existence and an interest in liberty. I argue that since agents possess these central interests they are full moral status holders.

I first argue that agency is necessary to have an interest in freedom from pain since pain is the frustration of desires and all beings capable of possessing desires are necessarily agents (as I argued in Chapter 3). I then argue for agency as necessary for an interest in liberty. I argue that liberty has a non-agent-affecting value for agents. Therefore in virtue of being agents, it is *pro tanto* good (through it may not contribute toward their well-being) because it allows them to determine the course of their own lives. Thus agents should be able to act as they choose and not have their choices interfered with (so long as there is no risk of significant harm).²⁶

I next argue for an interest in continued existence. I show that neither long-term future-directed desires, strong psychological connectedness over time nor categorical desires are necessary to possess an interest in continued existence. It is only final desires (that are not satisfied by death) that are necessary for an interest in continued existence. I then argue that were an autonomous agent (a paradigmatic full moral status holder) to lose their autonomy, but keep their agency, they would still possess these central interests and thus would still possess full moral status and be owed the corresponding duties.

²⁶ I will frequently refer to actions being *pro tanto* good, bad, right or wrong. Something is *pro tanto* good when it is good to some extent, though it may not be good *all things considered*. Something is *pro tanto* bad, right or wrong in the same sense.

Thus I argue that agents possess the central interests of full moral status holders and that plausibly both autonomous agents, and non-autonomous agents, have these interests for the same reason; they possess agency. Agency then is the necessary *and* sufficient for moral status I conclude. Finally I tie together the thesis by drawing some conclusions and identifying future research goals along with work that still needs to be done.

Chapter 1

What is Moral Status and Why is it Important?



Having moral status means that an entity possesses an elevated status above that of an ordinary 'thing'. It means that moral agents cannot simply treat one however they like, without concern for how one might be affected. One is owed moral consideration and this means that where the actions of moral agents affect one, they require justification. Moral status holders are also normally entitled to specific treatment in light of their elevated status. Common suggestions of what such treatment amounts to often include duties to aid them, to refrain from inflicting pain on them and ending or interfering with their lives.¹

There are many reasons for which moral agents may be required to act in certain ways toward an entity. For instance, I might be obligated to save a friend's car from destruction, not because its destruction will have any effect on the car itself, but because of the effects of the car's destruction upon my friend. The value that the car has is a result of my friend valuing it, it has no value in itself. Similarly, I might be obligated to save a work of art from destruction. In this case I ought to save the artwork because of its worth as a piece of art, but at best, this type of value could be non-instrumental aesthetic value; value it has because it is a beautiful thing in itself.²

Moral status holders possess a different kind of value to cars and artworks however. I am obligated to save a sparrow for instance, because her destruction will be bad for the sparrow herself. I ought to prevent the sparrow's destruction 'because [she] will get something out of continuing to exist'.³ This shows us the difference between moral status and other types of value. Beings that possess moral status not only possess intrinsic value, that is value in themselves for no further reason, but they matter *for their own sake*. This means that how they are treated makes a difference *to them*. Thus we ought to treat moral status holders

¹ These are merely suggested ways in which one may think moral status holders should be treated. I will give my own specification of the duties owed to moral status holders in Chapters 5 and 6.

² Kamm, *Intricate Ethics*, 228.

³ *Ibid.*, 229.

in certain ways because of the effects of our actions upon moral status holders themselves. Unlike things lacking moral status, when we fail to treat moral status holders in the appropriate way, we not only do wrong, but we also wrong *them*. To put this another way, moral status holders, unlike mere things, are the subjects of fairness and justice.⁴

Despite general agreement of this much, there is considerable debate over any further details of the concept of moral status. For instance, an important feature related to moral status is the notion of interests. The majority of theorists who discuss moral status seem to understand the possession of moral status to be intimately linked to having interests; this is certainly the case with the vast majority of those who defend the claim that animals have moral status.⁵ However, there seems to be significant disagreement over what it actually means to have interests and how they are linked to moral status.

An entity is also normally considered to have moral status in virtue of certain properties or characteristics that it possesses. Most theorists take these properties to be intrinsic to the moral status bearing entity i.e. properties that it has independent of its relations to other entities. However some philosophers argue that an entity can have moral status in virtue of the relations in which it stands to other objects in the world. These relations might be a power to produce some effect⁶ or belonging to a specific group such as a species, community or culture.⁷ However claims that the moral status can be grounded in relational properties appear to be in tension with the notion of moral status as a kind of intrinsic value.

It has also recently been questioned whether the notion of moral status is useful or even coherent. Ben Sachs has suggested that philosophers use the term in different ways and that there are at least three distinct understandings of moral status, all of them being problematic. As such, if we are to use 'moral status talk' (as Sachs refers to it), or if we are merely interested in making sense of what has already been said regarding moral status, we need to determine exactly what we mean when we use the term.

Finally, some philosophers have suggested that moral status comes in degrees, and that there is a moral hierarchy in which some beings have higher moral status than others.

⁴ Donaldson and Kymlicka, *Zoopolis*, 36.

⁵ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 27; Rollin, *Animal Rights and Human Morality*, 35–36; Donaldson and Kymlicka, *Zoopolis*, 36; DeGrazia, *Taking Animals Seriously*, 39; Cochrane, *Animal Rights Without Liberation*, 19; Regan, *The Case for Animal Rights*, 243.

⁶ Kamm, *Intricate Ethics*, 228; Liao, "The Basis of Human Moral Status."

⁷ Diamond, "Eating Meat and Eating People."

However, little has been said about what it means to have a higher or lower moral status or even if it makes sense to understand moral status as something that permits degrees. As such, the idea of degrees of moral status must be investigated, if we are to make sense of these claims.

In order to defend a robust account of moral status then, there are some problems that need to be resolved. In this chapter I tackle these problems and through doing so clarify exactly what I take the notion of moral status to mean. I will argue for an understanding of interests according to which one must have the capacity for subjective consciousness in order to possess morally significant interests. Further, since I will argue that interests are necessary for the possession of moral status, only those with a subjective mental life can possibly possess moral status. I will then argue, despite Sachs' claims to the contrary, that the notion of moral status is useful. Having moral status means one's interests should be considered and I will argue that an account of moral status ought to give some account of the central ways in which a moral status holder can be wronged. Further I will show that it is useful to know which beings possess moral status.

Next I will argue that the claim that moral status comes in degrees is either untenable or redundant. Therefore the most defensible notion of moral status is as a property that one can either entirely possess or lack. Hence, we should understand any individual who has moral status to hold it equally with any other moral status holder. Furthermore, as the Principle of Equality states we should consider similar interests of moral status holders to have similar weight. Finally, I will argue that any account of moral status must ground the possession of moral status in capacities and not relational properties, or else it cannot establish the universal obligations and intrinsic value that are integral to the notion of moral status.

I. The Moral Significance of Interests

A widely accepted implication of possessing moral status is that one ought to have one's interests considered. If this is true, then clearly a minimum requirement for the possession moral status is that one has at least some interests. There is some debate over what exactly it means to possess interests however. Peter Singer argues that unless one is sentient,

one cannot have any interests.⁸ This is because interests are at least partly dependent upon having a subjective mental life. If one cannot find any experiences pleasant or unpleasant, Singer argues, then nothing can make a difference to one's well-being or welfare.⁹ Thus if one cannot be affected by the actions of others, then there is nothing for us to take into consideration when we are deciding how we should treat such an individual.¹⁰

However this argument has been questioned by other theorists who contend that it is plausible to understand interests to mean merely having a goal or a good. What's more, they argue that we often use the term 'interests' in this way. This broader sense of interests appears to apply not only to beings with subjective mental lives, but also to non-conscious organisms. As such, defenders of this broader account of interests maintain entities such as plants could have moral status.

Through the next three sub-sections I will attempt to defend the claim that only the interests of beings with subjective mental lives are *bone fide*, morally significant interests. Whilst we might call some goals or goods 'interests', these do not carry any moral weight. I will thus argue that whilst plants and other non-conscious organisms may have instrumental, or even in some cases intrinsic value, they cannot possess moral status. Unlike beings with subjective mental lives, it is not possible for non-conscious entities to matter in themselves, for their own sake.

i. Two Accounts of Interests

Feinberg and Kuhse both independently attempt to provide an argument in defence of the claim that only the interests of subjectively conscious beings matter. They claim that although non-conscious organisms appear to have interests, in fact a plant's 'interest' in water is simply like a wall's need for paint,¹¹ or a car's need for oil.¹² Feinberg claims that in order to have interests one must have 'a sake to act for' or 'a behalf to act on' which non-conscious

⁸ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 27.

⁹ *Ibid.*; Singer, *Practical Ethics*, 57.

¹⁰ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 27. Others also argue for this account of interests in the same way: Donaldson and Kymlicka, *Zoopolis*, 36; Cochrane, *Animal Rights Without Liberation*, 37.

¹¹ Feinberg, "The Rights of Animals and Unborn Generations," 7.

¹² Kuhse, "Interests," 147.

entities lack.¹³ Feinberg also states that interests are rooted in desires and beliefs. Therefore, since one must have consciousness in order to have beliefs and desires, plants and other non-conscious organisms must lack interests.¹⁴

It is certainly true that at least some of the interests of conscious beings are formed from desires. However, it seems that we *can* plausibly act on behalf of, or for the sake of a plant by say, providing it with water, so why should we accept Feinberg's definition of an interest as a product of a desire/belief pair? Feinberg's understanding of an 'interest' seems to be unnecessarily strict and unsupported by substantial argument. As Kenneth Goodpaster notes, by accepting this analysis we 'abandon one sense in which living things [like plants] have interests'.¹⁵

Goodpaster offers us a different view, rejecting Feinberg's claim that having interests involves having beliefs and desires and that non-conscious organisms lack a sake for which to act. Goodpaster accepts a much broader notion of interests, understanding having an interest to mean something like having a goal or good.¹⁶ On this account all living things have interests and he assumes all living things have moral status. Unfortunately, much the same as Feinberg, Goodpaster neglects to provide an argument for why the 'interests' of plants, should be relevant to the question of who has moral status. In fact Goodpaster merely suggests that we are resistant to understanding plants as having interests and moral status because western society tacitly endorses hedonism. As a result, we are likely to see pleasure as good, pain as evil and non-conscious organisms (incapable of both) as unimportant.¹⁷

ii. Biological Functions as Interests

We have two accounts of having interests, the account forwarded by Singer, Feinberg and others, that interests require subjective consciousness, and the account put forward by Goodpaster, that having an interest simply means having a goal or a good. However we have no convincing argument showing either to be correct. Why should we accept either view over

¹³ Feinberg, "The Rights of Animals and Unborn Generations," 5.

¹⁴ *Ibid.*, 6.

¹⁵ Goodpaster, "On Being Morally Considerable," 319. The text in brackets is my own.

¹⁶ *Ibid.*, 320.

¹⁷ *Ibid.*, 310.

the other? Gary Varner argues that we should adopt Goodpaster's account over Feinberg's and Singer's. Varner suggests that interests in this more general sense of having a goal or good, must feature in any plausible account of interests even for conscious, highly rational beings such as humans. As such interest in this sense are clearly morally significant.

Varner contends that something need not satisfy a preference of mine to be in my interest. For instance, I can obviously have an interest in drinking water without desiring it. Therefore it seems interests unrelated to preferences must be morally significant. Furthermore, it is not the case that these interests are like those 'interests' of mere artefacts. The non-preference related interests of living organisms, can be differentiated from the supposed interests of non-living artefacts by the fact that they are biological and present due to natural selection. For any biological function, Varner claims, it is the case that: 1) the function belongs to an organ or subsystem *s* that was adaptive for one's ancestors and 2) one has the function because of the presence of subsystem *s*. The 'interests' of cars and walls fail to fulfil conditions 1) and 2). As such, Varner argues, the analogy between the interests of plants and artefacts is an over-simplification. Organisms with biological functions have interests whereas mere artefacts do not. Therefore organisms with biological functions can plausibly possess moral status.

Varner seems correct that the fulfilment of our biological functions is in our interests, however this does not give us sufficient reason to accept that all living organisms can possess interests. Varner's argument rests on the claim that we should consider biological functions to be interests wherever we find them, because they are considered to be interests in the most plausible accounts of *humans'* interests. It seems to me however, that we should consider more closely the relationship between humans' interests and biological functions, before we claim that biological functions *always* count as interests. Feinberg makes an important distinction that could be helpful here.

'A needs *x*' can mean either *a) x* is necessary for A to function correctly or *b) x* is good for A; its lack would be detrimental to A's well-being.¹⁸ If *x* is simply necessary for A to function then it need not necessarily be bad for A's function to go unfulfilled. However, it would be bad for A's function to go unfulfilled if this would be detrimental to A's well-being. Thus, in order for the fulfilment of a biological function to be an interest, it must be the case

¹⁸ Feinberg, "The Rights of Animals and Unborn Generations," 7.

that it is not just a need in sense *a* (but also in sense *b*). This seems to be the case for the biological functions of humans. However it is less clear that where the biological functions of non-conscious organisms go unfulfilled they are detrimental to their well-being. This being the case it is not clear that the biological functions of non-conscious organisms are interests in any morally significant sense.

In order to show that the biological functions of non-conscious organisms are needs in sense *b* then, it needs to be shown that the fulfilment of these functions is good for a non-conscious organism and not merely necessary for it to function correctly. To assess whether biological functions are good for non-conscious organisms I will consider why biological functions are good for humans. Let's take the function (or more properly the group of biological functions) responsible for taking in nutrition. Receiving adequate nutrition does not appear to be good for humans merely because functioning correctly is good for humans in itself. It seems much more plausible that it is good for us because receiving adequate nutrition confers *other* goods on us. By having adequate nutrition we avoid experiential suffering through lowering the risk of certain diseases and health problems, allowing ourselves to live qualitatively better lives. We also sustain ourselves, ensuring that we are not deprived of goods through our lives ending. Could receiving adequate nutrition be good for non-conscious organisms for the same reasons?

It seems doubtful. While adequate nutrition may lead to a plant avoiding certain diseases, since plants lack the capacity for conscious, it does not make sense to talk of them as having a qualitatively better or worse life. Further since they can have no preferences, views or thoughts, it could not matter to them that they are damaged through diseases or health problems. Thus, it seems that the explanation for why biological functions are considered to be interests in the case of humans, cannot justify why biological functions should be considered to be interests in the case of non-conscious organisms.

iii. Producing Offspring, Living and Flourishing

The reasons for which the fulfilment of our biological functions is good for us then, cannot be used to explain why the fulfilment of biological functions is good for non-conscious organisms. The goods that are conferred upon us by the fulfilment of our biological functions

are simply not available to non-conscious organisms. Perhaps though, the fulfilment of the biological functions of non-conscious organisms are good for them, because through fulfilling these functions they achieve some other good. Three effects of the fulfilment of biological functions might plausibly be said to be good for them (and thus explain why their biological functions should be considered to be interests). These effects are propagating, continuing to live and flourishing.

In order to attempt to discover why the fulfilment of these functions may be good for non-conscious life, I will first consider why they are good for humans (if they are good for humans at all). First I'll consider propagation. While it is true that many humans find great joy in having children, many have also had extremely rewarding, enjoyable and well-lived lives without ever having children.¹⁹ Furthermore, those who do gain something from having children, plausibly benefit from the relationships they develop with their children and through watching them grow and mature. Having children appears to be good in virtue of the experiences parents gain from it, or because there is some value in raising and forming a child. However it does not seem that having children is good for humans simply because producing a child is a good thing in itself.

Since non-conscious organisms lack any subjective consciousness, they couldn't possibly benefit from producing offspring in any psychological sense like humans can, nor could they 'raise' their offspring. As such, producing offspring couldn't be good for them for this reason. In fact, lacking any mental life it isn't clear how propagating could be good for them at all. Therefore there seems to be little reason to accept that producing offspring is good for non-conscious organisms.

Although, producing lives may not be good for a non-conscious organism, perhaps the continuing of its own life can be good for it. This is something we consider to be good for humans and so it could plausibly be good for other organisms. Why do we think that continued existence is a good thing however? Many theorists have suggested that life for both humans and animals is not good in itself but good only when, and to the extent that, it allows one to access other goods.²⁰ For example, a person's life may be valuable because it allows

¹⁹ Cochrane, *Animal Rights Without Liberation*, 74.

²⁰ Nagel, "Death," 74; Williams, "The Makropulos Case Reflections on the Tedium of Immortality"; Broome, "The Badness of Death and the Goodness of Life"; McMahan, "Death and the Value of Life," 70; Feldman, *Confrontations with the Reaper*; Harman, "The Moral Significance of Animal Pain and

them to have a family, have a successful career, and develop their artistic talents. One might even think that life is valuable because of the simpler things like being able to spend time with one's partner or friends or take one's daily walk through the park. Without life one wouldn't be able to pursue any of these goods. Further without any goods to pursue, life wouldn't have any value. Mere existence in itself, these theorists argue isn't worth having.²¹ Other theorists such as Oderberg have argued that human life is intrinsically valuable. Because of the goods that life allows us to pursue, Oderberg argues, living should be considered a good in itself. Thus, even if we are deprived of all other goods in our lives, we still have one good to pursue: living itself.²²

What can these two accounts tell us about the value of life for non-conscious organisms? Plausibly most of the goods which life enables us to pursue are experiential e.g. spending time with one's family, or at least require the capacity for experience in order to fulfil them e.g. having a successful career. Hence since such organisms lack any mental life they cannot be said to possess of the goods which humans pursue. Life cannot be said to be good for them because they gain further goods. Could the lives of non-conscious organisms be intrinsically valuable in the way Oderberg describes however? Though Oderberg claims that human life has intrinsic value, this value lies in the fact that life is the source of further goods. Thus if the life of non-conscious organisms does not grant one opportunities to pursue any other goods, then life cannot have this value. Since there don't appear to be any goods for non-conscious organisms to pursue, it doesn't seem that their lives could be intrinsically valuable in the sense Oderberg believes human life to be. Therefore it seems that death cannot be bad for non-conscious organisms and living cannot be good for them. The possession of subjective experience appears to be a necessary condition for living to be good for an individual.

Finally, one might think that the fulfilment of the biological functions of non-conscious organisms results in flourishing. Flourishing one might think is a good wherever it occurs and thus must even be good for non-conscious organisms. This has intuitive appeal, flourishing is something that we consider to be good for humans and does not appear to require subjective consciousness, after all we frequently refer to plants as 'flourishing'. However this appeal to

Animal Death," 730; DeGrazia, *Taking Animals Seriously*, 231–32; Regan, *The Case for Animal Rights*, 99–100.

²¹ Williams, "The Makropulos Case Reflections on the Tedium of Immortality."

²² Oderberg, *Applied Ethics*, 67.

the common usage of the term is misleading. It seems that flourishing may be being used equivocally in the case of plants and humans.²³

When we refer to flourishing being good for humans we mean something like ‘living a rewarding or worthwhile life’. In fact, claiming one is flourishing seems to be a shorthand way of stating that one is successfully pursuing goods or perhaps, advancing one’s interests. This being the case however, it can’t be true that the fulfilment of its biological functions is good for a non-conscious organism because it allows it to flourish. This would simply mean that it is good for an organism to fulfil its biological functions because in doing so it successfully pursues goods or advances its interests. Since it seems that there are no goods for a non-conscious organism to pursue, and no interests for it to advance, this justification is explanatorily empty. There would need to be at least one other good for an organism to pursue in order for flourishing to be good for it.²⁴ Flourishing, in any morally significant sense simply isn’t possible in the absence of goods to pursue.

Therefore neither living, propagating nor flourishing can be said to be good for non-conscious organisms. As such, whilst the fulfilment of biological functions is often good for sentient beings, there is no good reason to believe that the same is true of non-conscious organisms. Varner’s claim that biological functions are interests (whomever they belong to) is false. The biological functions of non-conscious organisms could only count as interests if the fulfilment of these functions would be good for them, or the lack of fulfilment of them would be bad for them. It seems that without a subjective mental life there are no goods which one can pursue, and as a result nothing can be good or bad for non-conscious organisms. Thus having subjective consciousness should be taken to be a necessary condition for having interests, and so moral status.

II. Is the Notion of Moral Status Useful?

An important difference between Feinberg’s and Goodpaster’s accounts of moral status is the work moral status is supposed to do. If one has moral status in Feinberg’s sense

²³ Feinberg, “The Rights of Animals and Unborn Generations,” 7.

²⁴ This doesn’t mean, of course, that we need to stop referring to plants as ‘flourishing’ or ‘having interests’ but simply that these terms, when used to describe non-conscious organisms, lack any moral significance.

then one matters and one is entitled to some specific treatment (Feinberg thinks that moral status holder have certain rights, though one needn't understand this point in terms of rights specifically). Thus if we know that one has moral status on Feinberg's account this can inform us that they matter and that we owe them some specific duties. However, Goodpaster wants something much less demanding. If a being has moral status on Goodpaster's account it merely has 'moral considerability', that is to say, it is the kind of entity that *can* matter morally. Whether any such beings do matter, or how much they matter, or how we ought to treat such individuals, are separate questions for Goodpaster. Clearly there is a distinct difference in these two uses of the term 'moral status'. Goodpaster and Feinberg see the role of moral status very differently. This brings us to some important questions regarding moral status: what work does the notion of moral status do? What work should it do? And is the notion of moral status actually useful?

Benjamin Sachs argues that all talk of moral status is redundant and often confusing.²⁵ He claims that there are three distinct ways in which one can understand moral status: *the No Further Fact View*, *the Further Fact View* and *the Expressivist View*, all of which he claims, are not useful. I'll explain each in turn before arguing that the circumstances are not as dire as Sachs believes and that the Further Fact View is a defensible understanding of moral status. What's more the notion of moral status is more useful than any of Sachs' options suggest.

On the No Further Fact View, moral status just *is* the possession of some property or properties which make one eligible to be considered e.g. consciousness, rationality or being alive.²⁶ The ascription of moral status is simply a linguistic convenience used to indicate the presence of some morally significant property (or properties). If this is the case, Sachs argues, then it will be easier to understand arguments for why a given being matters if theorists refer to the relevant morally significant properties themselves, rather than moral status. Referring to the notion of moral status merely adds extra premises to arguments that we could apparently do without. Thus, to assess any argument involving the notion of moral status, we would be required to translate the argument back into terms of morally significant properties.²⁷ As such it seems the concept of 'moral status' is redundant and unhelpful.

²⁵ Benjamin Sachs, "The Status of Moral Status," *Pacific Philosophical Quarterly* 92, no. 1 (2011), p. 89.

²⁶ Sachs refers to these properties as 'moral properties' however, in order to differentiate these from those properties we normally refer to as 'moral properties' (goodness, rightness, etc.) I shall refer to properties which ground moral status as *morally significant properties*.

²⁷ Sachs, "The Status of Moral Status," 92.

Sachs notes that some ethicists have argued that moral status talk is useful because it allows us to convey not just that a being has some morally significant property (or properties) but also that in light of these properties they are morally considerable. Thus if I claim that A has moral status, I am recognising that A's interests have moral weight and I ought to consider whether, and to what extent my actions will affect them when I act.²⁸ However Sachs argues that it isn't obvious that we have a duty to consider the interests of moral status holders *per se*. So long as we engage in the right action and have the right intentions, it does not matter whether we have arrived at this result through deliberating over the interests of moral status holders. Furthermore, Sachs maintains, even if we are obligated to consider who our actions will effect, it is not clear what this duty requires.²⁹ Should we consider the interests of all moral status holders before getting up in the morning? Obviously not. This would be far too demanding. If we were obligated to consider every moral status holders' interests before acting, we would never do anything. Plausibly we ought only to consider the interests of those that will be affected by any given action. However if this is the case then how can we identify who will and who will not be affected by some action?

Moreover Sachs notes, even if we could ascertain who would be affected by a given action, there are sometimes good reasons not to consider an individual's interests, even if they would be affected. For instance, where people have evil interests and where full deliberation of certain actions may cause offence.³⁰ For example, considering whether our action will frustrate another's racist interests seems unimportant (and perhaps morally questionable). It seems that only some of the interests of moral status holders ought to be considered. Thus since there is no clear way to determine which interests we ought to consider we cannot be morally required to consider the interests of moral status holders. Therefore Sachs concludes, moral status alone cannot inform us about whether we should consider an individual's interests.³¹

On the Further Fact View moral status is something one has in addition to the morally significant properties that one possesses. Moral status talk works as an intermediary between moral and non-moral facts. It appears to allow us to make moral claims such as 'A is owed duties' without appealing directly to non-moral facts such as 'A has sentience' or 'A has rationality'. By doing this it appears that one can successfully derive a moral fact from a non-

²⁸ Ibid., 94.

²⁹ Ibid.

³⁰ Ibid., 95.

³¹ Ibid.

moral fact but avoid the naturalistic fallacy. Instead of arguing that 'A is owed duties because A is sentient' one can argue that 'A is owed duties because A has moral status (and A has moral status because A has sentience)'. This understanding of moral status seems to legitimise certain kinds of non-moral to moral inference. Since this notion of moral status is not strictly defined, and is normally explained by referring to further claims in the form: 'to have moral status is to have *x*', Sachs maintains that using moral status in this way appears useful.³² It is considered as 'something we-know-not-what' (in Sachs' own words) and a useful buffer to sit in the chain of reasoning between moral and non-moral facts. So on this account, that one has moral status is a further fact about an individual, in addition to the fact that one has some morally significant property, such as sentience.

Unfortunately, Sachs argues, whilst using the term 'moral status' may make it appear as though the naturalistic fallacy has been avoided, since the meaning of 'moral status' is simply 'something we-know-not-what', this argumentative move fails. We are committing the naturalistic fallacy all the same by assuming that we can infer facts about moral status (which seems to be at least partially a moral property) from non-moral facts. Therefore, it would be better, Sachs argues, if we just go ahead and commit the naturalistic fallacy, rather than committing it and then attempting to hide our argument through using the concept of moral status. The inclusion of moral status in this chain of reasoning adds nothing to the argument and merely misleads the reader. As such, it seems that we would have clearer arguments if we simply did away with the notion of moral status altogether.³³ Sachs does add however that if it was the case that we *could* define moral status, and that moral status *could* mediate between moral and non-moral facts, then moral status talk *would* be a useful concept. However until this happens, he suggests, we should abstain from using such notions (at least in this way).³⁴

Finally, on the Expressivist View of moral status, claims such as '*x* has moral status' and '*x* should be granted moral status' are merely statements of one's intentions to treat given individuals in certain ways, and implore others to do the same. Moral status talk on the Expressivist View then, lacks any normative power.³⁵ Sachs is quick to point out that understanding moral status in this way is not helpful if we are trying to determine who ought to be given moral consideration. Expressions of intent are not what we need, he argues, and

³² Ibid., 96.

³³ Ibid., 97.

³⁴ Ibid.

³⁵ Ibid., 99.

have no place in moral reasoning on this subject.³⁶ We should endeavour to develop a theory to give us reasons to act on our intentions to treat some individuals in certain ways, not simply state them.³⁷

Of these three accounts it seems to me that the No Further Fact View is closest to how moral status should be understood. This is because it is the most intuitive, useful, and the least problematic. Sachs seems to be correct that moral status talk is simply a way to state that a being has morally significant property (or properties). Nonetheless, this does not mean that we should avoid using the term 'moral status'. While moral status talk adds extra complexity to one's argument, it doesn't seem to be as difficult as Sachs assumes to assess arguments that incorporate moral status talk.

Furthermore, moral status ascriptions seem to clearly and succinctly confer the message that an individual matters, in a way in which mere ascriptions of morally significant properties cannot. Thus, if I endorse the Sentience Account, when I state that 'A has moral status' I am claiming that A has the capacity for sentience. However I am also making it clear to my listener that I take this capacity to be morally significant and to be at least sufficient for one to matter. Conversely, when I state that 'A has the capacity for sentience' the implications of this phrase are not immediately clear and more information is needed to grasp the point that I am making. Thus using the term 'moral status' seems to better convey the implication that some individual matters. Hence, while it may be that moral status talk is redundant in the sense that we need not use it in arguments concerning which beings matter, there seems to be pragmatic reasons to keep this linguistic convenience.

Assuming my argument holds however, Sachs would still maintain that moral status understood according to the No Further Fact View is redundant in a deeper philosophical sense, because it is not useful to know that a given individual ought to be considered. Moral status would be a useful notion if we had a duty to consider who might be affected by our actions but we have no such duty Sachs maintains. So long as we do the right thing it doesn't appear to be necessary to deliberate over who ought to be considered. Further, even if this were not the case, there doesn't appear to be a clear way to fulfil this duty since there are individuals and interests that we need not (and perhaps ought not) consider when deliberating over any given course of action.

³⁶ Ibid.

³⁷ Ibid.

Sachs seems to be missing something here though, moral status doesn't seem to be redundant in this sense either. Firstly, the claim that so long as we do the right thing, from the right intentions, it doesn't matter whether we consider who might be affected by our actions strikes me as very odd. I cannot see how one could do the right thing, from the right intention *unless* one was to consider who would be affected. I surely cannot have the intention that stealing is wrong and that I shouldn't steal from my boss unless I believe my boss to be someone that matters and someone whose interests would be affected by my action. Of course I needn't deliberate over whether my boss's interests should be considered in a highly conscious way. However I will have to recognise in some sense that she is an individual who has moral status and thus ought to have her interests considered if I am to reach the conclusion that I shouldn't steal from her because doing so is wrong.

Secondly whilst Sachs may be right that there is no satisfactory formalised account of deliberative duties, this is beside the point. Just because no account has been specified (or can easily be specified) this does not mean that we are unable to fulfil such duties. We seem to be able to consider who or what our actions may affect fairly reliably without any specifiable method of doing so. We can with reasonable accuracy discern those individuals that will be affected by a given action from those that won't. In fact, we do this as part of any judgement about whether action x is permissible in circumstance y . For the most part we seem to engage in this deliberation rather intuitively, such considerations often do not feature as part of our conscious reasoning over the permissibility of a given action. Nonetheless we gain consistent results. Rarely am I surprised by who my actions affect. Moreover, when I am surprised, it is rarer still that I find that I have *drastically* miscalculated who my actions would affect.

When I consider whether I should steal money from the cash till at work to pay for an emergency operation for my wife, I understand that I should consider the interests of my wife, my boss, the other staff and myself, and that I need not consider the interests of a pack of stray dogs in Bristol, my 17th century ancestors or the Prime Minister. It may well be the case that an account of deliberative duties would help me to better assess who I should and should not consider. It may also be of particular help in borderline cases such as the interests of the customers of the shop in this scenario. However, it is certainly false that we cannot engage in reasoning over whose interests matter in any given situation, since we all engage in such reasoning on a daily basis. It is simply the case that we do this intuitively.

Moreover, Sachs is undoubtedly right that there are sometimes good reasons not to consider individuals' interests, their interests being evil for example. However, the fact that

there are some interests that we ought not to consider in certain circumstances seems irrelevant here. That one has moral status does not give moral agents an all things considered reason to satisfy one's interests (or not frustrate one's interests). Moral status, tells us whose interests should (and whose should not) be given due consideration before we consider the specifics of the situation or context. That an individual has moral status means that moral agents have a *pro tanto* obligation to consider an individual's interests. In other words, our obligation is defeasible. So if an individual's interests turn out to be evil for example, then we have good reason not to concern ourselves with their satisfaction.

Thus when the time comes for us to consider whether action x is permissible, one knows who, out of all those entities affected, one should consider and who one need not. For instance, if one was considering whether to have chicken for dinner, and held the belief that only apes, dolphins and humans have moral status, one need not even consider whether the chicken's interest in continued existence outweighs one's own interest in eating them. Or whether there are mitigating factors in play, such as the fact that there is no other food source available. Moral status claims allow one to shortcut around this unnecessary reasoning, eliminating the need to ask certain questions. So we might say that claims about who possesses moral status, prime us so we can work quicker when considering whose interests might be affected in any given situation. Thus it seems the notion of moral status makes moral deliberation *more* efficient.

Thus, the notion of moral status is useful. It can withstand the arguments put forward by Sachs and can inform us which individuals (*pro tanto*) ought to have their interests considered. It is also a useful linguistic convenience. In addition to this function I think that moral status has a further role. The most intuitive accounts of moral status are those which not only give us a plausible account of which individuals matter and why, but also inform us of the central and basic ways in which moral status holders *pro tanto* ought to be treated. Moral status accounts allow us to determine the central ways in which moral status holders can be benefited and harmed, through the property in which grounds moral status. Thus by appealing to the property that grounds moral status, we should be able to explain what is chiefly bad or wrong about the central ways in which a moral status holder can be harmed or wronged. Kant's and Singer's accounts of moral status exemplify this function of a moral status account well.

According to Kant, those individuals that matter are those that possess autonomy, from this we can determine that autonomy promoting actions are *pro tanto* good for moral

status holders and autonomy reducing actions are *pro tanto* bad for moral status holders. On Singer's account individuals matter because they are sentient and so they can be harmed by experiencing unpleasant experiences and benefited by experiencing pleasant experiences. From this we can easily work out some central actions that would be *pro tanto* good and bad for moral status holders on these accounts.

Of course an account of moral status cannot establish a whole normative theory, nor will it even establish a whole theory of well-being. In fact a moral status account alone cannot settle that *any* specific action is right or wrong. We will not be able to determine what is all things considered good, bad, right or wrong for a moral status holder merely from the fact that they possess moral status.³⁸ Whether for instance, inflicting x amount of pain on a moral status holder is wrong all things considered will depend not just upon an account of well-being, but also upon the normative principles one holds. One may endorse a consequentialist principle that happiness should be maximised, or one may think that rights protect individuals from suffering significant harms. These principles will play a large part in determining right and wrong treatment.

Further the wrongness and rightness of many actions also depends on one's relations to others, the consequences of one's actions, whether they are just and fair in the circumstances, whether they instantiate some virtue and many other factors besides. Thus I only suggest that we should be able to use a moral status account to determine the *central* ways in which something can be *pro tanto* right or wrong, or good or bad for a moral status holder. Where something is *pro tanto* good or bad for a moral status holder it will feed into an account of well-being which will in turn feed into an account of wrongness. Thus, assuming that on any plausible normative moral theory, contributing to the well-being of a moral status holder is *pro tanto* right and harming a moral status holder is *pro tanto* wrong, we can determine the central and basic ways in which it is *pro tanto* right and wrong to treat a moral status holder.³⁹

If moral status ought to be understood in this way, this may explain why Goodpaster's account seems so 'cheap'. Goodpaster maintains that being alive is the capacity that grants one moral consideration. Goodpaster's account does not give us any idea of how moral status

³⁸ I use good and bad as well as right and wrong since I take it that there is treatment that is neither good nor bad for moral status holders but which is right or wrong. For instance I argue (in Chapter 5 Sec.III) that restricting the liberty of a moral status holder is wrong but may not be bad for them.

³⁹ I will sometimes talk about what is *pro tanto* good and bad for moral status holders and at other times what is *pro tanto* right and wrong.

holders ought to be treated. In doing this he removes a function of moral status and stops moral status doing the work that I, and many others want it to do.⁴⁰ It appears to me that we should understand a being with moral status to be a being whose interests we have a *pro tanto* reason to consider and who *pro tanto* ought to be treated in certain ways.

In summary then, moral status talk is not redundant as Sachs claims. There are good reasons to use the notion of moral status from a linguistic perspective; it succinctly makes clear that one is making a claim about an individual's entitlement to be considered. Further the notion of moral status is not philosophically redundant either, since it is useful to know that certain beings have moral status and others do not, because it allows us to take argumentative shortcuts. Finally, it also seems to be the case that a moral status account plays a role in constructing an account of wrongness. A good moral status account will inform one's account of well-being and wrongness by making clear the central ways in which actions are *pro tanto* good and bad, and right and wrong for a moral status holder.

III. Does Moral Status come in Degrees?

Now we are clear on what moral status means, we can address the question of whether moral status come in degrees. Whilst Feinberg understands all beings that have moral status to matter equally and thus possess equal moral status, Goodpaster endorses a moral hierarchy. Goodpaster believes that moral status holders matter to different degrees and thus some beings have more or a higher moral status than others. Or, in Goodpaster's vernacular: while all living things are *morally considerable*, they do not have equal *moral significance*. In this section I will assess whether the notion of moral status can be understood to come in degrees. I will argue that there is no coherent way in which this can be the case if what I have said about moral status so far is correct.

The standard reason one might think that an individual has a higher moral status is because of the kind of interests they have. Often having stronger interests, more varied interests or some specific interests that individuals with a lower moral status lack, is thought to grant one a higher degree of moral status. Assuming that this is the case however, this doesn't tell us what exactly having a higher or lower moral status means. How does an

⁴⁰ As I note above both Singer and Kant rely on their accounts of moral status to give an account of wrong treatment towards moral status holders. These accounts of wrong treatment then play a central role in their respective normative theories.

individual's moral status make a difference to the treatment we owe to them? David DeGrazia describes two possible ways in which one might conceive of degrees of moral status: The Unequal Consideration Model and the Unequal Interests Model of Moral Status.⁴¹

On the Unequal Consideration Model the weight of any given interest is determined in part by who possesses the interest. As the name suggests, interests on this account are not given equal consideration where they are similar in content and strength. The higher degree of moral status one has, the greater moral weight one's interests are given. Thus, the Unequal Consideration Model suggests that it is worse to frustrate some interest where it belongs to an individual with a higher moral status, than it would be to frustrate the very same interest where it belongs to an individual with a lower moral status.⁴²

On the Unequal Interests Model the weight of any given interest is not affected by the moral status of the interest holder. The weight of an individual's interest is determined solely by the content and strength of the given interest itself.⁴³ Thus if two individuals with different degrees of moral status both have the very same interest in the absence of pain for example, their interests have equal moral weight and ought to be equally considered. The fact that one individual possesses a higher moral status is irrelevant to the question of how we ought to treat them.

If I am to show that moral status does not come in degrees then I must show that both of these models are untenable. I'll consider the Unequal Consideration Model first. I will argue that the Unequal Consideration Model fails to justify its central claim and seems to 'double-count' the interests of some moral status holders making it seem unfair and an untenable model of degrees of moral status. After this I will consider the Unequal Interests Model and argue that though it avoids this problem it is not clear why we should understand it to be model of degrees of moral status. It more plausibly seems to model equal moral status among moral status holders. As such, neither proposed models adequately make clear how moral status could permit degrees.

i. The Unequal Consideration Model

⁴¹ DeGrazia, "Moral Status as a Matter of Degree?"

⁴² *Ibid.*, 187.

⁴³ DeGrazia, "Moral Status as a Matter of Degree?," 188.

The Unequal Consideration Model suggests that having a higher moral status gives one's interests greater moral weight. Thus the very same interest when possessed by a being who has a higher moral status, is weightier than when it is possessed by a being who has lower moral status. However there appears to be no non-arbitrary way to defend this claim. Certainly when comparing two interests of different strength, the stronger interest should be satisfied over the weaker one (other things being equal). Thus if I have a slight interest in the absence of pain, but you have a very strong interest in the absence of pain, then clearly, where pain has to be inflicted upon one of us, it should be inflicted upon me, since it will harm me less. However it is not clear why, when two individuals share the very same interest, of the same strength, that these should be given different weights.

Let us consider a human and a cat. Plausibly humans have deeper interests, and a greater variety of interests, than cats. Humans then will have a higher moral status than cats. However let's also assume humans and cats have the very same interest, of the very same strength, in the absence of pain. According to the Unequal Consideration Model, it must always be morally worse (other things being equal) for a human to suffer x amount of pain than for a cat to suffer the x amount of pain. This seems odd. If one's degree of moral status is determined by the strength and/or variety of one's interests overall, then the moral weight of one's interest in pain for example is determined by the variety and/or strength of one's *other* interests. Why should this be the case? What does the fact that I have a deeper interest in being able to determine the course of my own life or the fact that I have an interest in becoming a successful artist for example have to do with the moral weight of my interest in the absence of pain relative to a cat's? Nothing seemingly, assuming that these other interests remain unaffected by the frustration or satisfaction of my interest in the absence of pain. That being the case, why should the weight of my interest in not experiencing pain be determined by the depth or variety of my other unrelated interests?

In other cases, where the interest being considered is also one of the interests that grants one higher moral status, there appears to be a double-counting of interests. Consider a situation where both a human and a cat have an interest with the same content, let's say an interest in liberty, but their interests have different strengths. In such a case the human's interest is a stronger interest and thus is morally weightier. However the human is also awarded a higher moral status because they have a stronger interest in liberty than the cat and this higher moral status makes their interest in liberty even weightier. Thus we have granted

moral weight to the human's interest because it is stronger and then we have granted more moral weight to this interest for the same reason. This doesn't seem right at all.

Furthermore, we tend not to use this model when comparing the competing interests of humans. In such cases we equally consider interests, taking the same interest in different individuals, to carry the same weight, despite differences in the depth and variety of one's other interests. For example, most would think it equally wrong (other things being equal) to inflict a moderate amount of suffering on a 2 year old child as it would be to inflict the same amount of suffering on a normal adult human, even though an adult has overall, much deeper and a greater variety of interests.⁴⁴ This being the case regarding other humans, it seems that we ought to adopt this position with regard to all moral status holders. We should equally consider similar interests whomever they belong to.

One should note that there are some important caveats to this point however. Firstly, we often label different interests with the same name. For instance, the interest in continued existence. A chameleon's interest in continued existence is different from a human's interest in continued existence because death for a chameleon and death for a human are different. Death for a human involves a greater loss and different kinds of losses, loss of certain opportunities unavailable to a chameleon. Thus we should not assume that because two individuals possess interests with the same name that these interests share the same content. If two interests do not genuinely have the same content then they do not require the same treatment.⁴⁵

Secondly one might point out, as Harman does, that *it is* in fact worse to inflict x amount of pain on a human than a cat because a human will suffer deeper psychological trauma from the same amount of pain. As such the human's and the cat's interests in not suffering x amount of pain are not in fact the same interest.⁴⁶ I will not take a stand on whether Harman is correct about this but I accept that if she is right then one *would* be morally required to inflict the pain on the cat rather than the human. However, this does not mean that we shouldn't accept equal consideration, it simply means that my example is faulty. As Singer notes, discerning the content of given interests is a difficult task which we should take great care in completing.⁴⁷

⁴⁴ DeGrazia seems to accept this point: DeGrazia, "Moral Status as a Matter of Degree?," 187.

⁴⁵ Singer, *Practical Ethics*, 59.

⁴⁶ Harman, "The Potentiality Problem," 180.

⁴⁷ Singer, *Practical Ethics*, 58.

A further problem in recognising the content of interests is that it is often difficult to isolate given interests. There may be knock on effects, and these may not be immediately apparent. For instance, failing to satisfy a child's interest now may affect them as an adult.⁴⁸ However, this again is a mere practical problem in determining the interests of a moral status holder. Despite the difficulties in discovering the content of an interest, in cases where the individuals concerned genuinely do share the same interest, it seems that these interests should be given equal consideration.

It appears to me that The Unequal Consideration Model is untenable. The model cannot plausibly justify its central claim that the interests of different moral status holders ought to be given different weights depending upon who possesses them. It seems that we ought to equally consider the similar interests of moral status as we do in the case humans with interests of different varieties and depths.

ii. The Unequal Interests Model

Perhaps the Unequal Interests Model can provide us with a more plausible model of degrees of moral status. Importantly the Unequal Interests Model allows for the equal consideration of interests and thus has *prima facie* intuitive appeal. The more plausible account of degrees of moral status then appears to be the Unequal Interests Model, according to which similar interests are given similar weight regardless of the holder of the interests. However, whilst understanding degrees of moral status in this way avoids the problems of the Unequal Consideration Model, it may not be useful and could even be a hindrance. As such I will argue, it is not clear that we should understand moral status to permit degrees at all.

On the Unequal Interests Model like interests are given equal consideration regardless of who possesses them. As such, having a higher moral status does not make one's interests weightier. However, since one's degree of moral status makes no direct difference to the weight of one's interests, or one's treatment, it is not clear what role it has to play. Consider again the human and the cat. The human has higher moral status, the cat has lower moral status and both have the very same interest in the absence of pain. According to this account since they share the very same interest in the absence of pain they are owed exactly the same treatment regarding this interest. Thus their degree of moral status plays no role in moral

⁴⁸ Harman, "The Potentiality Problem," 180.

decision making. Information about their degree of moral status is completely unnecessary to make any decision regarding whose interests we should preference. The depth of specific interests themselves can fully explain why we should preference the first individual's interest over the second individual's. So where degrees of moral status are secured by depth or variety of interests, talk of higher and lower degrees of moral status on this model (unlike the Unequal Consideration Model) seems redundant.⁴⁹

One might note that in fact degrees of moral status do play some role. If we know that an individual has higher moral status, we know that they have deeper or more varied interests than a being with lower moral status. Assuming that one has a clearly established hierarchy with distinct and determinate degrees of moral status, and that possessing a specific degree of moral status entitles one to specific treatment, knowing the degree of moral status one has may be useful. If for example we owe a strong duty of non-interference and a duty to not inflict pain to full moral status holders because of their strong interests in liberty and the absence of pain then the notion of full moral status is useful as a helpful linguistic convenience. By stating that some being has full moral status we can quickly and efficiently give some idea of the kind of treatment they are owed. Thus though it may not play a role in moral reasoning, conceptualising beings as possessing different degrees of moral status may be practically useful, much like how I suggested that moral status is useful.

However this practical use of degrees of moral status is mitigated by the fact that using degrees of moral status in this way is misleading. Talk of higher and lower degrees of moral status implies that those with higher moral status should have their interests prioritised over those with lower moral status, as is the case on the Unequal Consideration Model. However degrees of moral status do not have this function on this account. Thus when used as linguistic convenience to communicate an individual's possession of a set of interests it may easily be misinterpreted. This being the case it doesn't seem referring to one's degree of moral status will be all that practically useful. Furthermore, misinterpretation of the notion of higher moral status could lead to the undervaluing of some (often the most vulnerable) individuals' interests. At best degrees of moral status seem misleading on this model but at worst they are dangerous.

⁴⁹ DeGrazia accepts that talk of degrees of moral status on the Unequal Interests Model is redundant as does Harman: DeGrazia, "Moral Status as a Matter of Degree?," 191; Harman, "The Potentiality Problem," 187.

Finally, it is not clear how understanding moral status to come in degrees on this model is functionally different from understanding moral status to be a binary concept. On an account of moral status that does not permit degrees, like interests should be given equal consideration. On an account of moral status that permits degrees in the way this model suggests, exactly the same is the case. On the former account all moral status holders have equal moral status. On the latter account individuals have different degrees of moral status, yet an individual's degree of moral status is not necessary to determine what treatment is owed to them. Given the risk of misinterpretation and the apparent philosophical redundancy of degrees of moral status on this model there seems to be little motivation to conceptualise such an account as a moral status account that permits degrees. It seems prudent to forgo using such notions and instead conceptualise moral status as a binary concept that one either possesses or lacks. Such an understanding of moral status is clear and unambiguous.

In summary, the central claim of the Unequal Consideration Model is unjustified and the Unequal Interests Model seems to be more plausibly understood as a binary model of moral status. Therefore, unless some different model is shown to be successful there appears to be no satisfactory way in which we can understand degrees of moral status, at least not in any way where degrees of moral status play a functional role in moral reasoning. I suggest that we should understand moral status as being a binary concept which one can either possess or lack. Where one has moral status one is entitled one to have one's interests impartially considered.

IV. Relationalism and Individualism

On all the accounts of moral status discussed so far, the ground of moral status has been some intrinsic capacity: being alive, sentient or rational. However some theorists argue that the criterion for moral status is not some capacity or intrinsic property but a relation one has to others. Such relationist views often maintain that those individuals that matter are those with which we are co-members of a society, culture or species. Cora Diamond's view that those individuals that matter are those with which we share a human life is a paradigmatic example of such a position.⁵⁰ Following McMahan, whilst I will not deny the possibility of

⁵⁰ Diamond, "Eating Meat and Eating People." Accounts such as Diamond's are not the same as the 'nature-of-the-kind' accounts discussed in the introduction that maintain that being a member of the species *homo sapiens* grants one moral status. Relationist accounts maintain that some individual

one's relations granting one special duties towards some individuals, I do deny that such relations can ground moral status.⁵¹

Some argue that it is simply counter-intuitive that one's relational properties, like membership to a group could be morally significant.⁵² Others suggest that even if some relational properties may be morally significant, the relational property of belonging to a species is intuitively not one of these.⁵³ Such theorists support *moral individualism* which contends that whether someone matters can be determined only by their intrinsic properties. Thus one's biological, historical, social or any other kind of relations cannot ground moral status. Whilst these relational properties may entitle one to some special treatment (one might think one ought to treat one's mother better than a stranger *because* she is one's mother) they cannot ground moral status itself. Relational accounts, Rachels argues, attempts to determine how one ought to be treated on the basis of *others'* properties and this is irrational and unfair.⁵⁴ Indeed, I think Rachels is right, but more can be said against such accounts.

As stated at the start of this chapter, moral status is usually taken to be a kind of intrinsic or final value that beings possess. So possessors of moral status are beings that matter in themselves and for their own sake. What is more, if an individual matters then all moral agents are *pro tanto* obligated to give moral status holders certain treatment regardless of other facts about them. However on a relationist account of moral status all moral agents are not obligated to give the same treatment to moral status holders. Since relationist accounts determine moral status on the basis of other individuals' relations to oneself, they can only give *some* moral agents reason to treat a given group of individuals in certain ways. McMahan, explains this point in terms of the kinds of reasons which capacity-based and relationist accounts of moral status give us.

Relationist accounts only obligate those moral agents who share an appropriate relation to the relevant group of individuals. Thus relationist accounts only give us agent-relative reasons to treat the relevant group of individuals in a certain way.⁵⁵ For instance, Diamond maintains that those individuals that matter are those with whom we share a human

matters because of their relation to me. The former accounts however argue that being a human is a morally significant property independent of whether I possess such a property.

⁵¹ McMahan, "Our Fellow Creatures," 359; Cochrane, *Animal Rights Without Liberation*, 33.

⁵² Rachels, *Created from Animals*, 185.

⁵³ McMahan, "Our Fellow Creatures," 354.

⁵⁴ Rachels, *Created from Animals*, 187.

⁵⁵ McMahan, "Our Fellow Creatures," 360.

life. As such Diamond's account can only give other humans reasons to treat humans in certain ways. This account could not give a non-human moral agent reason to treat humans any differently from any other species. Of course, as McMahan notes, in reality we do not know of any moral agents who are not human, and thus this is not a problem that will manifest in practice. However this is beside the point.⁵⁶ Accounts like this fail to establish the independent value of moral status holders and instead argue that we should treat them in various ways merely because of our contingent relations to them.

In contrast, intrinsic capacity accounts provide all moral agents with reason to treat certain beings in specific ways. They are impartial in who they obligate and as such establish universal obligations applicable to all moral agents to treat those individuals that meet the criteria in appropriate ways. As McMahan puts it, capacity accounts give us agent-neutral reasons to treat a group of individuals in certain ways.⁵⁷ On capacity accounts then, assuming humans matter, even if one is not a human, one ought to treat humans in certain ways in virtue of their intrinsic capacities. The fact that one lacks any kind of relation to humans is irrelevant. The relevant group of individuals have independent value established by their possession of some capacity. This I maintain is necessary for a *bona fide* account of moral status. Unless an account can establish universal obligations it seems that it isn't an account of why some individuals matter, but merely an account of why they matter *to oneself*.

For this reason I do not believe that moral status can be grounded by one's relations to other beings but can only be grounded by the possession of some intrinsic capacity or capacities. However this does not mean that one's relations to other beings are entirely unimportant. Special relationships may grant us some moral entitlements and responsibilities. For instance, my view here is completely compatible with a position such as Donaldson and Kymlicka's or Anderson's in which one's membership of a society, community or culture grounds some duties or rights.⁵⁸ I do not deny the moral import of such relations but merely claim that such accounts cannot explain what is *fundamentally important* about individuals and thus such accounts will not be relevant to the question of what grounds moral status.

⁵⁶ Ibid.

⁵⁷ Ibid., 355.

⁵⁸ Donaldson and Kymlicka, *Zoopolis*; Anderson, "Animal Rights and the Value of Non-Human Life."

Conclusion

In conclusion, it seems that only beings capable of subjective consciousness can have moral status because they are the only type of being that can be said to have interests in any morally significant way. Despite claims that we should consider the biological functions of plants and other non-conscious organisms to be interests, it appears that we should not do so unless we have reason to believe that the fulfilment of these functions would be good for the organisms themselves. I have argued we do not have such reason because the most plausible justifications for why the fulfilment of these functions would be good for them (because it allows them to continue to live, propagate or flourish) have been shown to be inadequate. While living, producing offspring and flourishing may be good for humans and other conscious beings, this is because of the benefits that we gain from these in virtue of our capacity for subjective mental life.

I have defended the notion of moral status from charges of redundancy and argued that moral status has some clear functions to perform in a normative theory. I argued that having moral status means that moral agents ought to consider one's interests, and that knowing who possesses moral status allows moral agents to take various short-cuts in deciding whose interests one should consider, in any specific situation. Further, we should be able to determine the central ways in which a moral status holder *pro tanto* ought and ought not to be treated from an account of moral status. Thus by knowing that an individual possess moral status we know something about how they ought to be treated.

Finally, I argued that moral status does not appear to permit degrees as both proposed models of degrees of moral status have been shown to be inadequate. The Unequal Consideration Model appears unjustified and the Unequal Interests Model appears to be best understood, not as an model of degrees of moral status, but as an account of *equal* moral status. Thus, we should understand moral status to be a binary concept which one can either have or lack. Moreover, other things being equal, we should understand having moral status to mean that one is entitled to have one's interests impartially considered. Finally, we should understand the condition that grounds moral status to be a capacity rather than a relational property. Relational properties only give us agent-relative reasons to consider the interests of moral status holders. However any true account of moral status should give moral agents agent-neutral reasons to consider the interests of moral status holders and establish their intrinsic value independent of any agent-relative value.

Chapter 2

Why do we need an Agency Account?



In the previous chapter I outlined in detail the notion of moral status and its function. I argued that moral status cannot plausibly permit degrees, that similar interests of moral status holders ought to be equally considered, that we ought to be able to determine the central and basic *pro tanto* treatment owed to moral status holders in light of the capacities that they share and that moral status must be grounded in some capacity rather than some relational property. In this chapter I will evaluate some specific accounts of moral status. One should note that I will not discuss any moral status accounts that in principle reject the notion of animals possessing moral status. This is because, as I argued in the Introduction, sentient animals intuitively have at least some claim to moral status. Thus, any account which explicitly denies that animals possess moral status, I take it, must be inadequate and so is not worth discussing here.

I will consider three popular accounts of moral status: *the Sentience Account*, *the Subject-of-a-Life Account* and *the Kantian Account*. These accounts have laid strong foundations for debates in animal ethics, and have each played a significant role in bringing awareness to different facets of the plight of animals, both within the philosophical literature, and in the world at large. I find them all intuitive in some respect and I think each possess at least a kernel of truth. However despite their virtues, I will argue that none of them are quite right and a better account of moral status could be constructed. I will then build upon my criticisms of these accounts to determine what attributes an ideal account of moral status should possess. Finally I will elucidate *the Agency Account of Moral Status* and argue that this account is closer to an ideal account of moral status than the other accounts considered. This is the case I will suggest, because the Agency Account manages to adequately justify the central *pro tanto* treatment we owe to moral status holders in a compelling and straightforward way, it provides a plausible ground for the moral status of humans *and* animals and it can explain the intuitive appeal of the other accounts discussed here thus making it analytically reductive.

I. The Sentience Account

The Sentience Account is the most widely accepted account of moral status by those who recognise that we owe animals moral consideration. It goes back to the ancients, though its most well-known proponent with regard to animals is Peter Singer.¹ The Sentience Account is often assumed to be linked to consequentialism because of Singer's (and other proponents')² commitment to Utilitarianism. However there are plenty of defenders of the Sentience Account who have more or less deontological persuasions.³ So although I will predominantly discuss Singer's interpretation of the account, I will criticise only his account of moral status. Whether his utilitarianism is problematic in anyway is not of interest to me here.

The Sentience Account states that sentience grounds the possession of moral status. Often the account is summed up as the position that individuals who can feel pleasure and pain have moral status, but this is an over-simplified understanding.⁴ More precisely the account can be interpreted as claiming that being able to feel pleasant and unpleasant mental states grants one moral status. Thus being able to feel only mild disappointment as well as relief, whilst perhaps not pleasurable or painful states, are still unpleasant and pleasant states respectively. Therefore a being capable of possessing only these states is nonetheless sentient.⁵

Sentience theorists argue for this account by first arguing as I did in chapter 1: one must possess interests in order to possess moral status, since if an individual lacks interests then there is nothing for moral agents to take into account when considering how they should treat such an individual. Further, they argue, without being able to experience pleasant and unpleasant mental states, one cannot possess interests.⁶ Nothing can matter to an individual that lacks sentience since they cannot be negatively or positively affected by anything. Thus

¹ Peter Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals* (London: Granada Publishing Ltd, 1977).

² Notably Jeremy Bentham.

³ Cochrane, Donaldson and Kymlicka and Francione all take sentient beings to possess rights: Alasdair Cochrane, *Animal Rights without Liberation: Applied Ethics and Human Obligations* (Columbia University Press, 2012); Sue Donaldson and Will Kymlicka, *Zoopolis: A Political Theory of Animal Rights* (Oxford: Oxford University Press, 2013); Gary L Francione, *Animals as Persons: Essays on the Abolition of Animal Exploitation* (New York: Columbia University Press, 2008).

⁴ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 27; Jeremy Bentham, *An Introduction to the Principles of Morals and Legislation* (White Dog Publishing, 2010), chap. xvii.

⁵ I will discuss how we ought to understand the notion of sentience in detail in the next chapter, however for now, my points here should be sufficient for the purpose at hand.

⁶ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 27; Cochrane, *Animal Rights Without Liberation*, 24.

since possessing sentience is required in order to possess interests, sentience is a necessary condition for moral status.

Sentience theorists then show that sentience is sufficient for moral status by appealing to intuitions about the badness of pain. They suggest that it is plausible that pain is bad for the individual experiencing it, regardless of other facts about that individual. Bentham stated this clearly and forcefully enough that he is still quoted in most works on animal ethics to this day: '...the question is not, *Can they reason?* nor, *Can they talk?* but, *Can they suffer?* Why should the law refuse its protection to any sensitive being?'⁷ Further, if a being is capable of suffering then this suffering seems morally significant, thus it must be *prima facie* wrong for moral agents to inflict pain upon them without justification. Even if one denies that the suffering of some individuals is not as morally significant as that of others', it is implausible to deny any moral significance at all to the suffering of sentient beings. Therefore the possession of sentience is sufficient for moral status.

It is also worth noting that many theorists go further than this and argue that there is no principled reason to think that the suffering of one individual is morally worse than the suffering of another, where they are suffering in exactly the same way. That is to say such theorists have a commitment to the *Equal Consideration of Interests Principle*.⁸ Thus if I inflict *x* degree of pain upon a horse, this is no better or worse than inflicting the same degree of pain upon a human, other things being equal.⁹ Of course other things rarely are equal and there is considerable debate about how one's other capacities may affect the experiential and moral badness of experiencing pain. For instance one's capacity for anticipation may make painful experiences less unpleasant than they would be otherwise, because one knows that the pain will pass. However the same capacity might make painful experiences *more* unpleasant since the pain is augmented by the feeling of dread that the pain will continue and/or reoccur.¹⁰

The main virtues of the Sentience Account are its simplicity and intuitiveness. Because it is so intuitive and simple to grasp, it has significant argumentative force. However, as

⁷ Bentham, *An Introduction to the Principles of Morals and Legislation*, chap. xvii.

⁸ A principle closely related to the Principle of Equality discussed in the Introduction. Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1993), 21; Cochrane, *Animal Rights Without Liberation*, 48-49; Gary L. Francione, "Animals - Property of Persons?," in *Animals as Persons: Essays on the Abolition of Animal Exploitation* (New York; Chichester: Columbia University Press, 2009), 44-45.

⁹ Singer, *Animal Liberation: Towards an End to Man's Inhumanity to Animals*, 34.

¹⁰ *Ibid.*, 35; Singer, *Practical Ethics*, 60.

DeGrazia has noted, this simplicity comes at the expense of philosophical accuracy.¹¹ I will argue that despite the intuitive pull of the Sentience Account's central claim that the suffering of all beings matters, human moral status holders plausibly matter for reasons other than the fact they are sentient. Further, on the Sentience Account it is only experientially unpleasant mental states (or at best the risk of experientially unpleasant mental states) that is *pro tanto* bad for moral status holders. As such, I will argue, we cannot draw a convincing account of what is (*pro tanto*) centrally good and bad for moral status holders from the Sentience Account.

Firstly I should say I feel the intuitive pull of the claim that all sentience individuals matter and I think that this is correct. The Sentience Account has a plausible and intuitive extension. If one has interests and is capable of being affected by the actions of others, then one must surely matter to some extent. However, though it seems that sentient individuals possess moral status, it appears to me considerably less intuitive that sentient individuals possess moral status *because* they are sentient. Whilst many might find it convincing that non-human animals possess moral status in virtue of being sentient, this is not the reason we usually consider humans to possess moral status.

More often we characterise humans as individuals worthy of moral consideration because they have desires, plans, psychological lives, relationships, autonomy and more. It seems there is more to mattering than merely being capable of experiencing pleasant and unpleasant mental states. Whilst being sentient might form part of an explanation of what makes humans morally matter, it doesn't plausibly constitute an answer in itself.¹² However, this problem runs deeper than a mere suspicion that humans matter for other reasons. As I argued in the previous chapter, a good account of moral status should feed into the correct accounts of right and wrong action. It should do this by allowing us to draw out the central ways in which a moral status holder can be *pro tanto* harmed or benefited. Whilst the Sentience Account successfully explains why some treatment of moral status holders is *pro tanto* wrong, it doesn't account for some plausibly central ways in which moral status holders can be *pro tanto* wronged. Of course we need not be able to use a moral status account to determine what

¹¹ David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (Cambridge; New York: Cambridge University Press, 1996), 3.

¹² I should make clear here that Singer endorses a hierarchical account of moral status and believes humans have a higher moral status than merely sentient creatures because of their self-consciousness. However many defenders of the Sentience Account do not hold such a position and presumably take sentience to be the sole condition which grounds human moral status.

is *pro tanto* wrong with *any* action. However a moral status account needs to explain the wrong-making features of the central ways in which one can be wronged.

Let's look at the case when sentience does feature in the argument for the *pro tanto* wrongness of some action. On the Sentience Account, sentience plays an obvious and important role in explaining why we shouldn't physically assault moral status holders. Moral status holders are sentient beings. Sentient beings alone are capable of experiencing pain (an intrinsically unpleasant experience most suggest) and physically assaulting a sentient individual (in almost all cases) causes them to experience pain. Therefore physically assaulting moral status holders is in itself bad for them and therefore it is *pro tanto* wrong.¹³ Thus the property that plays the role of necessary and sufficient condition for moral status on this account, also plays a foundational role in determining the wrongness of assaulting moral status holders. That moral status holders are sentient explains why inflicting pain upon them is *pro tanto* bad for moral status and therefore *pro tanto* wrong.¹⁴

This is exactly the kind of argument that a moral status account should be able to provide to account for the wrong-making features of the central ways in which moral status holders can be wronged. However, the Sentience Account fails to provide such arguments with respect to other important wrongs. For instance, the Sentience Account does not establish a duty not to restrict the liberty of moral status holders. Plausibly humans, as paradigmatic moral status holders, are *pro tanto* wronged when they have their liberty restricted. Such restrictions can be some of the worst wrongs we can suffer in fact. As such it might be considered a central good for us that we possess liberty. This being the case, the Sentience Account should give us the resources to construct a clear explanation for the wrongness of (at least some) moral status holders having their liberty restricted. Furthermore, sentience should feature prominently in this explanation. However no such explanation is available. It is not clear how being sentient could explain why being deprived of liberty would *pro tanto* wrong one.¹⁵

¹³ Singer, *Practical Ethics*, 59.

¹⁴ A similar argument could be derived from the Sentience Account to show that we should promote pleasure in moral status holders.

¹⁵ One might think that the harm of death would be a better example here since Singer himself accepts that sentience creatures are replaceable and not harmed by death (in principle). However I think Singer is wrong about this and that in fact merely being sentient *is* sufficient for death to *pro tanto* harm and wrong one, for reasons I will give in Chapter 5. Singer, *Practical Ethics*, 121.

One option to attempt to circumvent this problem would be to adopt a two-tier system of moral status in which sentience grants one basic moral status, and some other condition grants one higher moral status, or what is often called 'full moral status'. Thus one might maintain that whilst merely sentient beings are not wronged by the restriction of liberty, individuals such as humans who qualify as full moral status holders, could be wronged in such cases in virtue of possessing some other property. Thus the Sentience Account *can* explain the central ways in which actions can be bad for and wrong *basic* moral status holders since the only central way in which they can be wronged is through suffering. A restriction of liberty does not *pro tanto* wrong basic moral status holders but only full moral status holders.

Further, one might suggest that autonomy grounds full moral status and the key wrong-making feature of restricting the liberty of full moral status holders is that it undermines their autonomy in some way. Whilst this would adequately explain the wrongness of these liberty restrictions, a hierarchical account of moral status is problematic for other reasons. As I argued in the previous chapter, I think hierarchical systems are untenable. There is no clear understanding of what it could mean to have higher or lower moral status that is justifiable. Moral status as I understand it is something that all moral status holders have equally. As such, this kind of two tier system is an unsatisfactory solution to the problem.

A related approach proponents of the Sentience Account could adopt is to argue for the equal moral status of all moral status holders on the grounds of sentience, but then rely on some other property (or properties) to explain the wrongness of various centrally wrong actions. For instance, one may argue that moral status holders have desires and if we restrict their liberty, we frustrate the pursuit of these desires. Thus restricting the liberty of moral status holders is wrong. However, this argument fails to show what being a moral status holder has to do with the restriction of one's liberty being wrong. Could an individual who has desires but lacks sentience (assuming such an individual could exist) be wronged by having their liberty restricted? Plausibly, if desire-possession grounds the wrongness of restricting one's liberty, it seems arbitrary to claim that one would need to be sentient in order to be wronged by having one's liberty restricted. Thus those that lack moral status must be able to be deprived of liberty. However that those lacking moral status can be wronged is incoherent. To accept that individuals that lack moral status can be wronged is a direct contradiction of what it means to possess moral status. As such, it cannot be the case that anyone but moral status holders could be wronged by being deprived of their liberty.

One might think that one must be sentient in order to possess moral status and that in order to be harmed by a restriction of liberty one must possess both desires and sentience. However if this is the case it becomes unclear why sentience should be taken as the necessary and sufficient condition for moral status when other properties such as desires play a significant role in underwriting the central duties owed to moral status holders. If another property, such as desire possession, plays a role in establishing the wrongness of some of the central ways in which one can be wronged, it seems this property should play some role in grounding moral status too. Sentience and this other property could be jointly necessary and sufficient conditions for moral status for instance.

Whilst hybridising the Sentience Account with some other account could seemingly resolve this problem, this solution seems non-ideal. A philosophically simpler and more straightforward solution would be to argue that a more fundamental property, the possession of which entails the possession of these other properties, is the condition for moral status. Or one could construct an account of moral status around a different property which can offer explanations for the wrongness of these different centrally wrong actions.

Hence, the Sentience Account provides us with a clear account of moral status that provides a strong explanation of the *pro tanto* wrongness of causing unpleasant experiences. However without further argument the account cannot show what is wrong with some other important ways in which one can be harmed. Whilst one could hold that individuals matter because they possess sentience and that it is wrong to treat them in certain ways (such as restricting their liberty) in virtue of further facts about them as individuals, a more ideal account of moral status would explain the wrongness of such actions by reference to the property or properties necessary and sufficient for moral status itself.

II. The Subject-of-a-Life Account

The Subject-of-a-Life Account was developed and defended by Tom Regan.¹⁶ Whilst this account is not widely endorsed, it has been widely influential, and for this reason is worth discussing. Regan argues that the criteria for the possession of moral status is being the subject-of-a-life. This means having desires and beliefs, perception, memory, a sense of future (including one's own future), an emotional life, sentience, interests (both subjective and

¹⁶ Tom Regan, *The Case for Animal Rights* (Berkeley: University of California Press, 1983).

objective interests), the ability to initiate action, a psychological identity over time and well-being.¹⁷

Regan states that one must have all these capacities in order to be a subject-of-a-life. If one lacks any one of these properties then one cannot be a subject-of-a-life. Furthermore, subject-of-a-life-ness does not permit degrees.¹⁸ Therefore one cannot be more or less of a subject-of-a-life. Even if one possesses all of these properties to a greatest degree one is just as much of a subject-of-a-life as any other individual that possesses these capacities. Importantly, unlike the other views discussed here, Regan's account does not give us a necessary and sufficient condition but only a sufficient condition for moral status.¹⁹ Hence on Regan's account, some individuals that are not subjects-of-a-life may possess moral status if they fulfil some other sufficient condition for moral status.

Regan argues for his account not by providing an argument for the moral significance of being a subject-of-a-life but by giving a conditional argument. Regan postulates the moral status of moral agents and reasons that many animals (adult mammals) are cognitively similar to moral agents in the ways listed above and can it seems be harmed in similar ways. As such, he concludes, if humans have moral status then adult mammals must possess moral status too. Further, since all (humans and non-human) adult mammals share the mental capacities that make one a 'subject-of-a-life', being a subject-of-a-life must be what underlies their moral status. Thus Regan concludes that being a subject-of-a-life must be sufficient for the possession of moral status.²⁰

The Subject-of-a-Life Account is an improvement over the Sentience Account in the respect that being a subject-of-a-life can explain more of the various ways in which moral status holders can be wronged. For instance, Regan believed that moral status holders could be wronged by having pain inflicted upon them, having their lives ended and having their liberty restricted, among other things. Being a subject-of-a-life directly and plausibly explains why we should promote pleasant experiences and refrain from inflicting unpleasant experiences. Being a subject-of-a-life entails being sentient and thus being capable of experiencing such pleasant and unpleasant states. Thus, Regan's account explains the wrongfulness of inflicting pain on moral status holders in much the same way as the Sentience

¹⁷ Ibid., 243.

¹⁸ Ibid., 244–45.

¹⁹ Ibid., 245–46.

²⁰ Ibid., 247–48.

Account. Unlike the Sentience Account however, the Subject-of-a-Life Account directly explains the wrongness of moral status holders having their liberty infringed (in some ways). Since subjects-of-a-life are capable of 'preference autonomy', roughly the ability to make preference choices, it is good for them to be free to make decisions. Thus it is wrong to interfere with the decisions of subjects-of-a-life.²¹ Regan's account also provides a more obvious explanation of the wrongness of ending the life of a moral status holder than the Sentience Account.

Regan's recognition of the value of what he calls 'preference autonomy' is particularly important. Preference autonomy as Regan defines it, means being able to choose between one's preferences and 'initiate action with a view to satisfying them'.²² This, I take it, is equivalent to what is commonly taken to be intentional agency. Preference autonomy is only one of the conditions for being a subject-of-a-life, yet it seems to do a significant part of the work in determining the kind of treatment moral status holders are owed. Not only is one's interest in being free to pursue one's preferences grounded in this, but possessing preference autonomy is necessary to possess any subjective (or as Regan calls them preference) interests.²³ Such interests are of vital importance when determining how to treat an individual. Regan himself sums up the value of preference autonomy aptly when he states that it 'provides the starting point for thinking about the interests of animals'.²⁴

However Regan's account does face some problems. Firstly it seems that Regan's distinction between those that are, and those that are not subjects-of-a-life, is counter-intuitive. Regan claims that unless one possesses all of the properties he lists, one isn't a subject-of-a-life *at all*. Thus there is a hard boundary between being a subject-of-a-life and not. It seems implausible to suggest that a slight difference in one's capacities amounts to the difference between possessing the strong duties that Regan argues are tied to being a subject-of-a-life, and lacking them (or at least lacking many of them). This hard boundary is especially implausible in cases where the only difference between two individuals is a sense of one's own future or the capacity for memory, for instance. Why should such a minor difference have such a significant effect? It seems arbitrary to grant moral status to those who fulfil the conditions but deny it to those who only just fail to meet them.

²¹ Ibid., 85.

²² Ibid.

²³ Ibid., 87.

²⁴ Ibid., 116.

One might attempt to solve this problem by weakening the conditions for being a subject-of-a-life. One might understand a subject-of-a-life to be an individual who possesses a sufficient number of the qualities that Regan lists, rather than necessarily possessing the complete set. Thus an individual that possessed all the properties save for memory, would still qualify as a subject-of-a-life. Whereas one who lacked memory, a sense of future and the ability to initiate action may not be a subject-of-a-life. One could even argue for an understanding of being a subject-of-a-life according to which there are certain central properties that one must possess, such as sentience and/or preference autonomy and other more periphery properties of which one need only possess a sufficient number.

One might point out that such a modification to the conditions for being a subject-of-a-life isn't necessary since being a subject-of-a-life is merely a sufficient condition for possessing moral status on Regan's account, not a necessary condition. However this does not plausibly allow us to account for duties towards individuals who do not qualify as subjects-of-a-life. What's more it seems to give rise to other problems. If some other sufficient condition can plausibly account for individuals that *just* fail to meet the conditions for being a subject-of-a-life, it seems that this condition will have to grant one many of the same entitlements as being a subject-of-a-life. If it fails to do this then the objection of arbitrariness will remain; why ought some individual who only just fails to meet the conditions for being a subject-of-a-life, miss out on the same benefits as a subject-of-a-life when there is little difference between their capacities?

However if this second sufficient condition does secure roughly the same entitlements as being a subject-of-a-life, then it appears to make the condition of being a subject-of-a-life redundant. Plausibly all subjects-of-a-life will fulfil this second sufficient condition for moral status since this condition is supposed to pick up those individuals who just fail to meet the conditions for being a subject-of-a-life; individuals that will have almost all of the same capacities as subjects-of-a-life. Therefore the second sufficient condition provides roughly the same entitlements as being a subject-of-a-life but has a more wider extension, granting moral status to individuals that the subject-of-a-life condition does not. Thus this second sufficient condition would be more efficient than the subject-of-a-life condition. What's more, it would be more intuitive as it grants moral status to more beings who plausibly matter. This being the case, it seems that we could do away with the subject-of-a-life condition entirely. By doing so we would have a more cohesive and theoretically simpler account of moral status than one

on which being the subject-of-a-life or possessing this second sufficient condition grants one moral status.

However even if one recognises the force of this argument, and weakens the conditions for being a subject-of-a-life as the reply above suggests in order to accommodate borderline cases, the Subject-of-a-Life Account still suffers from a problem. Merely sentient individuals have a claim to moral status. Thus a plausible account of moral status will have to grant them moral status or explain away their claim to moral status. Weakening the conditions for being a subject-of-a-life, may help to accommodate borderline case but it will not seemingly grant merely sentient beings moral status on this account. Further, arguing that sentience is a sufficient condition for moral status would seemingly give rise the redundancy problem. Since all subjects-of-a-life possess sentience, why does it matter whether one is a subject-of-a-life if merely being sentient is sufficient to possess moral status?²⁵ One cannot argue that being a subject-of-a-life grants one a higher moral status either since, as I have already argued, moral status does not permit degrees. Thus the Subject-of-a-Life Account cannot seem to accommodate the merely sentient beings.

Therefore whilst the Subject-of-a-life Account is intuitive in its explanation of why individuals matter and can explain the wrongness of some centrally wrong actions, it is problematic. This account suggests that two individuals with little and apparently rather insignificant differences can be owed significantly different treatment. This seems implausible. Further this account fails to grant moral status to merely sentient beings, who must plausibly be owed moral status.

III. The Kantian Account

Kant claimed that the possession of 'rationality' is necessary and sufficient for moral status. 'Rationality' in the relevant sense here means capable of acting, reflecting upon and revising one's reasons. Thus not only does one need to be capable of acting, but also capable of consciously reflecting and deliberating over one's action. Kant believed rational agents to be valuable because they alone are capable of acting on pure reason. That is to say they are

²⁵ Of course whether one is a subject-of-a-life will make a differences to the interests one will possess. However if we are interested in who possesses moral status, it seems odd to state that those who possess moral status are those who are either subjects-of-a-life or sentient since all subjects-of-a-life by definition are sentient.

capable of forming goals or aims, and acting in order to achieve these goals or aims, free from causal determination by their mere instincts and desires. In so doing they can decide how to live their lives. They possess autonomy and this makes them valuable.²⁶ As such, we should respect them by not interfering with them or treating them in ways which undermine the choices they make.²⁷

Kant's account of moral status directly and foundationally informs us how moral status holders ought to be treated. In virtue of being autonomous Kant argues, moral status holders ought to be treated never merely as a means, but always also as *ends-in-themselves*. By this Kant means that it is wrong to use moral status holders simply as tools without regard for them as individuals with their own plans and ambitions. From this primary duty, and the fact that moral status holders are autonomous, Kant derives all other duties, including a duty of non-interference and a duty not to kill. Thus the criterion for moral status plays a significant role in explaining the treatment we owe to moral status holders for Kant.

Despite close connection between Kant's account of moral status and his account of wrongness, his moral status account has a significant shortcoming: it excludes all animals and many humans.²⁸ Since humans and many animals have interests that are morally significant it seems that Kant's account will not pass as a satisfactory account of moral status. However, modern Kantians have argued that some animals and all humans can be said to possess moral status on a Kantian account if we reinterpret some of Kant's claims. Such arguments rest on the claim that animals possess a related but different kind of rationality to the notion Kant focused on. Christine Korsgaard has, I think, been the most successful in this endeavour and as such it is her account that I will consider here.

Unlike some other modern Kantians, Korsgaard's account does not involve much reinterpretation of Kant. She gives one of Kant's own arguments for the condition of moral status and shows that (despite Kant's claims) it is applicable to animals just as it is to humans.

²⁶ Immanuel Kant, *The Moral Law: Groundwork of the Metaphysics of Morals*, trans. Paton, H. J (London: Routledge, 1989), 96–97.

²⁷ *Ibid.*

²⁸ That is not to say that Kant thought we could permissibly treat animals however we like. Kant believed that we are obligated to treat animals in certain ways but not because they matter in themselves. We have such obligations regarding animals because treating animals badly cultivates cruel habits in rational agents and thus would lead to us treating rational agents (who do matter in themselves) badly. As such we have duties involving animals but not duties *toward* animals. Thus animals themselves do not have moral status although our treatment of them is restricted in certain ways. Immanuel Kant, *The Metaphysics of Morals*, trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1991), 238.

She argues that Kant's claim that rational beings have moral status should be understood as follows. Rational beings decide to pursue objects because they believe them to be good. No objects are good in themselves, yet still our chosen objects appear to be valuable. It must be that it is in the choosing of some object to pursue that we make it valuable; we make it 'a good'.²⁹ Further, if we are able to make our chosen ends valuable, it must be because we ourselves have some intrinsic value in light of this ability. Therefore rational beings possess moral status and are ends-in-themselves. Understanding Kant's claim in this way, Korsgaard argues, means that many animals as well as humans will qualify as moral status holders and thus ought to be treated as ends-in-themselves.³⁰

Korsgaard argues that while conscious non-human animals cannot evaluate their own actions and reasons like humans, they do choose to pursue objects. For instance, a dog can choose to lay in the sun or run on the lawn. Further, if animals can make choices over which objects to pursue, the objects of their pursuits can be said to matter to them.³¹ They would clearly be distressed and suffer if they could not pursue their chosen course of action. As such, they, like humans, through choosing some object, imbue this object with value. In other words they give themselves goods or ends.³²

Importantly, Korsgaard argues that goods don't have to be created or developed by the individual in order to matter to them. The content of one's goods may be 'set by nature'. What makes these objects 'goods' is that one endorses or desires them.³³ Thus being able to run for instance, may be a good for a horse, even though most horses may naturally desire to do this anyway. The fact that the horse chooses to pursue running is an endorsement of this object and makes it a good.³⁴ In this way Korsgaard argues, many animals can be said to set ends for themselves. As such, conscious animals as beings capable of imbuing objects with value, can be said to possess intrinsic value themselves, and like humans, be moral status

²⁹ Christine M. Korsgaard, "Kantian Ethics, Animals, and the Law," *Oxford Journal of Legal Studies* 33, no. 4 (2013): 11.

³⁰ Donald Regan provides a clear and concise account of Korsgaard's argument. Donald H. Regan, "The Value of Rational Nature," *Ethics* 112, no. 2 (2002): 271.

³¹ Christine M. Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals," *The Tanner Lectures on Human Values* 25 (2004): 103.

³² *Ibid.*, 104.

³³ Christine M. Korsgaard, *Christine Korsgaard on the Status of Animals*, Philosophy Bites, n.d., <http://philosophybites.com/2015/02/christine-korsgaard-on-the-status-of-animals.html>.

³⁴ *Ibid.*

holders who ought to be treated as ends-in-themselves. It is not merely autonomy that grants us moral status but the capacity to pursue our own goods.³⁵

Korsgaard's Account may appear to be almost exactly what I am looking for here. It appears to take something at least tantamount to agency to be the condition for moral status. Further, Korsgaard believes that it defends the moral status of roughly all sentient animals and the condition for moral status appears to play a fundamental role in grounding the duties owed to moral status holders. Unfortunately, despite its apparent virtues, this account has some problems.

Firstly, the account appears to grant moral status to non-conscious entities. Given that moral status entitles one to be treated as an end-in-oneself on this account, this seems implausibly strong. At first it appears that Korsgaard is only extending moral status from some humans to all humans and conscious animals. Korsgaard states that one has moral status if one has goods which one chooses to pursue. For something to be a good for one, that thing must matter to one in some sense, and a sufficient condition for something mattering to one, is one possessing a naturally occurring desire towards that object. However Korsgaard has stated that some groups of individuals such as a colony of bees, or termites, may possess goods and therefore moral status.³⁶

If such groups are pursuers of goods, then they must have things that matter to them. However if this is the case then Korsgaard cannot be using the term 'mattering' in the everyday sense to mean that one cares about or wants some object. 'Mattering' in the everyday sense of the term, clearly requires one to possess subjective consciousness, something bees and termites lack as individuals and as collectives. Thus, Korsgaard must be using 'mattering' in some other sense. Indeed Korsgaard's remark in the last pages of her Tanner Lecture reinforce this claim. She states that plants are capable of pursuing goods and admits that it seems that there is 'probably no distinct line' between stimulus and response mechanisms and 'full-blown consciousness'.³⁷ Thus whilst possessing desires may be sufficient, it does not seem to be necessary for the kind of mattering or pursuing of good that Korsgaard has in mind.

Thus it appears that non-conscious beings such as insects and plants (as well as collectives of these organisms) possess moral status on Korsgaard's account. Korsgaard does

³⁵ Korsgaard, "Fellow Creatures," 102.

³⁶ In personal correspondence

³⁷ Korsgaard, "Fellow Creatures," 106-7.

temper this claim by suggesting that moral status may come in degrees but for reasons I have discussed in the previous section and the last chapter however, this will not provide a satisfactory solution.³⁸ Putting aside the problem of degrees of moral status and the problems of collectives as moral status holders however, the claim that plants have moral status presents us with fresh problems. Specifically; if moral status holders are ends-in-themselves, ought plants to be treated as ends-in-themselves? It is not clear what exactly this means, and any interpretation of these terms applied to plants would be a radical departure from the way in which we use the term when referring to humans or animals. Interestingly this same problem does not arise in the case of considering animals to be ends-in-themselves. Humans and animals can be treated as ends-in-themselves in the very same sense.

Furthermore, even if we could get a good grip on what it means to treat a non-conscious organism as an end-in-themselves, claiming that plants ought to be treated as ends-in-themselves would appear to trivialise the claim that humans and animals ought to be treated as ends-in-themselves. In order to avoid this conclusion Korsgaard might suggest that animals (including humans) are ends-in-themselves and non-conscious entities are not. However if she was to take this line of argument it is not clear what grounds she would, or could, use for this distinction, considering that the criterion for moral status on her account, she suggests, is manifested in both conscious and non-conscious organisms. Further, even if there were some principled distinction to be drawn, why ought we to believe that non-conscious entities like plants possess moral status at all? As argued in Chapter 1, they don't seem to have any interests for us to take into account. The kind of 'goods' that they have don't matter to them. So it seems that Korsgaard's account has an implausible extension considering the protections/duties that she takes to come along with moral status.

Secondly, even setting aside the thorny issue of non-conscious entities as moral status holders, it doesn't seem that Korsgaard's criterion for moral status can secure the duties she believes moral status holders are due. Specifically, it is mysterious why we ought to treat non-autonomous beings as ends-in-themselves. Korsgaard argues that all individuals capable of possessing and pursuing goods have moral status and that this entails that they ought to be treated as ends-in-themselves. This requires moral agents to never treat a moral status holder as a mere means. She inherits this idea from Kant. However when Kant argued that rational beings are ends-in-themselves, and cannot be treated as mere means, he shows us that there

³⁸ Ibid., 106 (footnote 69).

is a clear connection between the nature of rational beings and their entitlement to be treated as ends-in-themselves.

Kant takes all rational beings to be autonomous (or more precisely capable of acting autonomously) as such it is good for them to live a life according to their own rules, and determined entirely by themselves. By failing to treat such individuals as ends-in-themselves we stop them from being able to make some autonomous choices and deprive them of some of their autonomy, or as Kant would say, we fail to respect their autonomy. Autonomy on Kant's account plays the key role in explaining the wrongness of treating moral status holders merely as means. However if we revise the account of moral status so that autonomy is no longer the condition that grounds moral status, as Korsgaard does, then the explanation of the wrongness of treating moral status holders as mere means disappears. Thus Korsgaard must supply another explanation for why such treatment of moral status holders is *pro tanto* wrong.

Korsgaard argues that that one must merely be able to pursue goods in order to have moral status. Whilst it seems to follow from this that moral status holders ought not to have the pursuit of their goods obstructed, I don't think it follows (merely from this claim) that such individuals must be treated as ends-in-themselves. It seems that one could allow some individual to pursue their goods without necessarily treating them as an end-in-themselves.

For instance, it would not be an obstruction to the pursuit of an individual's goods to drug them whilst they are asleep so that they develop a strong desire to pursue some new good. However to do so would surely be to treat them as a mere means and not as an end-in-themselves. It appears that moral status holders on Korsgaard's account (that lack autonomy) could be treated in this way. This may not be counter-intuitive. There plausibly is a difference between the treatment we owe to autonomous and non-autonomous moral status holders (as I will argue in the later chapters of this thesis). However, as it stands the criterion for moral status in this case does not ground the *pro tanto* treatment that the account claims moral status holders are due. In other words, if Korsgaard believes that all moral status holders ought to be treated as ends-in-themselves, she needs to provide a new argument for this claim, she cannot simply rely on Kant's argument to do the work for her, since it cannot explain why non-autonomous individuals ought to be treated as ends-in-themselves.

Finally, and most problematic, is the commitment that Korsgaard's account requires to a counterintuitive theory of value. For Korsgaard (and Kantians generally) nothing has intrinsic value other than moral status holders. Hence, the value of anything else in the world

is wholly extrinsic and contingent upon the value of moral status holders. Despite the widespread appeal of many parts of Kant's moral theory, this account of value is a hard pill to swallow. What is more, this account has a difficult time explaining how we choose between available options. According to Korsgaard's account, through selecting an object to pursue, we imbue it with value and make it a good for us. However, since one's available options don't have any value independent of us, we don't seem to have any standards to help us when deciding between two available projects. So it seems we simply have to arbitrarily pick one option and this option becomes valuable because we chose it. Or we simply pick the option that best satisfies our desires. However as Donald Regan rightly claims, most people don't think that their projects are good only because they have chosen them and that any other project would be equally as good had they chosen one of them instead.³⁹

More plausibly it seems we choose something because (we think) it is valuable, not the other way around. This is evidenced by the fact that when choosing between options, we usually believe ourselves to be considering the relative value of the options. If I consider whether I should devote myself to becoming an artist for instance or accept a job doing menial work, it seems to me that being an artist has more value than the menial job *independently* of my decision. Of course I could simply be mistaken. Though if this is the case some argument needs to be given to explain away this phenomenon and it is not clear what such an argument would look like.

More generally it seems intuitive that some objects have value independent of humans and further that some objects lack value despite one choosing to pursue them. The popular example of a life that is devoted to grass-counting seems to be as good an example as any. It seems simply false that an individual who chooses to pursue the project of endlessly counting blades of grass, pursues a valuable end. That they choose this end does nothing to convince us that this is a valuable pursuit in itself. Grass-counting simply seems like a life lacking in value. Of course that isn't to say that one shouldn't be free to pursue such a life if one wants to, but merely that in choosing such a life one does not plausibly make it valuable as Kantians are committed to maintaining.⁴⁰

Overall then the Korsgaardian account, fails for a variety of reasons. Firstly the account has an implausible extension including non-conscious entities such as plants and

³⁹ Regan, "The Value of Rational Nature," 274.

⁴⁰ *Ibid.*, 272.

insects as moral status holders. Secondly, it is unclear how the criteria for possessing moral status is supposed to ground the duty to treat non-autonomous beings as ends-in-themselves. Finally, it requires one to accept a counter-intuitive theory of value. It simply doesn't seem plausible that only moral status holders possess intrinsic value and all other objects or pursuits lack any such value. As such, having and being able to pursue goods (in the sense Korsgaard understands these terms) doesn't seem like a plausible sufficient condition for possessing moral status.

IV. Constructing an Ideal Account of Moral Status

Through my critique of these three accounts of moral status I have highlighted some problems with current attempts to justify the moral status of humans and animals. By taking note of these problems, I can build a list of desiderata for an ideal account of moral status. I can then use this to ensure that my account of moral status will succeed where others have failed.

Firstly on an ideal account of moral status we should be able to draw out the central and basic ways in which it is *pro tanto* wrong and *pro tanto* right to treat moral status holders. Furthermore the *pro tanto* wrongness and rightness of this treatment should be explained by the property that grounds moral status. Secondly, the ideal account should (if at all possible) have one necessary and sufficient condition for moral status which can ground the moral status of all those that plausibly matter. This is the theoretically simplest kind of moral status account and avoids the redundancy problem that the Subject-of-a-Life Account faces.⁴¹ Thirdly the account must provide a plausible explanation for the value of the property which grants moral status and explain why those lacking the property cannot possess moral status. As such the account should grant moral status to all sentient beings or provide a plausible explanation for why the interests of merely sentient individuals do not matter (since it has been established that they plausibly do).

It may not be possible to construct an account that manages to fulfil all of these criteria at the same time. However I will argue that an account that takes agency to ground moral status will come closer to realising these ideal conditions than any of the previously discussed

⁴¹ One should note that an account of moral status could have more than one sufficient condition, and still avoid the redundancy problem, if the groups of individuals that fulfil these conditions do not significantly overlap.

moral status accounts. As such it seems that we ought to adopt the Agency Account of Moral Status over its competitors.

V. The Agency Account

The Agency Account contends that the possession of agency grounds moral status.⁴² More specifically being capable of intentional action is the necessary and sufficient condition for possessing moral status. Importantly, I will argue through Chapters 3 and 4 that intentional action is significantly less cognitively complex than is often assumed, and that in fact all sentient individuals are necessarily capable of intentional action. Thus I will argue that all sentient individuals are moral status holders on the Agency Account since they are all necessarily agents.

A full argument for agency as a necessary and sufficient condition for moral status will occupy Chapters 5 and 6 but I will here briefly sketch out the overall argument. As I argued in Chapter 1, one must have interests in order to possess moral status. In this respect I agree with all other accounts of moral status discussed so far. However, I suggest that the only beings that possess interests are agents. This is because possessing interests requires the possession of desires (or some functionally similar motivating states).⁴³ One needs desires in order to find states of affairs pleasant or unpleasant and so one needs desires in order for things to matter to oneself. This does not mean that all goods are experiential, or even that all goods are agent-affecting, but merely that being capable of having desires, or some functionally equivalent mental state, is necessary for things to be good for one. If one possesses desires, then one is also capable of possessing beliefs. Finally, if one possesses desires and beliefs then one is capable of acting for a motivating reason and *this* is sufficient for intentional action. Whilst one may act intentionally without acting for a reason, it is uncontroversial that any action done for a (motivating) reason is an action done intentionally.

⁴² Since commencing the writing of this thesis two other individuals have independently argued in favour of agency accounts of moral status: Jeff Sebo, "Agency and Moral Status," *Journal of Moral Philosophy* 14, no. 1 (2017): 1–22; Thomas, *Animal Ethics and the Autonomous Animal Self*. However both accounts are underdeveloped. Neither plausibly establish the treatment owed to moral status holders in light of their capacity for agency.

⁴³ If some beliefs are motivating states, such as the belief one arrives at after a moral judgement, then possessing interests may merely require the possession of motivating beliefs.

Thus any being that possesses interests must be an agent and therefore being an agent is necessary for the possession of moral status.

Not only is agency necessary for moral status but it is also *sufficient* for moral status. Thus, one has moral status if and only if one is an agent. Agents necessarily have interests and thus only agents can be harmed, in other words, only agents can have higher or lower levels of well-being. Unless we have good reason to think that this harm is morally insignificant (and it doesn't seem like we do) agency in the minimal sense I articulate here must be sufficient for moral status. Further, agency is sufficient not just for mere mattering but for full moral status. I will argue that agents, merely in virtue of being agents, possess the interests that we consider to be central to possessing full moral status. Agents possess interests in being free from pain, continuing to exist and having liberty. As such, since these three interests are central to our notion of a full moral status holder, if individuals possess them it is likely because they are full moral status holders themselves. Therefore agency is a sufficient condition for full moral status.⁴⁴

This account of moral status fulfils more of the desiderata of an ideal account of moral status than the previous accounts reviewed in this chapter. This makes it a more ideal account than the alternatives. The criterion of the Agency Account plays a foundational role in an account of central *pro tanto* wrongful treatment of moral status holders. It can plausibly account for the moral status of humans and animals using the same single criterion. Furthermore, the account is analytically reductive, explaining away the intuitive appeal of rival accounts. I'll defend these claims in more detail next.

i. The Agency Account Determines Treatment

On the Agency Account, agency plays a foundational role in explaining what treatment is owed to moral status holders. I will argue in Chapters 5 and 6, that it is *pro tanto* wrong to inflict pain, end the life or restrict the liberty of a moral status holder and further that the wrongness of these actions are grounded in the capacity for agency. Inflicting pain on a moral status holder is *pro tanto* wrong because through inflicting pain we frustrate agents' desires. This I will argue is bad for agents. Thus by inflicting pain upon on a moral status holder we frustrate their desires, and in so doing, undermine their agency. Death wrongs

⁴⁴ This argument will be developed over Chapters 5 and 6.

moral status holders on the Agency Account because it robs one of the future satisfaction of certain desires. Thus likewise, by ending the life of a moral status holder, one undermines their agency. The Agency Account also shows that moral status holders are *pro tanto* wronged by certain restrictions of their liberty. I will argue that there is value in being able to exercise one's agency to determine the course of one's own life. Thus by restricting moral status holders' liberty we interfere with agents' choices and undermine their agency.

On the Agency Account, moral status holders' interests in pain, continued existence and liberty can be explained by reference to agency. Thus, the necessary and sufficient condition for moral status plays a foundational role in explaining the central ways in which moral status holders can be *pro tanto* wronged, and what *pro tanto* duties we owe them.

ii. Human and Animal Moral Status

The Agency Account gives us a plausible explanation for why both non-human animals *and* humans matter. The fact that (normal adult) humans are autonomous is intuitively an important part of why we think we possess moral status and at least one of the reasons that others cannot permissibly end our lives or interfere with our decisions. Accounts such as the Sentience Account suffer because they cannot incorporate this intuition. As I will argue in Chapters 5 and 6 however, the Agency Account can give an alternative explanation for this intuition, and provide an alternative and intuitively appealing account of moral status for humans and animals, based upon this explanation. It does this by demonstrating that those duties we believe to be owed to humans in virtue of their autonomy, are actually owed to them in virtue of their agency. Hence, it is because humans are agents that it is wrong to treat them in particular ways. Thus it is agency not autonomy that should be recognised as the necessary and sufficient condition for moral status.

However, one should note that this account does not reject the moral value of autonomy. Autonomy is simply considered *less* valuable on this account than it is usually taken to be. The Agency Account recognises that possessors of autonomy have some additional interests, and as such they are owed certain treatment that non-autonomous agents are not. Further they can be wronged in ways that non-autonomous agents cannot. However autonomous and non-autonomous agents are owed the same kind of basic and central treatment. Further the central ways in which they can be *pro tanto* wronged are the same. Thus on the Agency Account, autonomy, like some relational properties such as being a citizen,

entitles one to certain additional treatment, and allows one to be wronged in additional ways. However autonomy itself does not ground moral status.

The Agency Account also has intuitive pull for those like myself to believe that the intellectual capacities and individuality of animals is underplayed and undervalued by moral philosophy. By recognising animals as agents, the Agency Account offers opposition to the intuition that there is a deep divide between humans and animals. It suggests that we see ourselves and animals as fundamentally agential beings, and thus the same kind of beings (although of course different in degree). Furthermore, the claim that humans and animals matter for the same reason is what we should expect and is a benefit of the Agency Account over other available accounts. Humans and animals possess evolutionary continuity and psychological continuity. It seems that because of this they should also possess value continuity; they should be valued for the same reason. It just seems too unlikely that two closely related and similar groups of beings possess two distinct morally significant properties and thus matter for different, independent reasons. That animals and humans matter for the same reason on this account is also a benefit because it means the account avoids the redundancy problem regarding multiple sufficient conditions for moral status and that the account is theoretically simple, in terms of necessary and sufficient conditions.

iii. The Agency Account is Analytically Reductive

Agency is not just an attribute that is a plausible ground for moral status, it is the condition which encompasses the intuitive appeal of the other main contenders: sentience, rationality/autonomy and subjecthood. Agency is the conceptual link between these capacities, and sheds light on why it is that we might think that these other capacities are morally significant and could ground moral status. In this sense it is helpfully reductive, showing that opposing theories of moral status, have a shared underlying basis.

On the Agency Account all sentient individuals are agents since possessing sentience necessarily entails possessing agency. This is because (as I will argue in Chapters 3 and 4) sentience requires the possession of beliefs and certain intrinsic desires (or desire and belief-like states); psychological components that are sufficient for agency. Thus there is a clear connection between the value of sentience and the value of agency.

Not only does agency share a close tie to sentience but it also shares one with being a subject-of-a-life. One must have subjecthood in order to be an agent. One simply cannot act if one entirely lacks a self. Agency in the sense that is relevant to my account here, requires phenomenal consciousness and the possession of desires and beliefs (or functionally similar states). Selves are the only kind of thing that can possess these and use them to initiate action. Without some 'I' that acts, there is no *action*, only behaviour. Regan himself notes that desires and beliefs, as well as the ability to initiate action, are intrinsic features of a subject-of-a-life. Additionally many of the other features of a subject-of-a-life that Regan lists, are found (at least in some degree) in all agents of which we have knowledge; perception, memory, a sense of future, including their own future; an emotional life together with feelings of pleasure and pain; preference and welfare interests; the ability to initiate action in pursuit of desires and goals; a psychophysical identity over time and an individual welfare in the sense that one's experiential life fares well or ill for them.

In fact, Regan states that 'It remains possible that animals that are conscious but not capable of acting intentionally might be viewed as having inherent value'.⁴⁵ This suggests that Regan took subjects-of-lives to be the only individuals capable of intentional action and thus that he would accept that all agents are moral status holders. Further it is clear that agency had value to Regan as a central capacity of moral status holders.

Agency is closely connected with autonomy. The two notions are so closely related that autonomy is often mistakenly taken to be synonymous with agency. Frequently when we believe something to be wrong because it undermines autonomy, I will argue that it is in fact wrong because it undermines agency. Furthermore, the reason that agents and autonomous individuals possess many of the interests they do, I will suggest, is because these interests stem from their shared capacity for agency.⁴⁶

All agents also have goods to pursue in Korsgaard's sense and it is the case that all individuals capable of pursuing goods are agents on this account. Of course, contrary to Korsgaard, I deny that organisms lacking subjective consciousness like plants and insects are capable of pursuing goods (for reasons I give in Chapter 1). As such, though I disagree with Korsgaard's understanding of who may possess goods, I nonetheless agree that those beings that can possess goods are the beings that matter.

⁴⁵ Regan, *The Case for Animal Rights*, 246.

⁴⁶ I will argue in Chapter 6 that it is agency, not autonomy, which grants one interests in freedom from pain, continued existence and liberty for example.

Thus three qualities taken to be good candidates for sufficient conditions for moral status are all tightly linked to agency. In fact the link between some of these capacities is even closer than discussed here since all individuals of which we have knowledge who possess either agency, sentience or are subjects-of-a-life are coextensive. It seems likely that this is because it is nomologically necessary that agents possess these other features.⁴⁷

My argument also claims that agency is necessarily linked to the possession of interests. This is significant since having interests has been shown to be tightly tied to the notion of moral status. Agents necessarily have interests or else they could have no motivation to act in a given way. Without any motivation, one's actions would be purposeless and done for no reason at all. As such they could not truly be said to be 'actions' but merely behaviour. Thus, though one might be animate, one would lack intentionality and so not be an agent in any morally significant sense. Further, all interest holders are agents. As I argued in Chapter 1, possessing some subjective consciousness is necessary in order to possess interests (either subjective or objective interests) and as I will go on to argue in Chapters 3 and 4, having such consciousness entails that one is an agent. As such the sets of interest holders and agents are co-extensive, that is, all interest holders are agents and all agents are interest holders.

Thus, the Agency Account is helpfully reductive or unifying because it shows that all the moral status accounts I have considered here are aiming at different features of (or properties closely connected with) the same capacity: agency. The Agency Account clears up and simplifies the complex moral status landscape; this is a theoretical virtue. However it also explains the intuitive appeal of alternative accounts of moral status. If those individuals that possess sentience, agency, subjecthood and interests are closely related, then this can explain why all these accounts have an intuitive pull.

Conclusion

In this chapter I have considered three different accounts of moral status: the Sentience Account, the Subject-of-a-Life Account and the Kantian Account. I argued that despite their virtues and the importance of these accounts in bringing awareness to problems in animal ethics, none of these accounts are ideal. These accounts each suffer from their own problems as well as overall failing to either provide a plausible account of the central *pro tanto* duties

⁴⁷ I will argue in Chapters 3 and 4, that it is also logically necessary that sentient beings are agents.

owed to moral status holders or justify the condition for moral status. An ideal account of moral status, I suggested, would overcome these problems. Further the Agency Account is closer to the ideal account of moral status than the alternatives considered here. I argued that the Agency Account can give an account of the central *pro tanto* duties owed to moral status holders based on the necessary and sufficient condition of the account: agency. Further, the Agency Account provides us with account of moral status which can plausibly explain both human and animal moral status. Finally, because the account is analytically reductive it can capture the intuitive appeal of the other accounts whilst at the same time avoiding their pitfalls.

Chapter 3

From Sentience to Desires



In the preceding chapters I have argued that sentient animals ought to be considered moral status holders. I have also argued how best to understand moral status and that the Agency Account of Moral Status is the best account of moral status available. I will now bring these two strands of argument together and argue over the next two chapters that sentient individuals possess agency.

One way to argue that sentient individuals are agents is through appealing to evolution. Plausibly the evolutionary purpose of being able to feel pleasure and pain is to motivate behaviour that is conducive to one's survival (and the continuation of one's species) by incentivising beneficial actions and de-incentivising harmful actions. There would be no evolutionary advantage for non-agential beings to be able to experience pleasant and unpleasant feelings since they have no ability to modify the way they behave. If one lacks agency then one simply behaves in accordance with fixed routines, one cannot be motivated to act. As such, it is incredibly unlikely that sentience would become a pervasive characteristic of animals, or even of a single species, if those possessing it were not agents.¹ I think that this argument is correct. However at best, it shows that sentient beings that have come into existence through natural means, are likely to be agents. I think that there is a stronger connection between sentience and agency. I contend that all sentient beings must *in principle* be agents.

Therefore, through this and the following chapter, I will argue that the possession of sentience logically entails the possession of agency. That is to say that a being capable of experiencing affective feelings is a being capable of acting intentionally. Thus, despite claims

¹ David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (Cambridge; New York: Cambridge University Press, 1996), 136; Bernard E. Rollin, *Animal Rights and Human Morality* (New York: Prometheus Books, 1981), 41; Victoria Braithwaite, *Do Fish Feel Pain?* (Oxford: Oxford University Press, 2010), 27–28; S. F Sapontzis, *Morals, Reason, and Animals* (Philadelphia: Temple University Press, 1987), 186; Jonathan Balcombe, *Pleasurable Kingdom: Animals and the Nature of Feeling Good* (London: Palgrave Macmillan, 2006), 19–22; Michel Cabanac, "Physiological Role of Pleasure," *Science* 173 (1971); George John Romanes and Charles Darwin, *Mental Evolution in Animals: With a Posthumous Essay on Instinct by Charles Darwin* (Cambridge University Press, 2011), 108.

to the contrary, all sentient animals must be agents. First I will give an account of how I understand sentience then, once this is made sufficiently clear, I will argue that sentience involves possessing desires and then that possessing desires is sufficient for agency. I assume that this relationship is *not* one of bi-conditional entailment and thus, to be clear, I will not argue that possessing agency entails possessing sentience. I do not wish to rule out the possibility of purely rational non-sentient agents; much like how Kant understands angels.

I do not believe my account to be revisionary of the notion of sentience but merely a thorough elucidation of what the capacity amounts to. I also take it that there is sufficient scientific evidence to show that normal members of all vertebrate species are sentient, and that normal members of some invertebrate species (some cephalopods and crustaceans) are sentient in the sense I will defend. Although, whether this scientific evidence is accurate will not impact on my argument here. Since I claim that any sentient being is an agent, I am committed to classifying only those beings that actually are sentient as agents. Thus my argument, whilst not revisionary of the notions of sentience or agency, is revisionary of our ordinary conception of which individuals count as agents. However my point is precisely that the ordinary view of agency is under-inclusive in much the same way as prevalent views on which individuals have moral status are under-inclusive.

I. Evolution and Instincts

Before starting my main argument it is worth briefly considering a common objection many have to the idea of animals as agents. This is that animals cannot be agents because they act merely on instinct. If I was to ask a handful of people whether they think that animals are agents, the chances are that a good portion of them would reply with this objection. However this claim that animals cannot be agents because they act on instinct, is based on a misunderstanding of the notions of instinct and intentional action.

If acting on instinct is taken to mean that one's behaviour is governed by fixed stimulus and response mechanisms, like the mechanism that ensures that plants grow towards a light source, then beings that act solely on instinct cannot act intentionally. Some animals do seem to behave in ways governed by this kind of 'instinct'. The sphex wasp's burrow-checking

routine is a famous example of this kind of behaviour.² When returning to its burrow with paralysed prey, the sphex wasp will leave its prey at the entrance and go inside to check its burrow. Once done, it will return to its prey and bring it into the burrow. However, if the wasp finds that its prey is has moved since it has checked its burrow, it will relocate the prey back to the entrance of its burrow, then begin the burrow-checking routine again. If upon its return its prey has been moved again, the wasp will repeat the same routine again. This routine may continue a great many times (Wooldridge claims this occurred 40 times on one occasion).³

I have no doubt that some animals do operate according to such automatic procedures, however it doesn't seem that sentient animals, the kind of animals that I think are agents, are such animals. For one, behavioural evidence tells us that sentient beings act rather differently from sphex wasps. Moreover, as I argue above, it seems unlikely that sentient animals would operate according to such automatic procedures for reasons relating to evolution.

An alternative notion of instinct, one that is more plausibly the kind of thing that plays a part in sentient beings' actions, is a strong innate desire.⁴ Instincts in this sense may simply be desires that are so deeply ingrained that they cannot (or at least cannot easily) be resisted. If acting on instinct means acting on strong innate desires then beings that act only on instincts should still be understood to be agents.

If instincts are desires, or similar to desires, then they are motivational states. As such, they can only ever *motivate* one to act. Like desires, it seems that instincts can only bring about action when they are paired with an appropriate belief to form a motivating reason (I'll argue for this claim in Sec IV. of this chapter). However if one has motivating reasons I will argue, then one is an agent. The fact that all of an animal's motivational states may be formed from innate desires doesn't undermine the claim they are an agent. When one acts on an instinct one still acts on a motivating reason.⁵ Korsgaard offers a similar account of instinct: '...[A]n animal who acts from instinct is conscious of the object of its fear or desire, and conscious

² Some theorists deny that this often used example is as it seems however: Fred Keijzer, "The Sphex Story: How the Cognitive Sciences Kept Repeating an Old and Questionable Anecdote," *Philosophical Psychology* 26, no. 4 (2013): 502–19.

³ Dean E. Wooldridge, *Machinery of the Brain* (New York: McGraw-Hill Inc., 1963), 82–83.

⁴ A charitable interpretation of 'innate desire' here may be an innate predisposition to form some desire under certain circumstances. I'll discuss this further in Sec. IV below.

⁵ See Steward for a related but competing account of instinct: Helen Steward, "Animal Agency," *Inquiry* 52, no. 3 (2009): 225.

of it as *fearful* or *desirable*, and so as *to-be-avoided* or *to-be-sought*. That is the ground of its action...'⁶

Understanding acting on instinct in this way is also in line with how we commonly understand human instincts. Whilst the term 'instincts' is much more rarely used to describe human action, there are cases where it is clearly appropriate: sexual desire for example appears to be an instinct if anything is. This desire is widespread among our species, a powerful motivator and difficult to resist. It has all the hallmarks of an instinct. However when we act on this instinct we do not take ourselves to act without reason. Sexual desire, despite being powerful and deeply ingrained, merely motivates action, it does not determine our action by itself. It is not the case that we are blindly forced into a certain behaviour because of this instinct.

One important difference between humans and beings that can act only on instinct is that (most) humans can choose, consciously reflect upon and endorse their desires. Purely instinctual beings however cannot reflect upon and endorse (or reject) desires. As such if they have a desire they will presumably act on this desire if there is not counter-motivation. However this does not mean that instinctual beings are not agents. As I will argue over the next two chapters, though many humans possess the capacities to think about and endorse desires and reasons, it does not seem that they are necessary in order for one to intentionally act. Furthermore, it should be noted that cognitive ethnological studies show us that a great many animals do not act purely on innate desires and as such cannot be said to act purely on instinct. In this sense, the actions of many animals are motivated by learned behaviours. A lion's actions in a hunt, for example, are a result of her learning appropriate hunting behaviour from her mother. Thus whilst I accept that a being capable only of acting on instincts is an agent, it is false that instincts can account for all animal action. This being said I will now discuss the nature of sentience and show that it logically entails the possession of desires.

⁶ Contrary to my own thoughts, Korsgaard believes that one cannot be said to be acting for a reason unless one recognises one's reason *as* a reason and as such, though animals are agents, they are not capable of actions on reasons. I will argue for a weaker understanding of acting for a reason upon which acting for a reason requires merely that one is motivated by a reason. Thus on my account, acting on an instinct, is acting for a reason. Christine M. Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals," *The Tanner Lectures on Human Values* 25 (2004): 85.

II. Sentience

Sentience is often described as the capacity to feel pleasure and pain. This is roughly correct in that any sentient animal of which we have knowledge is capable of feeling pleasure and/or pain. However, more precisely, sentience is the ability to experience affective feelings, that is, feelings of a pleasant or unpleasant nature, mental states that have a certain positive or negative 'feeling' or phenomenology to them, feelings like pain, pleasure, anger, pride, loneliness, disappointment, thirst, deliciousness, enjoyment, etc.

Presumably philosophers often refer to sentience as the ability to feel pain and pleasure because these two mental states are paradigmatic instances of the two different types of affective feelings. Strictly speaking however there are many affective feelings difficult to categorise as either instances of pleasure or pain. There are many unpleasant experiences that do not appear to be painful and pleasant experiences that we may not call pleasures. Disappointment for instance, whilst certainly unpleasant, needn't be, and plausibly often isn't, painful. Imagine dropping a freshly-made sandwich before having the chance to take a bite, calling the disappointment of this event 'painful' sounds a little odd, one would surely be accused of dramatising the effects of the event. It would be much more appropriate to call this case of disappointment 'unpleasant'.⁷

Nonetheless, whilst unpleasant experiences may not be necessarily painful, it does seem to be the case that painful experiences are unpleasant. In fact 'painful' seems simply to be used to register that something is unpleasant in a particular way. A similar point can be made regarding pleasure. Feelings such as relief and achievement are pleasant though we may not call them 'pleasures' or 'pleasurable'. We don't take pleasure in the relief of reaching a deadline in the nick of time as we do a warm summer's day. Therefore, following others, I will refer to 'pleasant' and 'unpleasant' experiences as two forms of affective feeling rather than 'pleasure' and 'pain'.⁸

A capacity commonly confused with the capacity for sentience is phenomenal consciousness. A merely phenomenally conscious being would be capable of thinking but without the possession of any capacity to experience affective feelings or even sensations. One

⁷ Chris Heathwood, "The Reduction of Sensory Pleasure to Desire," *Philosophical Studies* 133, no. 1 (2007): 41.

⁸ Timothy Schroeder, "Pleasure, Displeasure, and Representation," *Canadian Journal of Philosophy* 31, no. 4 (2001): 513; Heathwood, "The Reduction of Sensory Pleasure to Desire," 41; Richard B. Brandt, *A Theory of the Good and the Right* (Oxford : New York: Oxford University Press, 1979), 34.

could for instance be an insular disembodied mind, with no experience of the world, that spends all of one's time solving philosophical puzzles. Of course, whether it is nomologically possible that a being that entirely lacked experience of the world and yet was phenomenally conscious could exist is questionable, however the answer to this question is not pertinent to my inquiry. One should merely note that the capacity of sentience is not obviously *logically* entailed by the mere capacity for phenomenal consciousness. Whilst having phenomenal consciousness is a necessary condition for sentience it is not sentience itself. Phenomenal consciousness and sentience (as I have defined it here) apparently come apart.

Another capacity sometimes referred to as 'sentience' is the capacity to feel sensations or qualia. One may for instance possess phenomenal consciousness as well as the ability to experience sensations. Thus one would possess not just some conscious life but be able to take in sensory information from the world around them. However one should also note that possession of the ability to experience sensations, does not logically entail possession of the ability to experience affective states specifically. Just as some beings have the sense of touch but not sight, or humans have the sense of vision but not the ability to use echolocation, having any given capacity to experience certain sensations does not imply that one has the ability to feel affective feelings. Thus it seems logically possible that one could be phenomenally conscious and capable of experiencing sensations yet be incapable of experiencing affective feelings. In fact, I think some such beings may exist in the world around us. Sea anemones, mussels and probably many insects might have mental lives like these (though of course empirical research should be deferred to in order to find out whether there is truth in these assumptions). This capacity then, whilst it is a concomitant of sentience in our world, is not sentience itself and thus not of interest to me here.⁹

III. Accounts of Sensory Pleasant Feeling

Now that the notion of sentience is sufficiently clear, I will argue that the possession of sentience entails desires. To do this I will first consider two kinds of accounts of the nature

⁹ Unlike phenomenal consciousness this doesn't seem to be a necessary condition for sentience. One could be sentient yet lack an ability to experience any sensations. For example one could in principle experience affective feelings resulting from thoughts on mental activities e.g. finding a theorem one has constructed beautiful, or being frustrated by one's inability to grasp why one exists. Perhaps however, one needs to at least be able to experience the passage of time in these cases and this may be considered a sensation.

of sensory pleasant experiences: *the Felt Quality Account* and *the Attitudinal Account*. I will then argue that the best account of the nature of pleasant sensory experiences is a desire-based form of the Attitudinal Account. As such, any sentient being must necessarily possess desires.

By pleasant sensory experiences I mean the pleasant experiences that occur with experiences of sensations. These experiences are not just the sensations that we come to experience through the five classic sensory inputs of taste, touch, hearing, sight and smell but any sensations we come to experience as a direct result of our body including, for example, temperature recognition and nociception (and echolocation and electroreception in other animals for instance). In contrast to these are pleasant non-sensory experiences; those pleasant experiences that result from other stimuli such as emotional or intellectual stimuli. Pleasant experiences such as having one's first child or being offered one's dream job would be instances of such experiences. I refer to these merely as pleasant non-sensory experiences to remain neutral on how best to categorise such experiences.¹⁰

Here I will focus predominately on accounts of pleasant sensory experiences rather than pleasant non-sensory experiences or pleasant experiences generally. My reason for doing this is that it seems to me that pleasant non-sensory experiences can be plausibly and intuitively explained by an attitudinal account similar to the one I give below. This I suppose is because such experiences seem to involve desires in an obvious way.¹¹ Pleasant sensory experiences however are a more difficult case and hence an attitudinal account of sensory pleasant experiences is more controversial. Thus, as the more difficult case, it is worth making it the focus of my argument. It should also be noted that though much of what I say here only refers to pleasant experiences, one should take an account of a very similar form to apply to unpleasant experiences. Thus the account I give here, though phrased in terms of pleasant sensory experiences, could with little adjustment, form an account of sensory affective feelings generally.

¹⁰ Heathwood recognises three categories of pleasure. Aydede also makes a distinction between sensory and non-sensory pleasure. Heathwood, "The Reduction of Sensory Pleasure to Desire," 28; Murat Aydede, "A Contemporary Account of Sensory Pleasure," in *Pleasure: A History*, ed. Lisa Shapiro, vol. 1 (Oxford: Oxford University Press, 2014), 4; Murat Aydede, "How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal," *Review of Philosophy and Psychology* 5, no. 1 (2014): 2.

¹¹ Heathwood also assumes this and takes up this method. Heathwood, "The Reduction of Sensory Pleasure to Desire," 28.

In giving an account of why we find given experiences pleasant, and others unpleasant, it is broadly accepted that there are several principal problems that any account should address. These problems are:

1) What unites all instances of pleasure? This is known as *the Heterogeneity Problem*.¹² A good account of pleasant sensory experiences will explain how all the different experiences we call 'pleasant experiences' are united. The pleasant taste of just ripe cherries is very different from the pleasant feeling of riding a rollercoaster, and both are different still from the pleasant experience of listening to one's favourite band playing live. There are many different experiences that we call pleasant and they seem to possess a wide variety of phenomenological traits. If these are all instances of pleasant experiences however they must all be tokens of the same type, and must all be the same kind of thing. Thus a good account of affective feelings should explain why these disparate experiences are all instances of pleasant experiences despite their apparent differences. (The same also applies to all cases of unpleasant experiences).

2) *The Problem of Opposite Values*.¹³ Pleasant and unpleasant experiences seem to be opposites yet the experiences which they accompany are not always opposite. For example one might find eating burritos pleasant and doing household cleaning unpleasant to the same degree. Yet eating burritos is not the opposite of cleaning one's house. As such the oppositeness of pleasant and unpleasant feelings cannot come from the experiences which they accompany (in this case eating burritos and cleaning the house). There must be something about the nature of pleasant and unpleasant feelings themselves that explains their opposite values. Any plausible account of affective feelings must explain what it is about these feelings that makes them opposites in this sense.

3) *The Problem of Motivation*.¹⁴ Ordinarily we feel motivation to pursue pleasant experiences (and motivation to avoid unpleasant experiences). However we don't feel motivation to pursue most other experiences in themselves. That we find something yellow, large or wet isn't enough by itself to motivate us to either pursue or avoid it. Pleasant and unpleasant feelings then, appear unique in that they have a very close link with motivation. This link with motivation seems to be an integral part of the notions of pleasantness and

¹² Ibid., 25; Aydede, "A Contemporary Account of Sensory Pleasure," 7.

¹³ Heathwood, "The Reduction of Sensory Pleasure to Desire," 26; Aydede, "A Contemporary Account of Sensory Pleasure," 11.

¹⁴ Aydede, "A Contemporary Account of Sensory Pleasure," 8.

unpleasantness. As such, an account of affective feelings needs to provide an explanation for why we are motivated to pursue pleasant feelings and avoid unpleasant feelings.

i. Felt Quality Accounts

According to Felt Quality Accounts pleasant and unpleasant feelings are a result of us simply finding experiences to be pleasant or unpleasant to some extent. Ben Bramble argues that the feeling of pleasantness of a given experience is a result of us making an unconscious normative ‘judgment’ that our current experience is good for us in some sense.¹⁵ As such, the experience is presented to our conscious mind as having a pleasant feeling. Likewise, experiences are accompanied by unpleasant feelings because we judge such experiences to be bad for us in some sense.

Felt Quality Accounts provide a straightforward answer to The Problem of Opposites. What makes pleasant and unpleasant feelings opposite to each other is that they supervene on our unconscious ‘judgments’ of what is good and bad for us respectively. So because ‘good for’ and ‘bad for’ are opposites, this can explain why we take pleasantness and unpleasantness to be opposites.

In providing an answer to the Heterogeneity Problem, Felt Quality Accounts offer two different alternatives. Some theorists, such as Bramble, endorse the claim that affective feelings have a *distinctive feeling* that unites all different instances of pleasant and unpleasant feelings respectively.¹⁶ Others argue that all affective feelings are experiences of the *hedonic tone* of given sensations.¹⁷ Distinctive feeling accounts seem to rely on our intuition to be able to reliably pick out the feelings that are pleasant, from the ones that are unpleasant or neutral. Hedonic tone accounts however, suggest that all experiences have a certain hedonic tone to them. The hedonic tone of an experience can be anything from extremely positive to extremely negative. Even experiences that don’t produce an affective feeling in us still have a hedonic tone value, albeit one that is affectively neutral between pleasant and unpleasant. So what

¹⁵ Aydede refers to this process as a ‘felt evaluation’. Ben Bramble, “The Distinctive Feeling Theory of Pleasure,” *Philosophical Studies* 162, no. 2 (2013): 213–15; Aydede, “A Contemporary Account of Sensory Pleasure,” 4; Aydede, “How to Unify Theories of Sensory Pleasure,” 13–15.

¹⁶ Bramble, “The Distinctive Feeling Theory of Pleasure,” 202.

¹⁷ *Ibid.*, 203; Aydede, “How to Unify Theories of Sensory Pleasure,” 4; Aydede, “A Contemporary Account of Sensory Pleasure,” 16.

unites all pleasant feelings on these accounts is that they are all instances of experiences with a positive hedonic tone.

The hedonic tone, much like the brightness of a visual experience or loudness of an auditory experience, seems to colour our experience in a certain way, and is part of the experience itself, rather than a result of the experience as attitudinal theorists take affective feelings to be.¹⁸ Aydede uses the example of the speed of a dance as an analogy of the hedonic tone of an experience.¹⁹ Whilst a dance can come in different varieties, a foxtrot, a waltz, etc. the speed of any given dance can vary in each case. Further one cannot have a dance with no speed at all, though a dance may be not particularly fast or slow (i.e. affectively neutral). And the speed of a dance is part of the dance itself and affects how one perceives the dance as a whole (just like how pleasant sensory feelings are part of sensory experiences).

Finally in answer to the Motivation Problem, Felt Quality theorists take pleasantness to be an intrinsically motivating feeling (and unpleasantness to be something that we are intrinsically motivated to avoid).²⁰ More can be said however if one endorses Bramble's claim that affective sensations are a result of unconscious 'judgments' about whether experiences are good or bad for one.²¹ One may argue that because we judge something to be good for us it motivates us to pursue it and where something is bad for us it motivates us to avoid it. So our (albeit unconscious) recognition that the source of pleasant feelings is good, and the fact that we are motivated to pursue that which we think is good for us, explain how we are motivated to pursue those things we find pleasant.

This is not to say that these 'judgements' provide us with motivating reasons to find given experiences pleasant or unpleasant. This cannot be the case since we do not have cognitive access to these 'judgments'. This explanation is merely supposed to show why pleasant feelings are desirable to us and thus why we are motivated by them. Though one should note that a given experience being pleasant is seen to be a motivating state by Felt Quality theorists (Aydede mentions this specifically).²² Thus pleasant feelings themselves are

¹⁸ Shelly Kagan, "The Limits of Well-Being," *Social Philosophy and Policy* 9 (1992): 172-73.

¹⁹ Aydede, "A Contemporary Account of Sensory Pleasure," 18; Aydede, "How to Unify Theories of Sensory Pleasure," 15.

²⁰ Aydede, "A Contemporary Account of Sensory Pleasure," 9.

²¹ Bramble, "The Distinctive Feeling Theory of Pleasure," 213; Aydede, "A Contemporary Account of Sensory Pleasure," 19-22; Aydede, "How to Unify Theories of Sensory Pleasure," 13.

²² Aydede, "A Contemporary Account of Sensory Pleasure," 4.

intrinsically motivating and these pleasant feelings are felt as a result of given experiences being 'judged' to be good for us at an unconscious level.

Feeling something as pleasant then is an attempt by our unconscious mind to motivate us. As Bramble himself notes we shouldn't take this 'judging' to be too cognitively complex since all those beings that are sentient must be capable of engaging in such 'judgments'.²³ Plausibly this unconscious 'judgment' seems to be little like a full evaluative judgment. Perhaps Bramble has something like Street's 'proto-judgments' in mind here. This being the case we should understand a judgment of this kind to be something more like an evolutionarily primitive 'yearning for' or 'going for'.²⁴ This tendency would be both unreflective and psychologically inaccessible, occurring at a level of the mind to which we lack cognitive access.

ii. Attitudinal Accounts

The more popular accounts of sensory pleasant feelings are *Attitudinal Accounts*.²⁵ According to Attitudinal Accounts pleasant feelings are a result of our attitudes being satisfied or fulfilled, or (according to some accounts) the appearance that our attitudes are satisfied. The most straightforward Attitudinal Account, and I think the most plausible and convincing, is a Desire-Based Attitudinal Account.²⁶ On a Desire-Based Attitudinal Account (henceforth merely Desire Account) pleasant affective feelings are felt when a certain desire of ours is satisfied. A similar story can be told with regard to unpleasant feelings: we feel an unpleasant affective feeling when a certain desire is frustrated.

Importantly, pleasant feelings do not result from the satisfaction of all desires. In order for one to receive a pleasant sensory feeling from the satisfaction of a desire, the desire must be intrinsic or what one might call a 'final desire'.²⁷ This does not mean that the desire must

²³ Bramble, "The Distinctive Feeling Theory of Pleasure," 216.

²⁴ Sharon Street, "A Darwinian Dilemma for Realist Theories of Value," *Philosophical Studies* 127, no. 1 (2006): 119.

²⁵ Heathwood, "The Reduction of Sensory Pleasure to Desire"; Schroeder, "Pleasure, Displeasure, and Representation"; Brandt, *A Theory of the Good and the Right*; Henry Sidgwick, *Methods of Ethics*, 7th ed. (Indianapolis: Hackett Publishing Co, Inc, 1981).

²⁶ Schroeder and Heathwood defend well developed accounts of this position. Schroeder, "Pleasure, Displeasure, and Representation"; Heathwood, "The Reduction of Sensory Pleasure to Desire."

²⁷ Kagan, "The Limits of Well-Being," 173-74. Kagan refers to the relevant kind of desire as an 'immediate desire' since the desire is for the immediate properties of an experience rather than any instrumental properties. I'll use the term 'final desire' since I think this is the clearest and least ambiguous term available.

be a desire for an object in virtue of its internal properties, but that the object of one's desire must be desired for that object's *own sake* and for no further reason.²⁸ Thus in order to gain a pleasant feeling from the taste of a ripe cherry I must desire the taste of the cherry itself, for its own sake, and for no further reason. An instrumental desire alone will not result in pleasant feelings. Hence if I merely desire to taste a ripe cherry because I grow cherries for profit and I want to know if my cherries are ripe for harvest, then the satisfaction of my desire to taste a ripe cherry itself would not be pleasurable.²⁹ Pleasant and unpleasant feelings on this account just *are* the satisfaction and frustration of final desires respectively.

For the satisfaction of a desire to result in pleasantness, it seems the desire must also have as its object, some sensory experience, rather than an object in the world.³⁰ If I eat an apple, but perhaps due to the strong aftertaste of another food, I cannot taste the flavour of the apple, I don't seem to have any pleasant experiences. Likewise, if I have a vivid daydream about eating an apple, I have a pleasant experience even though I haven't actually eaten an apple. This is the case because the object of my desire is not some interaction with objects in the world but the sensations that are caused by such interactions. I have a desire for the experience of the taste of an apple, rather than a desire for the actual eating of an apple.

Similarly the fact that the relevant desires have experiences as their objects can also explain why I may not experience a pleasant feeling from seemingly having a desire satisfied whilst I am intoxicated, distracted or under the influence of mind-altering substances. Since such circumstances affect the way one experiences the world, if I am distracted by an intense conversation whilst eating my apple for example, I may not experience any pleasant feeling since I do not actually experience the taste of the apple. Distracted by the conversation I simply don't pay attention to the relevant sensory input and fail to generate an experience of the taste of an apple. In cases of a lesser degree of distraction or intoxication it may be that I experience some taste sensation but do not experience the full taste sensation that normally accompanies eating an apple. As such I experience a lesser pleasant feeling than I would if I was to experience the full taste of my apple because, because my desire is not satisfied to the same extent as it would be, were I to taste an apple when I was clear-headed and free of distractions.

²⁸ Heathwood, "The Reduction of Sensory Pleasure to Desire," 30; Aydede, "A Contemporary Account of Sensory Pleasure," 15.

²⁹ Heathwood, "The Reduction of Sensory Pleasure to Desire," 30; Aydede, "A Contemporary Account of Sensory Pleasure," 3.

³⁰ Heathwood, "The Reduction of Sensory Pleasure to Desire," 29; Kagan, "The Limits of Well-Being," 173.

Since I am here only giving an account of pleasant *sensory* feelings, all such desires must also be desires for a mere sensory experiences and not more complex experiences. A final desire for feeling accepted by a group of people isn't a relevant desire here for example. Feeling accepted is not a sensory feeling but a complex experience that we can come to have through interpreting our sensations and inferring from them that certain events have occurred or will occur in the world. That is not to say that such experiences are not a source of pleasant feelings, but only that an account of pleasant sensory feelings cannot explain this kind of pleasant experience. This would be explained on an account of pleasant non-sensory feelings.

Not all Desire Account theorists accept that this is the case and Schroeder offers an alternative argument. Schroeder suggests that our actual desires don't need to be satisfied in order for us to feel pleasantness but that it only needs to *appear to us* as if our desires are satisfied.³¹ If one prefers to think of the Desire Account in this way then it will make little difference to my argument here, however I will continue to phrase my account in terms of desires for experiences being satisfied. This I think is a more straightforward and parsimonious way to understand the role of desires in affective feelings. One should note that I may refer to a desire for some object rather than a desire for the sensations or experience of some object. For instance, I may refer to a desire for water where I actually mean a desire for the sensation of drinking water. This is merely for convenience, one should not take this to be philosophically significant.

Heathwood also suggests that the desires that result in pleasant feelings must be *de re* desires for given sensations. Thus, Heathwood suggests, my desire for the taste of an apple should be understood as a desire for a particular sensation that I understand as the taste of an apple, and usually associate with eating apples (because this is the usual cause of the sensation). My desire should not be understood to be a desire for whatever sensations my eating of an apple causes me to experience. This stipulation is important because if I had a *de dicto* desire for the taste of an apple, and bit into a quince mistakenly thinking it was an apple (and it tasted just as I expected an apple to taste), then my desire for the taste experience of an apple would go unfulfilled, since I would have experienced the taste of a quince not an apple. As such, a pleasant feeling should not result from this experience, yet it seems highly likely, assuming I like the taste of apples, that it would. If however I had a *de re* desire for the taste experience of an apple in the same situation, my desire would be satisfied by biting into the

³¹ Schroeder, "Pleasure, Displeasure, and Representation," 516-19.

quince even though I am mistaken about the fruit I am eating. Since my desire in this case would be for the specific flavour profile, when I experience this taste, biting into the quince, my desire is satisfied. Thus despite the fact that this taste experience is generated by a quince not an apple, I experience a pleasant feeling. This is what we would expect and as such it must be that pleasant affective feelings result from *de re* desires not *de dicto* desires.³²

Finally, it seems that some of our final *de re* desires for sensations are more or less 'hard-wired' and determined through natural selection. The desire for the sensation of drinking water seems likely to be of this kind. This can explain why we seem to share some desires (desires for the sensations of eating food, drinking water, having sex, etc.) with other sentient animals. And it explains why we share many similar desires in common with our closest evolutionary ancestors and relatives. Other desires appear more malleable and can be changed through exposure to a given sensation.³³ The desire for the taste of coffee or sugar is a good example of this kind of desire. How exactly we are able to change these desires is not a question I will consider here however.

The Desire Account provides clear answers to the three main problems that an account of affective feelings needs to resolve. The Heterogeneity Problem can easily be dealt with by making reference to the attitudes underlying pleasant feelings.³⁴ What unites all pleasant feelings on the Desire Account is that they are all a result of the satisfaction of some final *de re* desire for some sensation (and unpleasant feelings a result of the frustration of final *de re* desires for sensations). Thus the Desire Account accepts that pleasant feelings can be phenomenologically very different from each other, yet also provides a good reason for understanding them to be experiences of the same kind.

It should be obvious how such an account provides an answer to the Problem of Opposites. The Desire Account again accepts that the experiences which we find pleasant or unpleasant, need not be opposites in order to explain why pleasant and unpleasant feelings are opposites. Since pleasant feelings arise through desire satisfaction and unpleasant feelings arise through desire frustration, and desire satisfaction and frustration are opposites, pleasant feelings and unpleasant feelings are themselves opposites.³⁵

³² Heathwood, "The Reduction of Sensory Pleasure to Desire," 31.

³³ Aydede, "A Contemporary Account of Sensory Pleasure," 18.

³⁴ Aydede, "How to Unify Theories of Sensory Pleasure," 6.

³⁵ Heathwood, "The Reduction of Sensory Pleasure to Desire," 27; Aydede, "How to Unify Theories of Sensory Pleasure," 6.

An answer to the Problem of Motivation also easily falls out of this account. According to the Desire Account, experiencing pleasant feelings *just is* satisfying a final *de re* desire. So those things which I find pleasant are the sensations which I have a final *de re* desire for. Since I desire these sensations, I am motivated to pursue or attempt to ensure the continuation of these sensations.³⁶ Thus sensations that I find pleasant are the sensations that I am motivated to pursue. Again a similar explanation can also be given regarding our aversion to unpleasant experiences.

It is important to note though that it is not the pleasant experiences of a sensation that motivate us according to the Desire Account.³⁷ Since pleasant feeling *is* desire satisfaction on this account, pleasant feeling cannot be what motivates us. It cannot be the case that I desire the taste of an apple because I find the taste of an apple pleasant. Desire Account theorists accept that this explanation is backwards and argue that in fact I find the taste of an apple pleasant because I desire it. Thus contrary, to psychological hedonism, we do not desire things because they are pleasurable but we find things pleasurable because we desire them.

IV. Pleasant Sensory Feelings as Desire Satisfaction

Attitudinal Accounts are more widely accepted than Felt Quality Accounts and it seems to me that there is good reason for this. Although they present some problems, of the two competing accounts, I maintain that a Desire-Based Attitudinal Account is the most plausible, and the account we should endorse. It has several benefits over its competitor and significant intuitive appeal. In this section I'll give my positive argument in defence of the Desire Account of Pleasant Sensory Feelings and then consider some objections that Desire Account theorists must answer.

A major benefit of the Desire Account is that it allows for the separation of sensory experiences and affective feelings. This has significant intuitive appeal since pleasant experiences don't seem to be inseparably linked to sensory experiences as Felt Quality Accounts seem to suggest. There are two examples that I wish to draw on to illustrate this point. Firstly consider the case of drinking too much of your favourite drink at a party, a case perhaps all too familiar to some of us. Dan's favourite drink is vodka and Coke. So when he

³⁶ Aydede, "A Contemporary Account of Sensory Pleasure," 12.

³⁷ Aydede, "How to Unify Theories of Sensory Pleasure," 6.

drinks it (and he's sober enough to fully experience the taste of it) he experiences a pleasant feeling. On his birthday this year Dan throws caution to the wind and ends up drinking a few too many vodka and Cokes. He regretfully spends most of the evening of his birthday in the toilet vomiting and the next day lounging around the house trying to get over his hangover. A week later however he's feeling fine and attends a party at a friend's house. On arrival his friend has already poured some drinks and hands one to Dan who immediately takes a gulp. It's a vodka and Coke. However Dan doesn't have a pleasant feeling when he tastes the drink, in fact he finds it repulsive and it make him wretch. It's the same brand of vodka and Coke that he usually drinks and he is clear-headed. What is more, when a friend asks Dan if the drink tasted different to how he expected, he replies that it tasted the same but he merely didn't like the taste anymore.

A plausible explanation of what has happened here is that the sensory experience has remained constant (or at least very similar) but that the accompanying affective feeling has changed. This is because, the Desire Account says, Dan's desire for the taste of vodka and Coke has changed. Since his birthday he has lost his desire for this taste and thus no longer finds it pleasant. According to the Felt Quality Account it would have to be the case that Dan's unconscious evaluative judgment of vodka and Coke has changed from being good for him to being bad for him. This seems *prima facie* implausible. It is not clear why or how this judgement could change. Some medical conditions reinforce Dan's case and also show that affective feelings can be separated from sensory experiences.

One such case is that of sufferers of *pain asymbolia*. Pain asymbolia is a condition caused by lesions on the brain. Sufferers of the condition (asymbolics) report that they are able to feel pain but do not have any aversion to their feeling of pain. Such individuals not only claim indifference to their pain, but often laugh and show signs of enjoyment, as well as showing no concern over the damage to their bodies. One model of what is occurring in the mind of the asymbolic suggests that the asymbolic has lost some desire(s) related to the experience of sensations we usually consider to be painful.³⁸ Thus, because they lack a relevant desire or desires for the absence of these sensations, they don't have any desires frustrated when they experience them.³⁹ So when asymbolics state that they are in pain, they are not experiencing any unpleasant affective feelings but in fact only experiencing sensations we ordinarily

³⁸ Colin Klein, "What Pain Asymbolia Really Shows," *Mind* 124 (2015): 6.

³⁹ Which desires they lack specifically is difficult to say. Klein suggests that the evidence tells us that they lack a desire for bodily integrity although I suspect they also lack at least a desire for the absence of certain sensations caused by nociception. Klein, "What Pain Asymbolia Really Shows."

describe as painful e.g. the tearing of skin or a sharp prick from a needle. However because they have always associated these sensations with strong unpleasantness, they still identify them as 'painful' even though they do not find them unpleasant.

Consider another disorder: Persistent Sexual Arousal Syndrome (PSAS) also known as Persistent Genital Arousal Disorder (PGAD). PSAS is a condition in which one's genitals are frequently, spontaneously and involuntarily stimulated.⁴⁰ The stimulation is so frequent that individuals with the condition can experience up to hundreds of orgasms a day. If orgasms are intrinsically pleasant experiences, as is often suggested, then one would expect this condition to have overall pleasurable effects. However, some sufferers of PSAS have reported that their involuntary orgasms are unpleasant, not just that they cause embarrassment, distress or depression (though sufferers of PSAS often state that is this case) but that the orgasms themselves as unpleasant.⁴¹

If sufferers of PSAS find the sensation of an orgasm unpleasant, and sufferers of pain asymbolia can find the sensation of their flesh being damaged affectively neutral, or even mildly pleasant, then this shows that sensory input can be divorced from affective feeling and that affective feeling is likely a reaction to sensory input, not part of the input. These accounts along with the case of Dan, give the Desire Account significant intuitive appeal. Other than the intuitive explanation in which the Desire Account can give in these cases there seems to be some further significant ways in which the Desire Account is superior to Felt Quality Accounts.

Since Felt Quality Accounts do not appeal to desires or desire like states (or at least they fail to do this at level of the mind to which we have cognitive access)⁴² they struggle to present a plausible solution to the Problem of Motivation. Unlike the Desire Account it is not obvious why we are motivated to pursue pleasant experiences on the Felt Quality Account. Both Bramble and Aydede claim that pleasant feelings motivate us because we view them as good and we are motivated to pursue what is good for us. However this explains little. As

⁴⁰ Sandra Leiblum et al., "Persistent Sexual Arousal Syndrome: A Descriptive Study," *The Journal of Sexual Medicine* 2, no. 3 (2005): 332.

⁴¹ *Ibid.*, 335.

⁴² Aydede refers to 'p-desiring' which he takes to be the exercising of a somewhat desire-like attitude. However such a process takes place at level at which we do not have cognitive access. Thus it is not clear to me how such a state could be anything like an attitude as we normally understand the term or factor into motivation at a conscious level. This is also, why I here categorise Aydede's account as a Felt Quality Account rather than a 'unified account' as he himself refers to his account. Lacking reference to anything recognisable as an attitude, his account seems much closer to a Felt Quality Account. Aydede, "How to Unify Theories of Sensory Pleasure," 13.

Aydede and Heathwood note, the Desire Account has the benefit of reducing affective states to another concept and thus is a more illuminating theory of the nature of pleasure.⁴³ Moreover the Felt Quality Account's solution of informing us that unconscious mechanisms present something to us as pleasant, and that pleasant feelings are motivating, seems to be merely a restatement of the problem rather than an attempt to provide us with an explanation of how this occurs. Even relying on the claim that we are motivated by what is good for us seems questionable. There are a great many things that we find pleasant that aren't good for us (and vice versa). But even if this wasn't the case, the Felt Quality Account doesn't explain how a belief that something is pleasant is sufficient to motivate us to act. There needs to be more explanation of how the motivation comes about. So it seems that Felt Quality Accounts don't provide an adequate answer to the Motivation Problem.

The Desire Account has a further advantage. It seems that the best account of non-sensory pleasant feelings is likely to be desire-based since hedonic tone doesn't appear to sit easily with non-sensory experiences. I will not attempt to provide a full defence of the claim that non-sensory pleasant feelings can be best accommodated by a desire-based account, but take it that this is *prima facie* plausible and intuitive. Assuming that this is correct, endorsing a Desire Account of pleasant sensory feelings gives a unified account of pleasant feelings generally that claims all pleasant sensory and non-sensory feelings are a result of desire satisfaction.

Moreover, if one endorses a Desire Account regarding pleasant non-sensory feelings but a Felt Quality Account with regard to sensory pleasant feelings it would entail that we gain pleasure through two different mechanisms. This seems unlikely (though is not impossible). More problematic however is that these two accounts of pleasant feelings in particular appear to be in tension. If it was the case that a Desire Account best explained pleasant non-sensory feelings, but a Felt Quality Account best explained pleasant sensory feelings, then this would mean that we could desire some objects because they are pleasant and that we could also find other objects pleasant because we desire them. This is implausible and as such it doesn't seem that a Felt Quality Account and a Desire Account could both be correct (even with regard to different types of pleasant feelings). A single unified account of affective feelings would neatly and intuitively explain the nature of pleasant feelings without this problem and it seems that a Desire Account is the best candidate for this role.

⁴³ Ibid., 6; Heathwood, "The Reduction of Sensory Pleasure to Desire," 24.

Of course it may be the case that pleasant sensory and non-sensory feelings are sufficiently different and thus need not be accounted for on a single unified account. However if one took this stance, it seems one would have trouble defending both sensory and non-sensory pleasant feelings as different forms of pleasant feeling and thus would not be able to give an adequate answer to the Heterogeneity Problem. Further, considering the phenomenological similarities between these two types of feeling, it seems unlikely that they are all that different. This, along with a worry that any account that tries to argue for two distinct kinds of pleasant feelings will take us too far from common sense, makes such an option untenable.

It is a major benefit to adopting the Desire Account then, that it can give us an account of pleasant feelings generally, allowing us to see how pleasant sensory feelings fit in the overall picture of pleasant feelings. Further the intuitive appeal of the Desire Account shows it to be a strong contender independent of this benefit. I will now defend the Desire Account against some objections.

i. General Desires

Firstly the problem of general desires for pleasant feelings. Most individuals have general as well as specific desires for pleasant feelings, or at least so it seems. Sometimes, for instance I want to eat something tasty. However I am not particularly discriminating about how such a desire is satisfied, so long as the means give me pleasant feelings, the desire will be satisfied. Desire Account theorists deny that we desire things merely because they are pleasant and argue instead that we gain pleasant feelings through the satisfaction of desires. It seems that the Desire Account will be unable to account for this, and other general desires, for pleasant feelings.

This is not the case. We can still have a general desire for pleasant feelings, or at least we can still have desires that are functionally equivalent, on the Desire Account. When one has a general desire for something pleasant tasting we can understand this as a desire for more of one's desires to be satisfied regarding taste sensations. Or, to put it another way, a desire for one's net desire satisfaction level to increase with regard to taste sensations. I can fulfil this desire by fulfilling any one of my more specific desires, for example, my desire for the taste of pizza or my desire for the taste of refried bean burritos. Thus I gain pleasant feelings through the satisfaction of my desire for the taste of burritos or pizza, and satisfaction of either of these

desires is constitutive of satisfying my desire to experience something tasty. However since this higher order desire for desire satisfaction is not a desire for sensations, its satisfaction does not grant us pleasant sensory feelings. Whilst the satisfaction of this desire may result in some pleasant affective feeling, in addition to the sensory affective feeling that arises from satisfying the relevant sensory desire, this pleasant feeling will be of the non-sensory variety.⁴⁴ Thus, plausibly it seems that one can still have general desires for pleasant experiences even though properly these are not desires *for* pleasure but merely desires for further desire satisfaction.⁴⁵

ii. Desires for No Reason

A further apparent problem one might have with the Desire Account is that it entails that we possess some desires for no reason.⁴⁶ Ordinarily we might say that we desire the taste of an apple because it is pleasant but on the Desire Account we cannot make such a claim. Desire Account theorists are committed to claiming that we simply desire the taste of an apple for its own sake and no further reason. If we could give a reason for our desire it wouldn't be a final desire, but an instrumental desire and so couldn't be the source of pleasant feelings.⁴⁷ As such, desire theorists must accept that we have some desires for no reason.⁴⁸

Having desires for no reasons shouldn't concern us though. If one rejected the Desire Account and endorsed a Felt Quality Account, one would still have to accept that we possess some desires for no reason. Even accounts like psychological hedonism that attempt to explain our motivations by suggesting that there is a only one single fundamental motivation (the desire for pleasure) have to accept that at least the desire for pleasure is a desire we hold for no reason. Desire Account theorists simply accept that there are various experiences that we desire for no reason, rather than one from which all other desires are motivated.

Further one should note that talk of having a desire for 'no reason' is somewhat misleading. It is not the case that no explanation can be given for why we possess these desires but only that a certain kind of explanation cannot be given. Specifically one cannot 'rationalise'

⁴⁴ The precise conditions under which this kind of pleasant feeling can arise cannot be laid down here since it is more properly the subject of an account of pleasant non-sensory feelings.

⁴⁵ Schroeder, "Pleasure, Displeasure, and Representation," 527.

⁴⁶ Heathwood, "The Reduction of Sensory Pleasure to Desire," 30.

⁴⁷ *Ibid.*, 37.

⁴⁸ Schroeder, "Pleasure, Displeasure, and Representation," 526; Aydede, "A Contemporary Account of Sensory Pleasure," 15.

why one has a given desire, or to put it another way, the possession of these desires is not something that is rational or irrational.⁴⁹ This is because these desires are not desires we are motivated to possess, these are desires that we simply have, and for the most part, we are stuck with. As such it just doesn't make sense to ask why we have chosen them over some other desires.

However we can give an evolutionary explanation for why we have such desires. For instance, it shouldn't be surprising that we desire the absence of the sensation of hunger since natural selection would surely have favoured individuals that had this desire, over those that didn't, considering the human need for nutrition in order to survive. Thus in one sense the reason I possess the desire for the absence of the sensation of hunger is because it was advantageous to my ancestors. Nonetheless the object of this desire is an object I am motivated to pursue for its own sake and no further reason.

iii. Moods

One problem for the Desire Account is accommodating moods, and more importantly, a problem for my argument that sentience entails agency is how a being capable only of moods could be an agent. In response to this objection I should say that moods present a problem for both Felt Quality Accounts and Desire Accounts, as such failing to accommodate moods shouldn't be considered a mark against Desire Accounts specifically. Secondly, moods seem to come about through rather complex circumstances and not the mere experience of a sensation such as hotness, bitterness or sweetness. They seem much more similar to complex intellectual or emotional feelings like grief, disappointment, excitement (perhaps even being instances of such feelings spread over time). Hence, Moods seem unlikely to be a type of affective *sensory* feeling. As such it should be no surprise that they aren't accommodated in my account of pleasant sensory feelings, they would be the subject of an account of affective non-sensory feelings.

Nonetheless for my argument in particular it must be shown that it is at least plausible that moods could be accommodated on an account of pleasant non-sensory feelings. This is not a problem however. I maintain that affective feelings in general result from similar final desires as affective sensory feelings do. Moods seem to be no exception. In many cases we can

⁴⁹ Heathwood, "The Reduction of Sensory Pleasure to Desire," 30.

pinpoint a satisfied desire as the reason for a good mood, for example, the fact that it's payday or the fact that I'm on holiday for a week. Similarly, the frustration of a desire can plausibly put one in a bad mood. Moods then, much like other affective feelings, can intuitively be understood as being a result of desire satisfaction and desire frustration.

One may point out that we can be in a mood without being aware of what the object of the mood is; without knowing why we are in a mood. However this doesn't seem to be an insurmountable problem either. Whilst this doesn't occur with sensory affective feelings this does seem to plausibly happen with other complex affective non-sensory feelings. I can gain pleasant feelings from playing a game with my friends without being directly aware of the why I am feeling pleasant feelings. Moods do not seem to present a unique problem to my argument here. Thus any individual that is capable of experiencing merely moods must have some desires to be satisfied or frustrated. If not they wouldn't be able to experience any such mood. Therefore an individual capable of moods must have desires (and therefore, I will argue, be an agent) just as any individual capable of experiencing merely pleasant sensory feelings.

iv. Lost Desires

Another objection to the Desire Account is the case of individuals not desiring given sensations yet still feeling pleasantness when experiencing such sensations. Take Bramble's example of an individual who comes to see their life as worthless and loses interest in all those things she once desired.⁵⁰ Such an individual should not find eating or drinking pleasant for example, yet she does. This being the case Bramble argues, her pleasant feelings must arise from something else other than desire satisfaction.

The answer here seems to be simply that this individual was mistaken about her desires. She does desire the taste of certain foods and drinks, even if she thinks she doesn't (or at least she desires the absence of the sensations of hunger and thirst). As mentioned earlier some desires are particularly 'hard-wired' and may be incredibly difficult to alter. It seems unlikely that she could cease to desire the absence of the sensations of hunger and thirst through sheer force of will. On the other hand, some desires are not seemingly so tough to change. It may be that this individual could lose the desire to leave the house and perhaps

⁵⁰ Bramble, "The Distinctive Feeling Theory of Pleasure," 215.

even the desire to continue living. But the lack of motivation to continue living doesn't nullify the motivation to avoid the sensations of hunger and thirst. These are desires most of us cannot change. Thus though we can choose not to satisfy them, we cannot choose not to possess them.

A similar point can be made to explain why one will still find the taste of one's favourite foods pleasant when one has formed the desire to stop eating them and lose weight. This higher order desire to lose weight, and stop eating one's favourite foods, does not relieve one of the desire for the specific taste of pizza, for instance. It merely introduces a counter-motivation to not pursue the desire for the taste of pizza. This fits with the Desire Account and our common sense intuitions in this case.

v. Novel Experiences

Sometimes when we experience sensations we have never experienced before, we experience pleasant feelings. However according to the Desire Account we must already possess a desire for some sensations in order to find them pleasant. How is it we can find such novel experiences pleasant then? It cannot be the case that we already have desires for those things which we have never experienced. Firstly it simply doesn't seem like this is the case, we aren't motivated to pursue specific sensations which we have never tried before. We sometimes speak of having a desire to try new things but this is different. We are not here concerned with a desire merely to experience previously unexperienced sensations (such desires appear instrumental). We are concerned with a final *de re* desire for some specific sensory experience previously unknown to us. Further if we lack any beliefs about what it is like to have some sensory experience *x*, we cannot have any desires toward it. One must plausibly know something about an object in order to desire it. If these points are correct then it doesn't seem that the Desire Account can accommodate pleasant feelings arising from novel experiences.

It seems to me that the Desire Account can avoid this unpalatable conclusion however. Whilst it is the case that we cannot possess a final *de re* desire for those sensations of which we lack experience, it could be the case that we are pre-disposed to form desires for certain sensations if we were to experience them. Thus, though I currently lack any desire for the taste of a quince (having never tasted one and so lacking knowledge of what they taste like), were

it the case that I was to experience the taste of a quince, I would immediately form a desire for this sensation, which would then be satisfied and result in pleasant feeling. In fact this seems to be how most desires come about in us. We are likely not born with any actual desires, since we are not plausibly born with any beliefs about what sensations feel like, but merely pre-dispositions towards sensations so that once we experience them we will form desires towards them (or against them).

Therefore overall it appears that the Desire Account, unlike the Felt Quality Accounts, can explain how affective feelings and sensory experiences come apart and how conditions such as pain asymbolia and PSAS occur. As such the Desire Account has intuitive appeal over the Felt Quality Accounts. Further despite the apparently backwards explanatory claim that we find things pleasant because we desire them, the account can stand up to scrutiny and can fend off the objections presented here. Therefore it seems to me that the Desire Account is the best account of sensory pleasant feelings available to us.

V. The Nature of Desires

If the Desire Account is the best way to understand the capacity for sentience then sentient beings must possess desires. However, if I am to show that sentient beings are agents I must give an account of the nature of desires and how desires feature in motivation. This will be the focus of this section. I will argue that assuming the Humean Theory of Motivation is correct, desires necessarily motivate us to act and thus sentient beings must necessarily be capable of motivation. I will adopt the Humean Theory of Motivation because it is intuitive and has widespread appeal. However one should note that if one rejects Humean Theory of Motivation this does not mean that sentient beings must not be agents but only that my argument here will be unsuccessful.

On the Humean Theory of Motivation desires are mental states that have a world-to-mind direction of fit. This means that desires are states that try to make the world reflect (or 'fit with') what is in one's mind. Desires can be contrasted with beliefs which have the opposite direction of fit: beliefs try to make the content of one's mind represent the world. This distinction can be understood another way. To use Smith's words; a desire that p tends

to continue to exist in the presence of a perception that not p.⁵¹ However, a belief that p tends to go out of existence in the presence of a perception that not p.⁵²

Other than being a mental state with a certain direction of fit, I take desires to be representational in that they have some content; a desire is always a desire for some object. Although they need not involve conceptual representation. Thus the content of a desire may be nonconceptual.⁵³ I take it that, like desires, beliefs may be occurrent or dispositional. Thus one may have a desire that one is not currently thinking about, and one may forget one has such a desire until the relevant object is available, or likely to be available, for example. Related to this, desires are also the kind of mental state that can, but need not, have a phenomenology, i.e. it can feel like something to desire the company of a friend. Although it is the case that we can also have a desire without it having any feeling whatsoever.⁵⁴ Finally, desires have a close relation to motivation. This motivating element is most commonly (and intuitively) explained by taking desires to be motivating mental states, that is to say, mental states that (when combined with an appropriate belief) constitute a motivating reason for one to do something. This at least is how the Humean Theory of Motivation (HTM) understands desires.⁵⁵

Of these claims about desires the final claim regarding motivation is in need of the most defence. According to HTM, being motivated to act, or having a motivating reason to act, amounts to having a desire and an appropriate belief. This is normally taken to be a means/end belief.⁵⁶ Thus if I desire the taste of an apple, in order to be motivated to walk to the fruit bowl, pick up and take a bite of an apple, I need a belief that will inform me that if I engage in this action I will achieve the object I desire; I will experience the taste of an apple. The belief then is the conduit through which the motivation travels. The belief ensures that the motivation is being used to reach the right goal.

⁵¹ Micheal Smith, *Moral Problem* (Oxford: John Wiley & Sons, 1994), 115.

⁵² I will discuss the nature of belief further in Chapter 4.

⁵³ I will argue for the possibility and plausibility of nonconceptual content in Chapter 4. The discussion in this section is focused on beliefs but the points made on specification and content are applicable to mental states generally.

⁵⁴ Jonathan Dancy, *Practical Reality* (Oxford: Oxford University Press, 2002), 89; Brandt, *A Theory of the Good and the Right*, 30.

⁵⁵ Smith, *Moral Problem*, 92.

⁵⁶ One should not take a 'means/end belief' to necessary to form a motivating reason however. One does not need a sophisticated belief about causal or metaphysical principles involving the concept of desire satisfaction. A belief that B (normally) occurs before A occurs seems to be a belief which, when combined with a desire for A, will give one a motivating reason. Such a belief could be gleaned merely through experience of the world. For instance, through watching the sequence of traffic lights one could come to believe that red follows amber, regardless of whether or not one understands what traffic lights are, their significance in the Highway Code or the laws of physics which govern their operation.

Assuming desires are understood to be as described, they clearly have a motivational element. However, if I am to show that sentient beings are necessarily agents then I will need to show not only that desires can have a motivational element, but that they are necessarily motivating. I will argue that in fact it is *necessarily* true that if one has a desire to ϕ , one is motivated to ϕ .

The claim that having a desire means that one is motivated is at least *prima facie* intuitive I take it. When we desire something it does seem we are also motivated to pursue it. Further it seems strange to imagine a case where this doesn't occur. How could one desire to drink a glass of water but not be motivated to do so? If one did claim that one had such a desire but was not motivated to satisfy it *at all*, this seems sufficiently far removed from the notion of desire for us to question whether one really had a desire, and not some other mental state. The 'at all' here is important.

One can of course possess a desire but not be *sufficiently* motivated to act on it. For instance, one could possess a desire for a glass of water but not be sufficiently motivated to go and get a glass of water because one is sat comfortably and enjoying a good film. This isn't a problem for the claim that desires are necessarily motivating however. Such an individual would still be motivated to get a glass of water because this desire would, under different circumstances, bring about the action of getting a glass of water. Having a desire entails being motivated *to some extent* but being motivated need not result in one acting to satisfy one's desire. The motivation that desires provide is defeasible and may be stifled from bringing about action by a stronger counter motivation.

i. Desires without Motivation

There are some more troubling counter-examples for this account of desires. Desires about states of affairs we cannot bring about, desires about the past, desires about that which one cannot change and desires about the logically impossible all seem to pose a problem for any account which suggests that desires are necessarily motivating. For example, the desire that my son got home safely last night, the desire that my medical test results don't come back positive or the desire that the answer I have given in my maths exam is the correct answer.⁵⁷ The desired outcome in these cases is something which we cannot possibly control. There is

⁵⁷ Dancy, *Practical Reality*, 88.

no action available to us that could bring about the desired outcome. Thus these accounts seem to be instances of desires which do not motivate us to act.

This conclusion seems to be too quick though. Just because there is no course of action available to one to bring about the object of one's desire, this doesn't entail that one is not motivated. We could very well be motivated in all these cases but simply unable to use this motivation since we cannot see any available actions to which we should apply our motivation. What may be occurring in these instances then is not that these desires are not motivating us, but that we lack an appropriate belief regarding the object of our desire. Thus we cannot form a motivating reason. Just like how one may desire to be successful yet lack knowledge of how to go about achieving this end. The desire to have given the right answer in my maths exam motivates me but I lack a belief to give this motivation a target. Though the explanation for why we lack an appropriate belief might be different, such cases are otherwise no different.

Further, one should note that this is supported by the fact that possession of these desires is often accompanied with a feeling of frustration.⁵⁸ This frustration is plausibly understood to be present because we *are* motivated by our desires but we lack an appropriate outlet through which to make use of our motivation. If we deny that such desires are motivating this frustration is difficult to explain.

Additionally, though one cannot pursue the object of these desires themselves, one may pursue related objects. For instance, taking Dancy's example, if one has a desire to die peacefully in one's bed, it seems plausible to say that one is motivated by this desire when one acts in ways that are conducive to this end or make the probability of it occurring more likely.⁵⁹ For instance, one may notify one's family of one's wishes, refuse to leave one's bed when ill, etc. Similarly, if one has a desire to be successful but does not know how to satisfy this desire, one may attempt to find out how to become successful. Plausibly such actions would be motivated by one's original belief to become successful. Since we would expect individuals with these desires to act in these ways, it seems fair to suggest that even desires which we cannot directly pursue are motivating.

One should note that I am not here claiming that some or any of these features necessarily obtain in cases where one lacks a relevant belief. It may be the case that one has a

⁵⁸ Ibid.; Brandt, *A Theory of the Good and the Right*, 26.

⁵⁹ Dancy, *Practical Reality*, 88.

desire that one's son got home safely last night and yet one feels no frustration, or that one has a desire to die peacefully in one's own bed but does nothing which is conducive to this end. The argument here is merely that these features are marks of motivation and their presence in a great many examples gives us sufficient reason to believe that desires are necessarily motivational.

ii. Non-Actable Desires

A more difficult counter-example for the account of desire I have put forward here is the case of *non-actable-desires*. Such desires have as their object, some a state of affairs which must be brought by someone or something other than oneself. Thus in principle one cannot do anything to bring about the object of one's desire. If any desires are going to lack a motivational element it seems like it will be these. Strawson's weather watchers famously have such desires.

The weather watchers, Strawson tells us, are nonsentient, inanimate beings that desire certain states of affairs to occur but do not desire that they themselves bring about such states of affairs (and in any case are entirely unable to bring about such states of affairs). Strawson gives us the example of a weather watcher watching a seed land in a crag on a riverbed and forming the desire for the seed to grow into a tree, but wanting this to occur without any intervention from itself.⁶⁰ Having such a desire then would apparently involve no motivation to satisfy one's desire, to make the satisfaction of one's desire more likely or to act in ways conducive to satisfying one's desire. Also consider a more everyday example of non-actable-desires. Harry, fancies Paula, and has a desire that Paula fancy him back, however Harry desires that Paula fancy him without him doing anything to influence her.⁶¹ Attila Tanyi, who gives this example, claims that Harry is not motivated to satisfy his desire or act in any ways that are conducive to satisfying it. Much like the weather watchers he has a non-actable desire. Thus non-actable desires present a clear problem for the claim that desires are necessarily motivating.⁶²

⁶⁰ Galen Strawson, *Mental Reality* (Cambridge, Massachusetts: MIT Press, 1994), 255.

⁶¹ Attila Tanyi, "Pure Cognitivism and Beyond," *Acta Analytica* 29, no. 3 (2014): 337.

⁶² Tanyi's target here is Pure Cognitivism but the case she presents is as much of a problem for the claim that desires necessarily motivate.

I'll consider Harry's case since it appears to be the more difficult of the two. One might argue, as with the cases discussed in the previous objection, that Harry is likely to feel frustrated and this is evidence that he *is* motivated to act but merely cannot find an outlet for his motivation. Not only is this possible but it seems plausible and highly likely; most of us know from personal experience that unrequited love is incredibly frustrating, even when we don't restrict ourselves from influencing the person of our affections! Harry's desire it seems can be straightforwardly dealt with, he is motivated after all. Tanyi however, argues that this reply is not sufficient. In order to show that desires are necessarily motivating it must not merely be likely that Harry is frustrated but *necessarily* the case. It must be impossible for Harry to possess his desire without a corresponding feeling of frustration. However Tanyi claims, it is clearly possible that Harry could have such a desire and yet not be frustrated. Therefore, if we assume that Harry possesses the desire for Paula to fancy him, and is not frustrated by his inability to have a hand in satisfying this desire, albeit as unlikely as this is, we have reason to doubt the claim that desires necessarily motivate.⁶³

There are a few things to say in reply to Tanyi. Firstly one should note feeling frustration is not a logical implication of being motivated and lacking an outlet for the motivation. As I stated at the start of this section, desires *can* have a specific phenomenology, but they need not. The fact that we can imagine Harry to lack frustration isn't conclusive proof that he must lack motivation. It could be that case that his motivation has no outlet and for whatever reason this hasn't manifested into a feeling of frustration. Perhaps he just has zen-like calm and simply accepts when he is unable to satisfy his desires. Of course this doesn't prove that Harry's desire *is* motivating either but it does at least rebut the objection that Harry's desire is not motivating because Harry is not frustrated. Admittedly however it would be strange to claim that Harry was motivated but not frustrated. This desire seems to be just the kind of instance in which one would get frustrated. This leads to the second point.

Desire and motivation are closely linked in most theories of desire, and in the common sense notion of desire. Further, in Harry's case more than most, it seems that he would be motivated. It seems not just unlikely that one could have a desire like Harry's and not be frustrated by one's inability to satisfy it, but it is bordering on unbelievable in the case of any normal human. It seems so unbelievable in fact that were we to meet Harry and he told us

⁶³ Tanyi, "Pure Cognitivism and Beyond," 338.

about his desire and lack of frustration we may be inclined to doubt that the mental state that Harry possessed was actually a desire.

One might retort that this is merely question begging; that desires are motivating is the point at issue here so I cannot simply state that because Harry isn't motivated he doesn't possess a desire. Indeed this argument would be question begging, but this is not my claim. My argument is that Harry's desire is already far removed from paradigmatic cases of desires because he specifically desires to have no part in bringing about the object of his desire. If it turns out that there is no motivation involved either, then this desire seems far removed not just from the notion of a desire on my account, but also from the notion of desire more generally. Whatever mental state Harry possesses seems less like what most of us would call a desire and more like another mental state such as a hope, wish and dream.⁶⁴ Since the example is in tension with the more general understanding of desires we should be sceptical that such a state is a *bone fide* desire. Whilst there may be examples of desires in which there appears to be no motivation, this one simply seems too far removed from our general understanding of a desire to be a plausible instance of such a desire.

My answer to Harry's case is the most intuitive to me, however I see that others may not share my conviction. If this is the case it need not be detrimental to my main argument in this chapter. If Harry's case does give us an example of a genuine desire, then it would seem that non-actable desires may not be necessarily motivating (though may be still be motivating in a great many cases). However it would still seem to be the case that actable desires *are* necessarily motivating. As such, so long as merely sentient beings possess some actable desires, they will possess motivation for action. Furthermore, it seems that merely sentient beings would possess such desires. Firstly unactable desires are more cognitively complex than actable desires and thus if any beings possess them it seems unlikely that it will be minimally sentient beings. Secondly, as argued above, sentient beings must possess some final *de re* desires for sensory experiences. These desires are actable and thus every sentient being will possess at least some desires that are necessarily motivating.

⁶⁴ I do not have a philosophical account of wishes, hopes or dreams in mind here but am relying merely upon a commonsense understanding of these terms. Tanyi argues that Harry's mental state should not be understood as a hope or a wish, however her remarks are brief and unconvincing. She states that since the object of wishes need not be logically possible, unlike the object of desires, Harry must have a desire not a wish (because the object of Harry's mental state is logically possible). This clearly doesn't follow. On Tanyi's own account, we can have wishes about logically possible things. What's more it seems that we do have such wishes: a wish for world peace for example. *Ibid.*, 343–44.

Conclusion

Therefore it seems that sentient beings possess desires and these desires must necessarily be motivating. Thus, sentient beings are beings that are capable of possessing motivation to act. I argued that this is the case because the Desire-Based Account of Pleasant Sensory Feelings is the most plausible and intuitive account of the ability to experience sensory pleasant feelings. I then argued for understanding desires as necessarily motivating, arguing against the possibility of non-actable desires (at least in sentient beings). Now I have established that sentient beings are necessarily capable of possessing motivational states, I will in the next chapter argue that this entails that they are agents.

Chapter 4

From Desires to Action



So far I have argued that being capable of feeling pleasant and unpleasant affective feelings, that is, being sentient, requires that one possess desires. This is because pleasant sensory feelings are a result of the satisfaction of *de re* final desires for sensory experiences. Further, desires (or at least these desires) are necessarily motivating. So thus far, I have shown that being sentient entails that one has desires. However possessing desires alone isn't enough for agency. In order to be an agent one must be able to act intentionally. Hence over the course of this chapter, I will establish that sentient beings are able to act for motivating reasons and therefore are agents.

First, I will defend the claim that merely sentient beings possess beliefs by arguing for an understanding of beliefs that includes certain nonconceptual representational states. I will then argue that such states grant sentient beings the cognitive resources to possess motivating reasons. Next I will suggest that acting for a motivating reason is both sufficient for acting intentionally and within the capabilities of merely sentient beings. Finally, I will argue that mental actions can sometimes be actions that are done for a reason and thus that agents needn't possess a body or the capacity for bodily action. I will end the chapter by making some other general claims about what it means to be an agent, thus concluding my argument that the possession of sentience entails the possession of agency.

I. The Nature of Beliefs

To recapitulate a little, I argued in Chapter 3 (sec. V) that having a motivating reason means having a desire for some object and a relevant belief regarding how one will satisfy one's desire. Thus a being's motivating reason to have a drink might be constituted by a desire for the taste of juice and a belief that by going to the fridge one will achieve this end. So far I have shown the sentient beings can possess the relevant kind of desires to form motivating reasons but I have not shown that merely sentient beings possess beliefs. I will now turn to this task.

I think a rather simple argument shows that sentient beings must be capable of possessing beliefs. Desires seem to be more complex mental states than beliefs. Desires are mental states with representational and motivational components. They are not only about something, but they also present it to one in a certain way and motivate one to pursue it. Thus a desire is a state that represents some state of affairs e.g. the tasting of an apple, but also presents it to one as desirable, something that one wants. Beliefs however, do not have this motivational component.¹ A belief is a mental state that merely represents some state of affairs e.g. that this is an apple. As such possessing desires seems to require the possession of more complex cognitive capacities than possessing beliefs. If this is the case, assuming that sentient beings can possess desires, they must also be capable of possessing beliefs.

However, there is widespread scepticism among philosophers and scientists regarding non-human animals' possession of beliefs and as such many will find this argument unsatisfactory. One may accept that the possession of desires entails the possession of beliefs but argue that this isn't sufficient grounds for the conclusion that sentient beings must have beliefs. In fact, the sceptic may argue, that sentient beings possess beliefs is a *reductio ad absurdum* of my view and that, it is so unlikely that merely sentient beings possess beliefs that we should conclude that the Desire Account of affective feelings must be faulty and sentient beings cannot possess desires *or* beliefs.

In response to this criticism, I will spend the rest of this section giving an independent argument to show that animals are capable of beliefs. Importantly my argument in this section will not show that the possession of beliefs necessarily follows from the possession of sentience. Instead I will draw upon empirical evidence to show that many sentient animals possess beliefs and consider two popular objections to this claim. Through doing this I aim to make the conclusion of my original argument, that all sentient beings necessarily possess beliefs, more convincing.

As I stated in Chapter 3 (sec. V), on the Humean Theory of Motivation beliefs are taken to be mental states that have a mind-to-world direction of fit. They are representational mental states that try to capture or reflect what the world is like. Beliefs are usually also taken to represent the world through propositional content. This means that they are representations constructed from concepts. Finally, and most important for my argument, beliefs have a

¹ At least most beliefs don't have this component. Moral beliefs may be an exception here but this shouldn't be a concern for my argument.

functional role in action. When appropriate beliefs and desires are combined they form motivating reasons which can motivate action.

Various kinds of beliefs can be combined with desires to form motivating reasons. Often a means/end beliefs is said to be necessary to form a motivating reason however it seems that simpler kinds of belief may suffice. One's motivating reason may be partly constituted by a belief about the causal relations that hold between the object of one's desires and various other objects, for instance, the belief 'that turning on the taps causes the water to flow'. However one's belief need not be so complex. One need not understand the underlying causal relations or reasons for events occurring; beliefs about the mere order or sequence of events in the world are often sufficient. For instance, a much simpler belief such as the belief 'that water flows after turning the taps' would be sufficient, when combined with a desire to drink water, to form a motivating reason to turn on the taps.

I will draw on several empirical studies to show that many different species of animals are capable of possessing representational mental states that can be combined with appropriate desires to form motivating reasons. I will then consider two objections: *the Argument from Specification* which contends that these mental states cannot be beliefs because we cannot specify their content, and *the Argument from Concepts* which contends that that these cannot be beliefs because beliefs are composed of concepts and animals cannot possess concepts.

I will argue that whilst the mental states of animals may not be specifiable in the way many human beliefs are, and they may not be conceptual, they nonetheless fulfil the functional role of beliefs and can plausibly be combined with desires to form motivating reasons. As such they ought to be understood as beliefs. However, if one is committed to the more orthodox notion of belief, then it is sufficient for my purposes that these mental states are belief-like in that they are representational states that can be combined with desires to form motivating reasons. That sentient beings are capable of possessing these mental states is all I need for my argument to show that merely sentient beings can plausibly possess motivating reasons.

i. Action-Guiding Beliefs

It seems sensible, if not intuitive to think that sentient animals possess beliefs. Considering that that we share an evolutionary history with animals, as well as being anatomically and physiologically similar to animals, it appears that we are probably at least at a fundamental level, psychologically similar to them. Since beliefs are a pretty basic part of our psychology, if animals are psychologically similar to us then they are likely to possess beliefs. Furthermore, animals engage in various kinds of seemingly complex behaviour such as hunting, hiding, navigating, nest building, avoiding danger, attracting a mate, etc. Since they are beings with minds, a plausible explanation of how they are able to do these things should feature beliefs, or at least some kind of belief-like representations. Thus *Prima facie*, it seems that animals possess beliefs.

More than an argument from analogy and an inference to the best explanation can be given in support of animals' beliefs though. We can strengthen these arguments by appealing to empirical literature. Empirical studies give us some specific examples of mental representations that, when paired with certain desires, can form motivating reasons. Thus by looking at the empirical studies we can see not just that sentient animals can possess beliefs *per se*, but that they can possess the kind of beliefs that are necessary for agency.

Many experiments, as well as cognitive ethnological evidence, provide support for the claim that animals possess representations that can be combined with desires to form motivating reasons. Firstly consider an experiment in which a chimpanzee competes with a human for food.² Food is placed in a glass booth with the human. There are two tunnels through which a chimp can put their hand to reach the food, one is clear and the other opaque. There is also an opaque divider present around the opaque tube such that the human cannot see if the chimp is approaching the opaque tube. If the human sees the chimp approaching they will snatch the food away before the chimp can reach it. Chimps in this experiment had a preference for reaching through the opaque tube rather than the clear tube, apparently aware of the fact that they will be seen reaching through the clear tube (and thus lose out on the food).

A variation on this experiment was also completed in which the same booth was used but the opaque tube was changed for another clear tube, and doors were fitted on the end of

² Brian Hare, Josep Call, and Michael Tomasello, "Chimpanzees Deceive a Human Competitor by Hiding," *Cognition* 101 3 (2006): 495-514.

both tubes.³ One door was noisy and the other silent. The human participant would look away during the experiment but would turn back if they heard the noisy door and take the food before the chimp could reach it. Chimps in this trial preferentially chose the silent door, implying that they were aware of what the human would hear if they were to make a noise. In both these experiments chimps used the opaque tube/silent door from the first trial i.e. before experiencing a negative result. In similar experiments dogs have also been shown to preference a silent method of procuring food over a noisy one when there is a risk of alerting a human.⁴

Other experiments show that chimpanzees will avoid collecting food which they think a dominant can see, and preference collecting food they think a dominant cannot see (because there is a physical obstruction between the dominant and the food).⁵ In fact, they may even try to misdirect a dominant con-specific away from food.⁶ These behaviours are not unique to mammals. Scrub-Jays, famous for caching foodstuffs and retrieving them months later, are also aware of others' attention towards their food. If there are conspecifics present they will wait until the onlookers are distracted before caching food, or cache their food behind some kind of barrier. If they notice that they are being watched whilst already engaging in caching food they will return later to move the food to a new location.⁷

In all of these cases we must attribute some representation to the animals involved. More importantly however, it appears that such representations could plausibly, when paired with appropriate desires, form motivating reasons. In the first case, the chimp appears to believe something like 'the human will take the food when they see me' and 'the human cannot see me here'. These representations combined with a desire to eat seem to give the chimp a motivating reason to approach the booth from the occluded side. Much the same can be said for any of the cases here. Whether the individual is a primate, dog or bird it seems that

³ Alicia P. Melis, Josep Call, and Michael Tomasello, "Chimpanzees (Pan Troglodytes) Conceal Visual and Auditory Information From Others," *Journal of Comparative Psychology* 120 (2006): 154–62.

⁴ Shannon Kundery et al., "Domesticated Dogs (Canis Familiaris) React to What Others Can and Cannot Hear," *Applied Animal Behaviour Science* 126, no. 1 (2010): 45–50.

⁵ Michael Tomasello, Josep Call, and Brian Hare, "Chimpanzees Understand Psychological States – the Question Is Which Ones and to What Extent," *Trends in Cognitive Sciences* 7, no. 4 (2003): 153–56.

⁶ E. W. Menzel, "A Group of Young Chimpanzees in a One-Acre Field," *Behavior of Nonhuman Primates* 5 (1974): 134–35.

⁷ Joanna M Dally, Nathan J Emery, and Nicola S Clayton, "Cache Protection Strategies by Western Scrub-Jays (Aphelocoma Californica): Hiding Food in the Shade.," *Proceedings of the Royal Society B: Biological Sciences* 271, no. Suppl 6 (2004): 387–90.

they plausibly possess a representation (or representations) that can be combined with an appropriate desire to form a motivating reason.

One may accept that some animals are capable of these kinds of mental states but nonetheless be sceptical that merely sentient beings could possess beliefs like this. However in addition to such examples, there are more simple kinds of representations that are possessed by many different kinds of animals. These mental states, I will argue, should also be understood to be beliefs, and are more plausibly the kind of representations that merely sentient beings could possess.

One example of a simpler, or less cognitively demanding representational mental state is a cognitive map. Cognitive maps are a suggested way in which individuals represent information about their environment, specifically where objects are and how to navigate from one place to another. Cognitive maps may be understood to be pictographic representations of an environment much like their real-life namesake, though this need not be the case. There are different accounts of the exact structure of these representations but a general definition of cognitive maps is that they are mental representations that encode geographical information in a non-propositional format.⁸ Helpful to understanding the notion of cognitive maps is knowing what it is like to use one. Heck describes possessing a cognitive map as being able to find a location or object within the area represented by the map, but not being able to give directions to it (without first imagining the route one would take).⁹ One might also think of being able to quickly locate something in one's messy office but being unable to describe the location of the object to one's spouse over the phone.

There is evidence for a variety of animals possessing cognitive maps. Rats are one example. Rats possess the ability to reliably navigate to a specific location within a familiar environment even when they start in an unfamiliar position. In Morris' water maze experiments, rats were tasked with navigating to a platform in a pool of water. The liquid was made opaque with the addition of milk powder and the platform was slightly submerged so that the rat attempting to find the platform would not be able to locate the platform visually from a distance. After a few days of training, rats were able to navigate to the platform (using

⁸ Rescorla discusses various interpretations of cognitive maps here: Michael Rescorla, "Cognitive Maps and the Language of Thought," *The British Journal for the Philosophy of Science* 60, no. 2 (2009): 381.

⁹ Richard Heck, "Are There Different Kinds of Content?," in *Contemporary Debates in Philosophy of Mind*, ed. Brian P. McLaughlin and Jonathon D. Cohen (Oxford: Blackwell, 2007), 129.

an almost direct route) when started in a novel position.¹⁰ This is the case, it has been argued, because they ‘map’ an environment when they enter it so that when they return, even if they enter the environment from a different and unfamiliar point, they can orientate themselves and find their goal area.¹¹ They do this by forming representations of the relative distances between various landmarks of the area and the goal area/object (the platform in this case). These representations together form a cognitive map. This kind of representation isn’t unique to rats though. There is evidence that cognitive maps are possessed by cats.¹² Others have argued that snakes, lizards and turtles¹³ as well as fish¹⁴ make use of cognitive maps. Seemingly cognitive maps may also explain the ability of pigeons to ‘home’ and the ability of scrub jays, and other food-caching birds, to find previous food-caches, months after caching.¹⁵

Though cognitive maps are not propositional in format they seem to fulfil the functional role of beliefs. It seems that a cognitive map is the kind of representational state that could, when paired with an appropriate desire, form a motivating reason. For instance, consider the rat in the water maze test. If we assume that the rat has a desire to reach the platform in the pool (because she will get some reward), her cognitive map of the pool, which includes her current location as well as the location of the platform, appears sufficient for her to form a motivating reason to swim to the platform. Plausibly then, cognitive maps should be understood as beliefs. Moreover, if beliefs like this are possessed by animals as varied as reptiles, fish, mammals and birds, perhaps merely sentient beings could possess such beliefs.

There are many objections to the claim that animals possess beliefs. I do not have time or space to consider them all, and in any case not all of them are worth considering. However I will consider two objections which appear to be widely supported among those sceptical of animals’ beliefs and which I expect are most likely to be launched against my argument here. The first argues that in order to ascribe any beliefs to some individual we must be able to ascribe some specific belief(s) to this individual. This I call ‘the Argument from Specification’. The second argument contends that beliefs are propositional and thus constructed from concepts. Further having concepts is possible only if one possesses language. Thus since

¹⁰ R. Morris, “Developments of a Water-Maze Procedure for Studying Spatial Learning in the Rat,” *Journal of Neuroscience Methods* 11, no. 1 (1984): 47–60.

¹¹ Bruno Poucet, “Animal Navigation and Cognitive Maps,” 2006.

¹² Bruno Poucet, “Spatial Behaviour of Cats in Cue-Controlled Environments,” *The Quarterly Journal of Experimental Psychology*, 1985.

¹³ C. Broglio et al., “Hippocampal Pallium and Map-Like Memories through Vertebrate Evolution,” *Journal of Behavioral and Brain Science* 05, no. 03 (2015): 110.

¹⁴ *Ibid.*, 114; Victoria Braithwaite, *Do Fish Feel Pain?* (Oxford: Oxford University Press, 2010), 84–87.

¹⁵ Poucet, “Animal Navigation and Cognitive Maps,” 6.

animals lack language we cannot ascribe any beliefs to them. This I call 'the Argument from Concepts'.¹⁶

In opposition to these arguments I will claim: firstly that being unable to ascribe precise content to an individual's beliefs is not sufficient reason to doubt that they possess beliefs *at all* and secondly that neither language possession nor concept possession is a prerequisite for beliefs. I will also argue that even if one is reluctant to accept that these belief-like states of animals are *bone fide* beliefs, since they are representational states that fulfil the functional role of a belief, they are sufficient to form a motivating reason (when combined with an appropriate desire). Therefore possession of these belief-like states and appropriate desires, I will argue, are jointly sufficient for agency.

ii. The Argument from Specification

First I will consider the Argument from the Specification. The Argument from Specification states that in order to attribute a belief to an individual one must be able to specify the content of this belief, or its content must be 'expressible'.¹⁷ At first this might not seem to pose a problem for the ascription of beliefs to animals. In the case of the chimpanzee in the experiment with the glass booth for instance, I could ascribe the belief 'that the human will take the food when they see me'. However for the chimp to possess a belief with this content she would have to share my concepts of HUMAN, FOOD, and SEE or at least relevantly similar concepts.¹⁸ Yet it doesn't seem likely that any chimpanzee would possess these concepts.¹⁹

In order to possess the relevant concept, HUMAN, the chimp would need to have some knowledge of different kinds of species, for instance she may need to understand that chimps are not humans, nor are gorillas and understand that humans are mammals.²⁰ A

¹⁶ I focus on beliefs here since it is attributing beliefs to animals that most often attracts criticism. However these problems affect attribution of mental content to animals more generally. From what I say regarding beliefs it is hopefully clear how these problems can be dealt with more generally regarding the attribution of other kinds of mental content to animals.

¹⁷ This argument is famously supported by Davidson and Stich. Stich is less committed to the claim that animals lack beliefs but clearly supports the argument. Donald Davidson, "Rational Animals," *Dialectica* 36, no. 4 (1982): 317-27; Stephen P. Stich, "Do Animals Have Beliefs?," *Australasian Journal of Philosophy* 57, no. 1 (1979): 15-28.

¹⁸ Following convention I will use uppercase to denote concepts.

¹⁹ Stich, "Do Animals Have Beliefs?," 18.

²⁰ I will discuss the nature of concepts below when considering the Argument from Concepts. I ask my reader to hold off on criticism of use of concepts until then.

similar point can be made for the other concepts involved too. Unlike a human, these are the kinds of discriminations a chimpanzee likely doesn't make. So it seems highly unlikely that the content of the chimp's belief would involve the concept of HUMAN that I use when I believe 'that the human will take the food when they see me'. If the chimp has some belief about what will happen when the human sees her, this belief must have some other (albeit similar) content. However, unless we can specify this content, the Argument from Specification says, we cannot take her to have *any* such belief. Further, since this problem of specificity occurs with all proposed beliefs, defenders of the Argument from Specification conclude that in fact animals do not have any beliefs.

At first the problem of specifying the content of animal beliefs may seem easy to circumvent. We could just clarify that our content specifications are *de re* in nature. Thus when I state that the chimp believes 'that the human will take the food when they see me' I do not mean that the chimp has a belief about a human *as* a human, but that their belief picks out the object that I pick out by using the concept HUMAN. In other words, the chimp has a belief about the human under some description. The same could be said for any other concepts involved too. However in doing this we face another problem. Such a specification is indeterminate between different possible beliefs that the chimp could have. For instance, this does not tell us under what description the chimp possesses a belief about the human. The chimp could have beliefs about 'a hairless ape' or perhaps 'a tall and thin animal'. By using a *de re* ascription then we do not discriminate between different possible senses for the same referent. As such, the specified belief lacks the requisite 'finesse-of-grain'.

In cases such as these we say that the attribution of such a belief lacks *opacity*.²¹ Belief attributions of the form 'A believes that *x*' have opacity. This means that they do not allow for substitutions of referentially equivalent terms without a risk of changing the truth value of the belief. Thus taking the attribution 'Susan believes that the human will take her food', I change the truth value of this belief if I change the term 'human' for another term with the same referent such 'Sarah' or 'the 32 year old woman'. Whilst it may be true that Susan believes 'that the human will take her food', it is false that she believes 'that Sarah will take her food' or 'that the 32 year old woman will take her food' since Susan doesn't know that the human in question is called 'Sarah' or that she is 32 years old. Thus, the belief attribution 'Susan believes the human will take her food' has opacity because modifying the terms that

²¹ Davidson, "Rational Animals."

make up this attribution for other terms that pick out the same thing, could lead to the attribution of a false belief.

Thus specifying the content of animal beliefs is a genuine problem. Assuming animals have beliefs, these beliefs must have some determinate content yet using *de re* ascriptions which fail to differentiate between senses cannot plausibly capture the true content a belief. It seems there is no easy way to specify the content of animals' beliefs. Nonetheless, that we are unable to specify the precise content of any animal's specific beliefs, I think we should resist moving to the conclusion that animals lack beliefs altogether. After all, though we cannot give opaque belief attributions, we can it seems give more or less accurate approximations of the content of animals' beliefs. Furthermore we determine that they do not possess certain beliefs. As such it seems that we can determine that they have beliefs with some content, we simply cannot give a precise specification of this content.

DeGrazia brings out this problem with a thought experiment: Let us imagine that we discover Martians with seemingly human level intelligence. These Martians act in a way that leads us to believe that they have beliefs; many of their actions are understandable to us and they do the kind of things we would expect them to do were we in their position. However they have sensory organs and mental representational systems that are radically different from our own. These systems are so different that belief attributions in any human language is impossible. Should we conclude that Martians do not have beliefs? Of course not.²² We should merely conclude that though we cannot attribute specific beliefs to them, they likely still have beliefs with determinate content. The same should be said for animals.

Difficulty in determining the content of beliefs isn't only present when specifying content across species barriers either. We have trouble ascribing content to those whom we do not share a language or culture, and even our own children. Consider Peter Smith's example of the two year old child Zoe. Zoe states that her father teaches philosophy. Yet despite her testimony, she surely doesn't mean what I would mean if I was to express that my father teaches philosophy. Her understanding of what philosophy is at such a young age would be significantly impoverished compared to mine. Thus it wouldn't be accurate for me to attribute to her the belief that her father teaches philosophy. What does she believe then? Well, we don't know exactly. Nonetheless, though we cannot specify the precise content of

²² David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (Cambridge; New York: Cambridge University Press, 1996), 145.

Zoe's belief we can say *something* about her belief. For instance, she believes that her father teaches something called 'philosophy' Smith suggests.²³ We might also be able to say that she distinguishes philosophy from art, but not science, and that it is something that is engaged in by reading books not lifting weights. Though we are unable to specify the content of her belief, we can describe it; it's a belief about her father's work and its content is somewhat similar to the claim that 'my father teaches philosophy' in various ways *x*, *y* and *z*.²⁴

Thus it does not seem that one needs to be able to specify the content of a belief in order to ascribe beliefs to an individual. We can be unsure of an individual's belief content whilst still being reasonably sure that they do indeed have a belief about something.²⁵ For example in Zoe's case, we can be reasonably sure that she believes something about philosophy or at least that she believes something about the referent of the term 'philosophy'.²⁶ We should not be concerned if it is somewhat indeterminate which beliefs one possesses, this does not entail that the content of the beliefs itself is indeterminate. Of course if we want to ascertain not just whether an individual has beliefs, but which beliefs they have, it will be necessary to give precise content ascriptions. Importantly however such ascriptions aren't necessary to understand that an individual has beliefs *at all* or even (roughly) what they are about.

A relevant question in response to this argument is: If animals are capable of possessing beliefs, why are we so woefully unable to specify the content of their beliefs? Though our specification of the content of children's beliefs is vague, there is considerably more vagueness in our specification of the content of animals' beliefs. We seem to be able to list a determinate set of possible content ascriptions for children. However this does not seem to be something we are able to do for animals. Many find this disparity troubling. If animals possess beliefs just as human children do, then it seems that we ought to be able to specify them at least in principle.

There are two plausible potential answers here. Firstly animals may represent at least partly in a format (or formats) different to the format that humans predominately use (as I will argue in the next section).²⁷ Thus when giving a specification of the content of an animal's

²³ Peter Smith, "On Animal Beliefs," *The Southern Journal of Philosophy* 20, no. 4 (1982): 508.

²⁴ Andrews too makes the same point. Kristin Andrews, *The Animal Mind: An Introduction to the Philosophy of Animal Cognition* (New York: Routledge, 2014), 87.

²⁵ DeGrazia, *Taking Animals Seriously*, 155.

²⁶ *Ibid.*, 155–56; Smith, "On Animal Beliefs," 509.

²⁷ Jacob Beck, "Why We Can't Say What Animals Think," *Philosophical Psychology* 26, no. 4 (2013): 538.

belief, we are attempting to express something non-linguistic in a linguistic format. This is incredibly difficult if not impossible to do with any precision. Beck uses the example of trying to use language to precisely express the content of the Mona Lisa. Such an act cannot be done, Beck suggests, since the pictorial format of the Mona Lisa simply doesn't translate into words. Even if one were to specify the placement of every brushstroke, something would still be missing from the translation. It would still not capture the true content of the painting. Further even if it were possible, it would be incredibly difficult to produce a translation of such detail.²⁸

Secondly, even if the content of animal beliefs is in a format that we use, or could in principle be translated into a format that we use, such content is likely to be so unfamiliar to us that we cannot easily express it in language.²⁹ Animals have different ways of experiencing the world, different cognitive capacities and in some cases, senses that we cannot imagine having, such as echolocation and electroreception.³⁰ Further, animals are likely to represent the world very differently to us due to these capacities and senses, not to mention the radically different ways in which they live their lives and the background beliefs they have based on such experiences.³¹ Thus it should not surprise us that trying to specify the content of the representations of a non-linguistic being in language with any certainty is incredibly difficult.

Others suggest that our inadequacy in this area is because of a deeper difference between humans and animals: animal lack concepts. This is the argument I will address next.

iii. Argument from Concepts

Beliefs are commonly thought to be constructed from concepts. Further concepts are commonly thought to only be possessed by language users. Thus, because animals are non-linguistic beings they must not be able to possess beliefs since there is no possible content that these states could have.³² The possession of concepts is often taken to require the possession of language because of the Language of Thought Hypothesis. The Language of Thought Hypothesis suggests that thought must have a language-like structure. Specifically this means that the content of one's thoughts must have discreet parts that can be combined and

²⁸ Ibid., 536.

²⁹ Ibid., 533.

³⁰ Ibid.

³¹ Hans-Johann Glock, "Can Animals Judge?," *Dialectica* 64, no. 1 (2010): 26.

³² Davidson, "Rational Animals"; Stich, "Do Animals Have Beliefs?"

recombined according to grammatical (or grammar-like) rules to form different possible content.³³ These discrete parts are usually understood to be concepts. Thus the constituents of thoughts (concepts) should be able to be recombined according to certain rules to generate new thoughts, just as we are able to do with words to create new sentences.

Thoughts must possess this recombability, it is argued, because it is strange to suppose otherwise. Imagine I believe 'that the orange is smaller than the watermelon' I should also be able to believe 'that the watermelon is smaller than the orange'.³⁴ It would appear odd if one was able to form the first belief but not the second. Unless one could understand the second we would doubt that one truly understood the first. This, we think, is because these two sentences involve the same meaningful parts, merely combined in a different way.

According to the Language of Thought Hypothesis, if I am able to possess the belief 'that the orange is smaller than the watermelon' it is because I possess the concepts of ORANGE, WATERMELON and SMALLER THAN. This means not just that I understand that 'the orange is smaller than the watermelon' as a whole sentence, but that I understand the constituent parts in isolation. As such, it should be possible for me to rearrange them following the syntactical rules of language (or rules relevantly similar to these) to form new beliefs. If I am unable to do this, most would not be willing to ascribe the possession of any of these concepts to me or the belief that the orange is smaller than the watermelon. Without language it seems that our thoughts must lack a language-like syntax and thus cannot be composed of concepts. Therefore it has been argued if animals lack language then they cannot possess beliefs.

There are two broad lines of defence that have been argued against this position recently: first showing that concept possession does not require language possession and that animals do in fact possess concepts, and second showing that representational states can have content that is not constructed from concepts. I will argue in defence of both these claims. I will argue firstly that concept possession does not require language and thus that the content of beliefs can be conceptual but non-linguistic. I will then draw from empirical studies to show that there are cases in which animals can plausibly be said to possess concepts. Next I will argue that whilst many animals may possess conceptual capacities, representational states are not necessarily conceptual, and that many animals possess nonconceptual representational

³³ Gareth Evans, *The Varieties of Reference* (Oxford: Oxford University Press, 1989), 104; Jerry Fodor, *Psychosemantics: The Problem of Meaning in the Philosophy of Mind* (London: MIT Press, 1987), 134–38.

³⁴ Jerry Fodor and Zenon Pylyshyn, "Connectionism and Cognitive Architecture," *Cognition*, 1988, 39.

states. Further, such states nonetheless appear to be sufficient to form motivating reasons when combined with appropriate desires. Thus if minimally sentient beings possess states like these, and not full-blown propositional representations, they will still be agents.

One should firstly note that just because an individual lacks language it clearly does not follow that they lack thought with a language-like structure. Thus, the mere fact that animals lack language isn't sufficient to show that they lack beliefs. However, what reason do we have to believe that animals possess concepts? Arguing that animals possess concepts directly is difficult, not least because there is widespread disagreement over what exactly concepts are. There are several competing accounts and it would be beyond the scope of this section to consider whether animals can possess concepts in all of these senses.

Furthermore, it seems near impossible to gather empirical evidence for the claim that animals possess concepts in some of these senses. For instance, on the Classical Theory of Concepts, concepts are complex representations made up of simpler representations. Thus the concept ORANGE is composed of the simpler representations, citrus fruit, juicy flesh, sweet tangy taste, etc. In order to have the concept ORANGE one needs to be able to represent all of these simpler representations. Thus to possess the concept ORANGE one must be able to represent JUICY, SWEET, TANGY, etc. As Chater and Heyes note, one needs language to represent any of these simpler representations that make up our concepts and so it seems highly unlikely that animals could possess any of our concepts on the Classical Theory.³⁵ Of course, animals do not need to share *our* concepts in order to possess concepts *tout court*.³⁶ Nonetheless, if the Classical Theory of Concepts is correct and animals have their own concepts which are complex representations composed of simpler representations in their own languages (of languages of thought), since we lack these concepts, we will still not be able to gather any empirical data about which individuals possess concepts.

As such I propose investigating the possession of concepts in animals through an indirect route. Assuming the Language of Thought Hypothesis is true, we can investigate whether animal thought has the right kind of structure for animals to be possessors of concepts. If they are unable to recombine the content of an apparent representation in appropriate ways, then their thought lacks the appropriate structure to be conceptual and thus

³⁵ Nick Chater and Cecilia Heyes, "Animal Concepts: Content and Discontent," *Mind & Language* 9, no. 3 (1994): 214.

³⁶ I'll remain agnostic regarding whether we share any of our concepts with animals here. DeGrazia, *Taking Animals Seriously*, 155; Stich, "Do Animals Have Beliefs?," 22–24; Hans-Johann Glock, "Animals, Thoughts and Concepts," *Synthese* 123, no. 1 (2000): 54.

they cannot possess concepts. Further, if it is the case that they appear able to recombine the constituent parts of their representations in appropriate ways, then their representations must be composed of concepts and thus be *bone fide* beliefs.

Consider the experiments I discussed above involving chimpanzees. Can we say that the chimps in these experiments possessed representations that had the appropriate kind of structure? In the booth tests the chimps seem to be capable of believing that the human can't see them, as well as that the human can't hear them and the negations of these beliefs. Further in the experiments where chimps competed against conspecifics for food, they seem to have similar beliefs, in addition to beliefs that attribute social/dominance relations. Thus they appear to be able to represent different properties belonging to the same subject, for instance, that a particular human can see/hear them or that a particular human cannot see/hear them.

Further, they seem to represent different subjects possessing the same properties. They seem to be able to represent humans and conspecifics as not being able to see them and they can represent a conspecific as either dominant or subordinate. These experiments seem to show that there is a re-combinability to chimpanzee's representations. Moreover, this is the kind of re-combinability that we find in language. A chimpanzee is capable of representing not just that one can see them but that one cannot see them, they are capable of representing not just a conspecific as being able to see them, but also a human as able to see them. Assuming no counter evidence comes to light, there seems to be good reason to think that chimpanzees have a language-like structure to (at least some of) their thoughts. This being the case it seems that they likely possess concepts and so have *bone fide* beliefs.

However, primates are considered to be among the most cognitively complex animals. If I am to show that merely sentient beings are capable of beliefs it will not be enough to show that *some* animals possess concepts. Let's consider my other example involving cognitive maps. Cognitive maps, I argued are the kind of representations that mere sentient beings likely could possess. Thus if I can show that these kinds of representations are beliefs, then I will have given a good case for the plausibility of belief in merely sentient beings.

It doesn't seem that cognitive maps have a language-like structure in any obvious way. Heck and Recorla note that one cannot possess the negation of a cognitive map like one can of the belief that 'it is raining'. Neither can one possess the conjunct of two maps, as one can of the beliefs that 'it is raining' and that 'it is cold'. The structure of maps simply does not

allow for this kind of manipulation of content.³⁷ Further, it doesn't seem to be the case that one could even translate the content of a map into a format with a language-like structure.³⁸ The information in a map just cannot be given in sentences or in a sentence-like format. Of course parts of the content can be given in sentences, for example the library is 100 metres north of the town hall. However one couldn't capture all of the information contained in a map in a language-like format. This isn't just because maps are informationally dense. There doesn't appear to be one true description of any given map just as there doesn't appear to be a single true description of any given painting, even if we could give one in infinitesimal detail. Further it isn't clear that maps have any discrete elements. Again, much like a painting, you cannot break down a map into its basic constituents as you can language-like thought.

Certainly cognitive maps do have some kind of structure though. For instance, if on one's map the cabin is represented as over the river, and the wood is before the river, one can determine that the cabin is past the wood. Similarly if one heard that otters had been spotted at the river, one could determine that the otters were past the wood and before the cabin, and thus accurately navigate to the otters' position. There is a clear structure present. Cognitive maps also have versatile re-combinability too. One could in principle represent the otters' position as after the cabin or before the wood, or in any location which is not already occupied. Further if one revisited an area for which one possessed a cognitive map, one could update one's map by adjusting the placement of various elements on one's map. For instance, if one visited a building for which one had a map, and the furniture had moved around, one could update one's map accordingly.

Yet this re-combination of components seems too free. We don't think that concepts can be combined in *any* order to create new content. There are strict language-like rules. To return to my earlier example, one couldn't believe that 'the bigger is watermelon than the orange'. The concepts of WATERMELON and BIGGER THAN cannot be used in this way, a WATERMELON is an object concept and BIGGER THAN is a relational concept that holds between two or more objects. This kind of combination of concepts does not follow the sentence-like structure that the Language of Thought Hypothesis requires. The structure and rules that govern the re-combination of elements in maps are far removed from the structure and syntax of language. They are considerably more plastic. Plausibly then cognitive maps

³⁷ Heck, "Are There Different Kinds of Content?," 129; Rescorla, "Cognitive Maps and the Language of Thought," 380.

³⁸ Heck, "Are There Different Kinds of Content?," 129; Beck, "Why We Can't Say What Animals Think," 537.

are unlikely to have conceptual content at least according to any orthodox understanding of conceptual content.

Does this mean that cognitive maps are not genuine beliefs? I do not think so. As Camp notes, it isn't clear why the constituent parts of thought, or the rules that govern the recombination of these parts, must mirror the parts or syntax of language. Of course since humans are language users it is likely the case that the rules and parts of *our* thought for the most part mirror the parts and syntax of language. Nonetheless, it seems that the Language of Thought Hypothesis itself actually only establishes the weaker claim that representational states ought to be constructed from *some kind* of parts, that are recombinable according to *some* principles, in order to form new content.³⁹ Seemingly then, cognitive maps could have the requisite structure for *bone fide* beliefs. However they may not be conceptual. This needn't worry us particularly though. This would mean only that we are wrong to think that concepts are the necessary building blocks of thought.⁴⁰ Perhaps while concepts may constitute the essential elements of one format of representations, one may represent content non-conceptually. Many theorists have recently suggested that this is the case.⁴¹

One may insist that because such states aren't conceptual they cannot be 'beliefs proper' but at best mere 'proto-beliefs' (and indeed some do). However doing so will not under-cut my argument here. What matters for my argument is not whether we recognise that these states are genuine beliefs, but whether these states perform the same functional role as beliefs. It seems that when combined with an appropriate desire, such states *are* plausibly sufficient to form a motivating reason and thus motivate action. Therefore the fact that animals may lack concepts does not pose a problem for my argument that sentient beings possess agency.

My aim here was to fend off two arguments against the claim that merely sentient beings are capable of beliefs (and mental states generally). I have satisfactorily fulfilled this aim. I have shown that the Argument from Specification fails since we need not be able to

³⁹ Elisabeth Camp, "Thinking with Maps," *Philosophical Perspectives* 21, no. 1 (2007): 152.

⁴⁰ Whether in fact animal thought is best characterised as conceptual or non-conceptual is a question I cannot fully explore here, however it may be that my understanding of concepts is too narrow and that such representations could be classed as conceptual under some theory of concepts.

⁴¹ Jose Luis Bermudez, "The Distinction between Conceptual and Nonconceptual Content," in *Oxford Companion to the Mind* (Oxford: Oxford University Press, 1998); DeGrazia, *Taking Animals Seriously*, 151; J. Beck, "The Generality Constraint and the Structure of Thought," *Mind* 121, no. 483 (2012): 563–600; Adina L. Roskies, "A New Argument for Nonconceptual Content," *Philosophy and Phenomenological Research* 76, no. 3 (2008): 633–59; Heck, "Are There Different Kinds of Content?"

precisely specify the content of beliefs in order to conclude that an individual has any beliefs at all. What's more we often cannot specify the contents of humans' beliefs yet we do not doubt that they have beliefs at all. I have shown that the Argument from Concepts fails since animals can plausibly possess concepts in some situations and in other cases clearly have beliefs whether we understand them as having concepts or not. Thus I have shown that sentient animals can possess beliefs, and in doing, I have hopefully made plausible my claim that sentient beings necessarily possess beliefs, and thus the cognitive resources to form motivating reasons. I will now show that acting on motivating reasons is sufficient for intentional action.

II. Acting Intentionally

What counts as an intentional action is a hotly debated topic where there is no strong consensus among philosophers. Settling what counts as intentional action then would require much more work than I could possibly devote to the task here. However what is accepted by most philosophers of action, is that an action done for a reason is always an intentional action. This is not to say that actions done without a reason are unintentional. There are plenty of things we do for apparently no reason: doodling, kicking stones along the pavement, stroking our chins. These actions are not involuntary and most would say that these are intentional.⁴² I don't wish to deny this. Thus I claim here that acting for a reason is merely a sufficient condition for intentional action.⁴³ As such, I will argue that if sentient beings can be shown to be capable of acting for reasons, then they must be capable of intentional action, and thus must be agents.

i. Acting for a Reason

Assuming that sentient beings possess motivating reasons, what does it mean to act for such a reason?⁴⁴ Acting for a reason can occur in various ways. In the paradigmatic case of acting for a reason, one considers what one desires to do, considers the method through which

⁴² Some suggest that actions such as those described here can be understood as done for reasons and that acting for a reason is the sufficient and necessary condition for acting intentionally. Audi himself does not endorse this but suggests how such an account might look: Robert Audi, "Acting for Reasons," *The Philosophical Review* 95, no. 4 (1986).

⁴³ G. E. M. Anscombe, *Intention* (Basil Blackwell, 1966), 11.

⁴⁴ Henceforth, whenever I use the term 'reason' I mean 'motivating reason' unless I state otherwise.

to bring this about, and then puts one's plan into effect. Commonly we understand such a process as forming an intention and acting on that intention. We ordinarily act in this way when a situation requires careful thought, for example, when considering whether to apply for a new job or deciding what colour to redecorate the bedroom. However, there are a great many cases where we do not engage in conscious, thorough planning. Often we act spontaneously, reacting to a fast changing situation where there is no opportunity to engage in such planning. Nonetheless, we still seem to act intentionally and for a reason. Nobody wants to claim that when we make cup of tea we fail to act for a reason because we don't consciously deliberate over the methods available to fulfil our desire for a cup of tea, or because we didn't give any consideration to whether we *really* wanted a cup of tea.

Nomy Arpaly makes this point using the example of a tennis player: The accomplished tennis player doesn't have time to deliberate over her reasons when engaged in a tennis match. She has to react quickly to the changing situation. Yet it would seem strange to say she doesn't act for reasons. After all, we can say that she acted rationally or irrationally, praising or criticising her moves on the court. We can say 'that's a great shot' or 'why did she do that?' These statements seem to be claims about the reasons for which she acted. Further being an accomplished tennis player (rather than a mere beginner) means being able to respond to the changing situation on the court *in the right way*. In order to understand there being a better or worse way to play tennis, it seems we need to assume that one can act for good or bad (or better or worse) reasons.⁴⁵

Arpaly further goes on to note that if one thinks that acting for a reason requires conscious deliberation, then one will struggle to explain how deliberation itself is possible. Deliberation seems to be an action and, if actions require conscious deliberation, we must deliberate on whether to deliberate. However this would mean that we would also need to deliberate on whether to deliberate, on whether to deliberate and so on *ad infinitum*. One may be tempted to avoid the vicious regress here by arguing that deliberation itself isn't a genuine action or at least not an action in the same sense as bodily actions. As I suggest in section III of this chapter however, we shouldn't be so quick to dismiss certain consciously directed mental activities as actions. There isn't a clear or easy way to differentiate 'physical' from 'mental' action, furthermore conscious deliberation, as something that is seeming done for a

⁴⁵ Nomy Arpaly, *Unprincipled Virtue : An Inquiry Into Moral Agency* (Oxford: Oxford University Press, 2004), 52.

reason, seems as good a candidate as any for a 'mental action'. Hence it seems that one can act for a reason without consciously deliberating.⁴⁶

How can we know if an action is done for a reason? Following Anscombe, I take it that being able to give reference to a reason when asked 'why did you do that?' seems to be a good rule of thumb. If one acts for a reason then one's action has a rational explanation. Or to put this another way: one possesses a reason that can explain why one engaged (or is engaging) in said action. For instance the answer 'I fancied a cup of tea to drink and making some tea is a means to this end' provides a rational explanation for the action of drinking tea. In contrast, an answer such as 'it's involuntary' doesn't provide a rational explanation of the action of drinking tea, and thus (assuming the testimony is correct) shows that one did not act for a reason.⁴⁷ Importantly though, it isn't a requirement of acting for a reason that one be able to correctly state one's reason for action. As we have all likely had experience of, introspection is not completely reliable, humans are fallible and can be mistaken about their motivations. We can even be self-deceptive. More importantly, sometimes we don't actually have cognitive access to our reasons, even if we try to discover them.

As Arpaly argues, in the tennis player case, there were a lot of complex factors that were at play and the player didn't even consciously recognise what reasons she was acting for at the time, she just acted on them.⁴⁸ Further it is likely that the tennis player isn't able to reconstruct an account of the reasons she acted upon during her match once she has finished either. We can say the same about other actions too; driving a car whilst thinking about something else for example. Thus it seems not only that sometimes we don't explicitly notice the reasons for which we act, but we actually lack cognitive access to them. Hence one must be able to act for a reason even if no reasons feature in one's conscious thoughts before, during or after the action.⁴⁹

Thus, so long as a reason motivates me to act, I act for a reason; it doesn't matter whether I am able to consciously recognise the reason for which I am acting or not. In terms of cognitive abilities required to act for a reason, one need not be able to consciously recognise one's reasons for action. It seems that one merely needs to be able to appropriately respond to or be sensitive to reasons. And finally if one acts for a reason then it is the case that a rational

⁴⁶ Ibid., 57.

⁴⁷ Anscombe, *Intention*, 11.

⁴⁸ Arpaly, *Unprincipled Virtue*, 52.

⁴⁹ Audi, "Acting for Reasons," 520; Rowland Stout, *Action* (Routledge, 2005), 28.

explanation of one's action, could possibly be given (although not necessarily by the agent themselves or any other agent).⁵⁰ Such an explanation allows us make sense of one's action through appealing to the motivational role of one's motivating reasons.

One should note that this account of acting for a reason isn't causal. It is not the case, according to this account, that one necessarily acts for a reason when one's action is caused by a reason. If this were the case then this would make so-called deviant causal chain cases intentional actions. For instance, consider the day dreaming athlete: The athlete has a motivating reason to trip a fellow athlete, but can't bring himself to do so due to fear of being seen to be a bad sportsman. Lost in reflecting upon his reasons however, he falls and indeed does trip his fellow athlete. In this case his tripping his fellow athlete is caused by his motivating reason to trip his fellow athlete. Clearly however his action isn't intentional. The account of acting for a reason I have proposed here would not class such an action as an action done for a reason, since the athlete's response to his reason isn't rationally evaluable.⁵¹ This reason, or in fact any of his reasons, don't feature in a rational explanation of this action (even though the action is done because of the presence of the reason) because there is no rational explanation for this action, it was involuntary.

ii. Reason Super Blind-Sighted

Our ability to recognise reasons and give reason explanations is not necessary to act for a reason then. In fact, often we do neither of these things when acting for a reason. Rowlands makes the same claim regarding moral reasons in particular and gives the example of an individual with blindsight to illustrate his point.⁵² This example is, I think, worth considering here to illustrate how merely sentient beings could act for reasons.

Sufferers of blindsight (a condition caused by lesions on the visual cortex of the brain) claim to be unable to see objects placed in a particular area of their field of vision. However if asked to guess what the object is, they guess correctly at a rate significantly higher than chance. What this shows us is that phenomenal consciousness and functional abilities can

⁵⁰ 'Rational' in the weak sense that it is an explanation about rationality rather than a non-rational kind of explanation. I do not wish to claim that acting irrationally means that one is failing to act for a reason. An irrational explanation would still be sufficient to show that one acted for a reason (just that this was a bad reason for which to act on).

⁵¹ Audi, "Acting for Reasons," 517.

⁵² Mark Rowlands, *Can Animals Be Moral?* (Oxford: Oxford University Press, 2012), 165.

come apart. And specifically that being able to successfully identify objects doesn't require phenomenal experience of said objects.⁵³ Now consider the (theoretical) case of a super blindsighted person; an individual who entirely lacks *any* phenomenal visual experiences of the world around her. This individual, despite her impairment, is able to recognise objects when asked, catch objects that are tossed her way, and navigate her environment with the competence of a sighted person. In fact she can do any task that demands the use of sight as reliably as a sighted person.⁵⁴

We should further note that the super blindsighted person is not able to engage in these actions in virtue of some sixth sense. Neither is it the case that she can 'see' the world as a result of the heightening of her other senses like the superhero Daredevil. Just like sighted and blindsighted persons, she is able to engage in these activities because light bounces off of objects arounds her, hits her optic nerve and then information from the nerve is processed in her visual cortex.⁵⁵ Whilst such an individual would seem to lack visual consciousness, Rowlands suggests, it seems less plausible that she should not be considered a visual subject.⁵⁶ She can reliably carry out tasks that require visual recognition and does so through the same apparatus as a paradigmatic visual subject. Thus it seems as though we should recognise that she is a visual subject.

This case seems analogous to the case of the individuals who can act for reasons but lack the capacity to recognise, recall or explain their reasons for action. They may lack these capacities that paradigmatic agents have but nonetheless reliably respond to reasons. And, if we were able to peer inside their minds, we would see that there are reasons for (some of) their actions. Furthermore, these reasons move them to act in the way that they move a paradigmatic agent to act. Such 'reason super blindsighted' individuals possess a desire, a relevant belief regarding the desired object and are motivated to act in a given way because of these states. Thus they are capable of acting for a reason, though they are not be capable of various metacognitive functions involving reasons that typical human agents possess. I maintain that minimally sentient beings act in the same way as reason super blindsighted individuals. They meet the sufficient requirement for acting for a reason and thus are beings

⁵³ Ibid.

⁵⁴ Ibid., 166.

⁵⁵ Ibid.

⁵⁶ Ibid.

capable of acting for reasons, though they lack the ability to recall, scrutinise or recognise their reasons for action.

Through the next two sub-sections I will consider the relationship between agency and autonomy. In the first sub-section I will defend the claim that agents can lack autonomy and moral responsibility for their actions. In the second sub-section I will argue against the claim that autonomy is necessary to grant one control over one's motivations.

iii. Autonomy and Moral Agency

Autonomy is often taken to be a necessary condition for agency. Where does what I have said regarding agency leave autonomy then? If sentient creatures can act for reasons, are they autonomous? And if they are autonomous does this mean that they are moral agents and so morally responsible for their actions? If this were the case my argument would be grossly unpalatable. As such, I will need to explain the relation between the notions of agency, autonomy and moral agency in order for my account to maintain plausibility. I cannot hope to give a full account of autonomy, moral agency or moral responsibility here but I will attempt to go some way towards clarifying the relations between these notions.

It is not unpopular to think that agency requires autonomy (as Kant did for instance) and it is easy to see why this is thought to be the case. Autonomy is often understood as 'self-governance' or the ability to govern oneself, and agents must in some sense be able to govern themselves as this is just what it means to be an agent. However 'autonomy' has been used to mean many different things, by many different philosophers, sometimes ambiguously or inconsistently.⁵⁷ Some of the distinct ways in which the notion of autonomy has been understood include, but are not limited to, the following:

1) *Self-Control* - The ability to decide to follow one's motivational states, the ability to exercise some kind of self-control through using second-order desires/the ability to reflectively endorse actions through second-order desires.⁵⁸

⁵⁷ Dworkin notes the many problems in pulling together the various, somewhat disparate notions understood as 'autonomy' under one single capacity. Gerald Dworkin, *The Theory and Practice of Autonomy* (Cambridge: Cambridge University Press, 1988), 9–12.

⁵⁸ Arpaly refers to this 'Agent Autonomy', I will use 'Self-Control' to avoid confusion with agency itself and other notions of autonomy. Arpaly, *Unprincipled Virtue*, 118.

2) *Personal Efficacy/Independence* – The ability to get along well in the world without requiring the help of others.⁵⁹

3) *Normative Autonomy* – The ability to make one’s own decisions without the need for paternalistic intervention by others. One might think of this type of autonomy as the capacity that teenagers often bemoan isn’t being recognised by their parents.⁶⁰

4) *Authenticity* – Being free to act as one’s true self or acting in line with one’s most core values. Further, being able to have one’s actions genuinely attributable to oneself (or one’s ‘true/ideal’ self) rather than being partly attributed to someone/something else as might be the case if one is manipulated, coerced, confused, etc.,⁶¹

5) *Responsive to Reasons/Rational Autonomy* – Being capable of responding to reasons.⁶²

One can also be autonomous to different degrees; one can be more or less autonomous. For example with regard to ‘personal efficacy’, depending upon whether one is more or less independent in the world or, in the case of self-control, whether one is more or less controlled in one’s actions. One can also possess autonomy with regard to different domains, for instance one could possess ‘independence’ autonomy regarding physical actions and thus be capable of physically moving about in the world unaided, and yet lack ‘independence’ autonomy with regard to financial actions, requiring assistance in order to make financial decisions. So one might be said to have ‘independence’ autonomy but only with regard to one’s physical actions. Autonomy, then, is not an all or nothing property.

Whilst some theorists suggest that these various capacities that fall under the label autonomy are linked together or are different facets of the same capacity, I think we should be hesitant to accept this.⁶³ Sentient beings on my account are clearly autonomous in some of these senses, yet it is not clear that this implies that they are autonomous in *all* of these senses. For instance, what degree of personal efficacy one has will likely depend upon the body one

⁵⁹ Ibid., 119.

⁶⁰ Ibid., 120.

⁶¹ Ibid., 122; Darwall calls this ‘Agential Autonomy’ I will avoid this term in the aim of clarity. Stephen Darwall, “The Value of Autonomy and Autonomy of the Will*,” *Ethics* 116, no. 2 (2006): 165.

⁶² Arpaly simply refers to this as ‘Responsiveness to Reasons’ whereas Darwall opts for ‘Rational Autonomy’. I’ll stick to the former to avoid confusion with Kant’s notions of autonomy and rational. Arpaly, *Unprincipled Virtue*, 125; Darwall, “The Value of Autonomy and Autonomy of the Will*,” 165; Jeff Sebo, “Agency and Moral Status,” *Journal of Moral Philosophy* 14, no. 1 (2017): 15.

⁶³ Kant for instance, seemingly believed that all of these capacities mutually entail each other.

has and the world in which one lives. Further, one could have autonomy in the independence sense and yet completely lack autonomy in the authenticity sense.

If one possesses the minimally sufficient conditions for acting for a reason one must have autonomy in the sense of being responsive to reasons. One must also have personal efficacy to some extent. However, since animals do not plausibly possess the ability to make second-order decisions, minimally sentient beings won't possess 'self-control' autonomy. Further it seems an open question whether a minimally sufficient agent has autonomy understood as 'authenticity' or what degree of 'personal efficacy' they have.

Most importantly however mere agents are not autonomous in the 'self-control' sense of autonomy. This kind of autonomy is what is most often meant when the term autonomy is used in an unqualified sense and is the kind of autonomy that was central to the Kant's notion of autonomy (perhaps the most famous of all accounts of autonomy). What is more, it is this kind of autonomy, I will argue, that is linked with moral agency and moral responsibility.⁶⁴

How is 'self-control' autonomy linked to being a moral agent and morally responsibility then? To be morally responsible for one's actions plausibly requires that one be able to reflect upon and modify one's reasons for action.⁶⁵ After all, one cannot be morally responsible for acting on a reason if one does not have the capacity to evaluate the goodness or badness of acting on said reason. Since one would need the capacity to have second-order thoughts about one's reasons in order to do this, only those that are 'self-control' autonomous could be morally responsible for their actions. Thus, mere agents then cannot be morally responsible for their actions.

This does not mean that mere agents are unable to act in ways that result in, and even are motivated by, contributions to the well-being of others. Empirical evidence shows that many animals are capable of acting in apparently altruistic ways, and so mere agents may be capable of acting in similar ways.⁶⁶ This would not make them moral agents however. Being

⁶⁴ Arpaly notes that those beings that are moral agents and those that possess 'self-control autonomy are co-extensive. Arpaly, *Unprincipled Virtue*, 145.

⁶⁵ Sebo suggests a similar condition for moral responsibility. Sebo, "Agency and Moral Status," 16.

⁶⁶ Rachels gives the example of (unethical) tests that have shown some chimpanzees, monkeys and rats for example refuse food even when hungry if securing that food would lead to the electrocution of a conspecific. Further, no doubt the reader will have heard of at least one news story of dolphins saving humans stranded at sea or whales trapped in shallow water, or of a dog refusing to leave the side of their dying companion (human or canine) - Mark Rowlands gives several example in the opening page of his paper 'Moral Subjects'. Even some lizards have been documented to behave in altruistic ways, lying beside their dead (monogamous) partner after death for several days: James Rachels, "Do Animals Have a Right to Liberty?," in *Animal Rights and Human Obligations*, ed. Peter Singer and Tom Regan

a moral agent requires not merely that one is capable of acting on the right kind of reasons, but acting on these reasons *because* it is good to do so. Plausibly the person who acts in ways that they deem to be good, but has no understanding of what it means to be good or bad, cannot be said to be a moral agent. Therefore, since mere agents are unable to form second-order thoughts about their motivations, they are unable to consider or recognise the goodness or badness of their reasons.⁶⁷ Further, the notions of good and bad are complex and abstract concepts. Possessing these concepts may require the ability to use language. If this is the case then even some beings that possess self-control autonomy may not be moral agents. Mere agents cannot be said to be moral agents then, and cannot be morally responsible for their actions.

iv. Control without Autonomy

One might argue that though it is possible to act for a reason without being able to consciously consider or recognise one's reason for action in specific instances, acting for a reason requires that one at least has the *capabilities* to evaluate and adjust one's reasons. It is this metacognitive ability that gives one control and choice over one's action: the hallmarks of being an agent.⁶⁸ Without this ability one isn't *motivated* to act but merely *caused* to act.⁶⁹ Thus unless merely sentient beings possess the ability to recognise and modify their reasons i.e. 'self-control' autonomy (which they plausibly do not), they must at best be caused to act by their reasons and thus cannot be said to be agents.⁷⁰

This objection is an instance of what Rowlands calls '*the miracle-of-the-meta fallacy*'.⁷¹ This is the fallacy of assuming that metacognitive abilities or metacognition have miraculous powers.⁷² In this case, the specific reasoning error is assuming that having second-order

(Englewood Cliffs, N.J.: Pearson, 1989); Mark Rowlands, "Moral Subjects," in *The Routledge Handbook of the Philosophy of Animal Minds*, ed. Kristin Andrews and Jacob Beck (London: Routledge, 2017); Mark Fletcher, "The Mating Game" (BBC, 2013).

⁶⁷ However they may be able to act on moral reasons. Rowlands argues that some animals are 'moral subjects'; agents capable of being motivated by moral reasons but are not moral agents since they are unable to evaluate and reject or endorse acting on these reasons. I think that he is correct though it will not bear on my argument here. Rowlands, *Can Animals Be Moral?*

⁶⁸ Audi, "Acting for Reasons," 534.

⁶⁹ Joëlle Proust, "Rationality and Metacognition in Non-Human Animals," *Rational Animals?*, 2006, 268.

⁷⁰ Again this criticism parallels a criticism Rowlands considers regarding animals acting for moral reasons: Rowlands, *Can Animals Be Moral?*, 170.

⁷¹ *Ibid.*, 171.

⁷² *Ibid.*, 178.

thoughts about one's motivational states gives one control over one's motivational states that one would otherwise lack. This is a mistake since it overlooks that the supposed problem that occurs at the first-order level (a lack of control) will also occur at the second-order level.⁷³

What reason do we have to think that being able to evaluate our motivating reasons gives us control over them? The obvious answer that comes to mind is that this higher-order thinking allows us to stand back from our motivations, coolly consider them, and then make a choice. However, it isn't clear that this is the case. In possessing the ability to evaluate our reasons, we are able to make choices about our actions based upon considerations that individuals capable of mere first-order thoughts cannot. For instance, we might consider whether a given action is prudent, the most efficient available or morally permissible. However, why should we think that being able to do this grants us control over our first-order motivations? This could only be the case if the choices we make at a second-order are free themselves and this isn't obviously the case.

It might be that we pursue these second-order considerations with little introspective access to the motivating reasons for why we pursue them. So just as a reason super blindsighted individual merely appropriately responds to her reasons without being able to consciously recognise them, perhaps we do the very same at a second-order level.⁷⁴ If this is the case then our ability to evaluate and modify our reasons doesn't imbue us with any more control than reason super blindsighted individuals possess. The question of control merely moves one level up.

One may reply to this counter objection in one of two ways. One option is to accept that having higher-order considerations always grants one control over lower-order motivations, and argue that we have control over second-order deliberation because we have third-order considerations. However this is problematic since we will have to appeal to fourth-order considerations to explain our control over third-order considerations and so on *ad infinitum*. Needless to say, in accepting this explanation we will never succeed in satisfactorily explaining how we have control over our highest-order considerations and therefore will be unable to explain our control over even our first-order motivations.

Alternatively, one might argue that second-order considerations are just the kind of thing over which one has control in virtue of their nature. Thus we can stand back from

⁷³ Ibid., 186.

⁷⁴ Ibid., 154.

decisions that take place at a second-order level in a way that we can't when making a decision at first-order level. If this is the case, then we don't need any further higher-order thoughts in order to make a *bone fide* free choice and have control over our actions if we are capable of second-order deliberation. However, what is so special about second-order thoughts that could justify this claim? If this is true of decisions that take place at a second-order level, why shouldn't we take this to also be the case with decisions that take place at the first-order level? Unless one can say more about why second-order thought is so radically different from first-order thought, this line argument seems arbitrary and implausible.

Therefore, despite claims to the contrary, one need not be able to deliberate over one's motivating reasons in order to act for a reason. It is sufficient for intentional action that one is able to appropriately respond to reasons. Thus it is unproblematic that agents lack 'self-control' autonomy. However it should be noted that whilst the ability to evaluate and adjust one's reasons does not grant one any more control over one's actions than having first-order motivations, it does seem to allow one to consider a larger variety of different reasons. As Arpaly notes, though deliberation is not necessary for action in itself, it does seem to be necessary for certain kinds of actions, for instance those pertaining to philosophy, law, science, etc. and of course moral agency.⁷⁵ The possession of 'self-control' autonomy (more than other kinds of autonomy) does still seem to have significant moral import. As such, despite possessing autonomy in some senses, because mere sentient beings lack 'self-control' autonomy, I will refer to them as *non-autonomous agents*. This should help clearly separate the kinds of beings capable of higher-order thought and moral agency from those beings capable of merely acting intentionally.

One should remember that here I am attempting to show merely what capacities being sentient entails. As such, one should not take the mere agent discussed here to be a representation of any real or potential living being. Such an agent is merely a hypothetical model of an agent who fulfils the minimally sufficient conditions for agency: a being that possesses only sentience and the capacities that are entailed by this. Though it will not impact the argument here, I take it that many terrestrial animals may be capable of metacognition and different senses of autonomy to varying degrees. Though I take it that determining which animals have which capacities is incredibly difficult.⁷⁶

⁷⁵ Arpaly, *Unprincipled Virtue*, 63.

⁷⁶ See 'Do Animals Know What They Know?' for an account of some of the tests that are used to test for metacognition in non-linguistic creatures and support for the claim that some monkeys have

In this section I have shown that acting for a reason means that one's action can be rationally explained by reasons present in one's psychology. I have further shown that this alone is sufficient for being able to act for a reason since, despite how they appear to us, the metacognitive capacities often understood as 'self-control' autonomy do not offer us some special kind of control over our motivations that we would otherwise lack. Whilst most human agents have the ability to consciously consider, explain and modify their reasons for action, such capacities are superfluous to being able to act for a reason.

III. The Nature of Agency

So far I have shown that a sentient being, in virtue of having the capacity of sentience must have desires and beliefs. Thus they are able to possess motivating reasons that can explain their actions. Further they can act on these reasons, despite lacking any ability to recognise, reflect upon or modify their reasons for action. Therefore, sentient beings fulfil the sufficient condition for acting intentionally and are agents. However, we often assume that something more than this is necessary in order for one to be an agent. This something is a (functioning) body. Without a body, we commonly think, it doesn't matter how intellectually sophisticated a being is, they will not be able to act and thus not be an agent.

i. Agency without a Body

If beings that lack control over a body cannot be agents then this will present a problem for my account of moral status. Individuals with locked-in syndrome for example will appear to lack moral status on the Agency Account. This is implausible. It seems that despite their inability to move, such individuals still have interests which matter and thus we still owe them significant duties. As such, if I hope to defend a plausible account of moral status I must be able to explain why individuals that are unable to engage in physical actions, still possess moral status. I will do this by arguing that understanding agency to be limited to bodily actions, and actions that are constituted by bodily actions, is unnecessarily restrictive. There is no good reason, I suggest, to rule out some mental activities as *bona fide* actions and thus I will argue that being capable of mere 'mental agency' is sufficient to make one an agent.

metacognitive capacities: Sara J. Shettleworth and Jennifer E. Sutton, "Do Animals Know What They Know?," *Rational Animals*, 2006, 235–46.

A paralysed person with a fully functioning brain is ordinarily thought of as an individual that lacks agency. This is because actions are often thought to be things which we do with our bodies. But why should this be the case? One reason not to take mental activities to be actions is that it seems to lead to an infinite regress of explanation when we try to explain or understand our reasons for action.⁷⁷ For instance, if forming an intention is an action, then my intention to go to the shops cannot constitute a full explanation of my action of going to the shops. Since forming an intention itself is an action, to understand why I went to the shops we must also understand why I formed an intention to go to the shops. To explain this intention I might reference a belief I possess but then, since the formation of this belief is an action, the formation of this belief will also need explanation. So if we assume that mental activities are actions, we cannot explain our actions with reference to mental states. By doing so we merely explain our actions with reference to other actions. Thus an infinite regress looms and reason explanations of actions apparently become vacuous.

I am not sure that the claim that all mental activities are actions is as problematic as this argument makes out. In any case, whilst it seems unlikely that *all* mental activities are actions, I do think that *some* mental activities are actions. It may be problematic to take the forming of a belief or an intention to be actions but there seems to be good reason to understand certain mental activities as actions.⁷⁸ For instance, doing mental arithmetic, trying to remember one's friends' birthday or making a 'mental note' to book the cinema tickets when one gets home. All such activities seem very similar to bodily actions in that they can be voluntary and, at least apparently, can be done, and normally are done, for a reason. Unlike forming a belief, we can give a rational explanation for why we did these things.⁷⁹ We can give an answer to the question 'why did you do that?' that isn't 'I didn't realise I did' or 'it was involuntary'.

Another way in which these mental activities are similar to paradigmatic 'physical actions' is that they pass the 'Try Test'. Proust suggests the 'Try Test' as a rule of thumb for discerning whether some activity is an action rather than a mere behaviour. Every action is a successful trying, so if something can meaningfully be tried, it is likely an action. For instance

⁷⁷ Proust cites Ryle as the source of this objection: Joëlle Proust, "A Plea for Mental Acts," *Synthese* 129, no. 1 (2001): 212.

⁷⁸ Stout, *Action*, 145.

⁷⁹ Proust suggests understanding mental agency in this way (though ultimately rejects it); Joëlle Proust, "Mental Acts," in *A Companion to the Philosophy of Action*, ed. Timothy O'Connor and Constantine Sandis (Wiley-Blackwell, 2010), 209.

whilst I can try to walk, I cannot try to faint.⁸⁰ Using this test we can see that directed remembering as well as doing mental arithmetic and 'making a mental note' are all things one can try to do and as such are likely actions. This, combined with the fact that such actions can plausibly be done for reasons, suggests that some mental activities can plausibly be understood as actions. If this is correct, then theoretically, one need not possess a body in order to be an agent.⁸¹

One might point out that many mental activities appear more cognitively demanding than basic bodily actions. As such, though mental actions may be possible, mere agents may not be capable of engaging in them. The most plausible cases of mental action appear to be cases where one directs one's attention to a mental activity that is normally automatic and engages in this activity intentionally and purposively. If this is the case then one seemingly needs an awareness that one is capable of engaging in a mental activity, and some metacognitive capacities, in order to engage in a mental action.⁸² For instance, in order to make a mental note to remember something, one must come to know that one is capable of remembering.

Perhaps some mental actions do require the use of metacognition and sophisticated cognitive capacities. This doesn't mean that all mental actions are cognitively demanding though. It seems to me that there are *some* possible mental actions that are within the grasp of mere sentient beings. Proust gives the examples of directed hearing and seeing. By directing attention to the information we are taking in through our ears we are, as she puts it, listening not hearing.⁸³ Think of the difference between merely sitting in a noisy bar and sitting in the same bar actively trying to pick out the conversation of two people sat on a nearby table in the same bar. These activities are clearly different and the latter plausibly appears to be an action that can be done for a reason.

Even a merely sentient individual, suffering from a condition like locked-in syndrome, unable to move their body at all, could nonetheless engage in directed sense perception. As such, whilst they may be entirely unable to physically move they would still be an agent in virtue of their ability to engage in the mental act of directed perceiving. Of course they may lack the senses of vision and hearing but, if directed perception of sight and sound can be

⁸⁰ Proust, "A Plea for Mental Acts," 108.

⁸¹ One may think one needs to possess a body in order to be sentient. I take no stand on this here.

⁸² Proust, "Mental Acts," 211; Proust, "A Plea for Mental Acts," 10.

⁸³ Proust, "A Plea for Mental Acts," 5.

done for a reason, it seems directed perception of other sensory modalities could be done for a reason too. Thus, since all sentient beings are at least capable of experiencing pleasant and unpleasant affective feelings they could engage in attention directed experiencing of affective feelings. Thus, even a minimally sufficient sentient being who also happens to lack perceptual input from the world, would be capable of at least one kind of action; directed perceiving of her own affective feelings. Thus possessing a body is not necessary to be an agent. Any being that is sentient must also be an agent.⁸⁴

In addition to the claim that agents need not be able to engage in bodily action, some other facets of the notion of agency are worth considering here. I will argue in the next subsection that agency should be understood to be a capacity that one possesses regardless of whether one exercises this capacity. Thus even an individual that never acts for a reason is still an agent in light of their capacity to act for reasons. It also seems that we should not restrict the title of 'agent' to those beings which are only able to act in the here and now. Plausibly there are times when agents lack the ability act for a reason for brief periods of time yet they do not cease to be agents (and thus do not lose their moral status).

ii. Agency as a Capacity

One may understand agency to be a capacity one has. Up until this point I have been discussing agency as a capacity. However one can also understand agency as an ideal one strives to reach. This distinction is commonly seen in the autonomy literature. Some theories of autonomy suggest that it is something one aspires to achieve and thus one is autonomous if, and to the extent to which, one manages to exercise control over one's motivations/actions. Other theories of autonomy suggest that autonomy is the *mere ability* to control one's motivations/actions. Thus one is autonomous in virtue of having the capacity for autonomous action even if one never exercises this capacity. Similarly, we can understand agency as something one has if one sufficiently exercises one's capacity to act or we can understand agency as the mere capacity to act. I take it that the relevant kind of agency for the Agency Account of Moral Status is the *capacity* to act for reasons. Thus we should understand a moral

⁸⁴ It is worth reminding my reader that I do not maintain that such beings exist but merely use them as a theoretical tool to show that the possession of sentience logically entails the possession of agency. That being said I do not deny that the existence of such beings is logically possible, though I recognise that there may be various reasons for why a sentient being that completely lacks any sensory input could not exist within our world.

status holder to be any being with the capacity to act for reasons, regardless of whether they exercise this capacity.

This is right, I think, because in merely having the capacity for agency one makes use of one's agency. Even in not acting one makes a choice and decides not to act. Whether or not such 'non-actions' are true actions is not a debate I shall enter into here, but it is noteworthy that such non-actions are available to (and only available to) agents. Without the capacity for agency, an individual would be causally forced into certain behaviour, whereas agents, even if they simply engage unreflectively in the action that strikes them at any given moment, or simply do nothing at all, choose to act in this way. In being an agent then, one cannot help but to use one's capacity to some extent, even if it is to engage in absolutely nothing. Thus it seems that merely possessing the capacity for action should make one an agent regardless of whether one decides to exercise it and act for a reason or not.

Further it does not seem to me that being able act at any given moment is necessary for one to be an agent. We often think of individuals as agents even though they currently lack the opportunity to exercise their capacity for action. Whilst I sleep for example I am not acting, nor am I in a position to act. I cannot exercise my agency even if I wanted to. Nonetheless I am plausibly still an agent whilst asleep. If I am knocked unconscious and regain my consciousness later that day then again it seems I am, and have been an agent, all day, despite my loss of the ability to act. It also seems that an individual who is restrained and under the influence of heavy drugs does not lose their agency either. They remain an agent despite being unable to currently act intentionally. At some point, if the period of time for which one loses the ability to act is too long or permanent, one ceases to be an agent. For instance, an individual who falls into a coma from which they could never recover plausibly ceases to be an agent. Therefore I suggest that, though we may cease to be agents if we permanently lose our ability to act, or lose it for too long, being an agent remains a property of an individual though brief periods in which one is unable to make use of one's agency.

Thus being an agent means possessing the ability to act intentionally although not necessarily exercising this capacity through action. Furthermore, agents may go through short or temporary periods in which they are unable to act intentionally. They may also lack a body (in principle) and be unable to act in the physical world. They may only engage in 'mental' actions. As such, since minimally sufficient sentient beings at the very least, are able to engage in intentional mental actions (such as directed perception), possession of sentience must imply

the possession of agency. Therefore all sentient beings must be agents in the sense defended here.

Conclusion

In this chapter I have argued that the possession of desires requires the possession of beliefs, and thus sentient beings must possess beliefs. Further I argued that such beliefs need not be specifiable in language and nor conceptual in content. And that in possessing desires and beliefs, sentient beings are capable of possessing motivating reasons. I then argued that the ability to act on a motivating reason is sufficient for intentional action. Acting for a reason doesn't require one to be able to consciously recognise, explain or reconsider one's reasons and thus doesn't require any metacognitive capacities which a minimally sentient being would lack. Finally, I argued that there is good reason to consider some mental activities to be actions and that directed perception is an example of such a mental action. Such a mental action, I maintain, could even be engaged in by a sentient being lacking a functional body. Therefore, I have shown that sentience entails agency and thus that if a being is sentient so must she be an agent.

Chapter 5

The Agency Account of Moral Status: Liberty and Pain



Over the next two chapters I will argue that having agency, understood as acting in a way that can be rationalised by beliefs and desires, is a sufficient and necessary condition for possessing moral status. I will start in this chapter, by arguing that agency is a necessary condition by bringing together the argument that has been laid through Chapters 1, 3 and 4. I will then argue for agency as a sufficient condition for moral status. In order to do this I will show that agents, merely in virtue of their agency, have three interests that are considered to be central to the notion of a full moral status holder: interests in continued existence, liberty and freedom from pain. In this chapter I will argue that agents possess interests in freedom from pain and liberty, and in Chapter 6 I will argue that agents possess an interest in continued existence. I will then show that in fact these same interests, when possessed by paradigmatic full moral status holders (autonomous agents), are grounded in agency, not autonomy as is often thought. Since it is these interests that we consider important agency not autonomy is a minimally sufficient condition for full moral status.¹

I. Agency as a Necessary Condition for Moral Status

First I will show that agency is a necessary condition for moral status by demonstrating that without agency one is not owed moral consideration. In order to have moral status, one must have morally significant interests. As argued in Chapter 1, many things can have interests in a conditional sense, that is to say, things can be good for them in order for them to achieve some states of affairs, for example in the way cars or daffodils have interests. However these interests are not morally significant. Cars may be said to have an interest in oil in order to function and daffodils an interest in sunlight in order to remain alive. However, dying and

¹ Throughout this chapter I will focus on showing that non-human agents incapable of autonomy possess these interests since they are the most controversial cases.

ceasing to function don't matter to cars and daffodils themselves. Cars and daffodils won't be affected by these interests being frustrated, because cars and daffodils cannot be affected at all.

Having morally significant interests I continued, requires the possession of desires (or some functionally equivalent motivating states).² If one has no desires or preferences then one will be entirely indifferent to how one is treated, and one will be unaffected by anything that is done to one. Thus one needs desires in order for things to matter to oneself. If one has desires, then when one's interests are frustrated or satisfied, it impacts one's well-being. As I argued earlier, this does not mean that all goods are experiential, or even that all goods are person-affecting, but merely that being capable of having desires, or some functionally equivalent mental state, is necessary for things to be good for one.

Finally, I argued in Chapters 3 and 4 that if one possesses desires, then one is capable of acting for a motivating reason and this is sufficient for possessing agency. Therefore, since any being with desires is an agent, if an individual isn't an agent, it simply cannot be the case that one has (morally significant) interests. Further, since we have nothing to consider when deciding what we morally owe to a non-agent, such an entity cannot have moral status. Agency then, is a necessary condition for possessing moral status since without it we would lack any interests.

II. Agency as a Sufficient Condition for Moral Status

Not only is agency necessary for moral status but it is also *sufficient* for moral status. Thus, if and only if one is an agent, one has moral status. The argument in support of this claim follows quite straightforwardly from what I have said above. Agents necessarily have interests and thus only agents can be harmed, in other words, only agents can have higher or lower levels of well-being. Unless we have good reason to think that this harm is morally insignificant (and it doesn't seem like we do) agency, in the minimal sense I articulate here, must be sufficient for moral status.

However whilst possessing agency may be sufficient for mattering to *some* extent, humans, persons or autonomous agents are often thought to have some kind or degree of

² If some beliefs are motivating states, such as the belief one arrives at after a moral judgement, then the possession of certain beliefs would be sufficient for moral status.

moral status over and above this. Autonomous agents, it is often suggested, have *full moral status* for which autonomy is a necessary condition. Having full moral status not only means that one's interests matter, but that there is a strong presumption against being treated in certain ways, specifically, having pain inflicted upon one, having one's life ended and having one's liberty restricted. If this is true then little ground is gained by advocating for the Agency Account of Moral Status. I think this claim is mistaken and will argue that mere agents, in virtue of being agents, can be *pro tanto* wronged in the same central ways as full moral status holders and thus possess full moral status just like autonomous agents.

I have already argued that status does not permit degrees in Chapter 1. Thus, assuming my argument here is correct, any plausible understanding of full moral status can interpret the notion not as a genuine degree of moral status, but only as a degree of moral status in the sense defended by the Unequal Interests Model of degrees of moral status. Understood in this way, the claim that only autonomous agents are full moral status holders means merely that the subset of interests in freedom from pain, continued existence and liberty is unique to autonomous agents. Assuming this to be the case, my argument here is not so much an attack on degrees of moral status but on the claim that this subset of interests is unique to autonomous agents.

The structure of my overall argument to establish that agents have full moral status comes from Jeff Sebo. Sebo also argues that (some) animals possess moral status in virtue of the possession of agency and does this by arguing that agential animals possess the typical interests we take to be hallmarks of full moral status holders.³ However my implementation of this argumentative strategy differs from Sebo's in a few ways. Firstly, I consider an interest in freedom from pain instead of an interest in property to be one of the three central interests of full moral status holders. Secondly, my argument for mere agents' possession of these central interests is significantly more in-depth than Sebo's. Thirdly, the specific interests I take mere agents to possess in continued existence and liberty are different from the kinds of interests Sebo takes agential animals to have in these goods.⁴

One may think that my understanding of full moral status is too narrow and that full moral status is better characterised by the possession of some entitlements other than freedom

³ Jeff Sebo, "Agency and Moral Status," *Journal of Moral Philosophy* 14, no. 1 (2017): 1-22.

⁴ It should also be noted that, unlike myself, Sebo takes agency to be a sufficient condition for moral status but not a necessary condition. *Ibid.*, 14.

from pain, continued existence and liberty. I do not deny that there may be other ways in which full moral status holders may be wronged, and further specific treatment that they may be owed. Moreover, there are certainly other goods that are available only to individuals with specific capacities; goods related to the capacity for self-consciousness and the capacity for autonomy, for example. However an absence of pain, continued existence and liberty are, I think, the *central* goods of full moral status holders. Being denied any of these goods, or not having access to them in the first place, are the three most significant ways in which one can *pro tanto* wrong a full moral status holder. Further, it seems fair to say that these are the three most common ways in which full moral status holders *actually* are harmed and wronged in everyday circumstances. Hence focusing on these goods seems appropriate.

Over the remainder of this chapter, and most of the next chapter, I shall argue that autonomy is not necessary in order to possess an interest in any of these goods, and that mere agents possess interests in all of these goods. Further, I will show that agents possess these interests in virtue of their capacity for agency. Therefore it seems that mere agency itself is sufficient for the possession of these interests, and thus being an agent is sufficient for possessing full moral status. This means that agents do not *just matter*, but that in light of their specific interests in continued existence, liberty and freedom from pain, we have strong *pro tanto* obligations to refrain from inflicting pain upon them, ending their lives and restricting their liberty.⁵

Though this argument will show that all agents have interests in freedom from pain, continued existence and liberty, this argument will not establish that all moral status holders have exactly the same interests in these goods. Since the mental and physical capacities of agents vary significantly, as one would expect, different agents will have interests of different strengths in these goods. Autonomous agents will usually have a stronger interest in continued existence for example, since they have more to lose through death. However what I will show is that all agents have the same *kind* of interest in these goods. Thus, whilst two

⁵ One may wish to use the term 'persons' to denote agents in light of this claim. I will not since there is little consensus among philosophers regarding the conditions of personhood and personhood is often taken to be synonymous with the highest degree of moral status. Since I do not recognise degrees of moral status, the question of whether agents are necessarily persons is moot. The persons/non-persons distinction will either line up with the moral status/no moral status divide, and thus be a redundant distinction, or merely mark out agents that have a specific set of interests as persons, something that could more precisely and unambiguously be carried out by merely listing the specific interests in question.

agents may be harmed to different degrees by the infliction of pain, they will nonetheless be harmed (and wronged) in the same way.

In what follows I will consider each of the three central interests of full moral status holders in turn, demonstrating that agency is sufficient for the possession of these interests. I will focus on agential animals throughout these sections since it is considered more controversial that animals can be ascribed these interests than any other non-autonomous agent. Through making this argument I will show that all sentient beings, in virtue of being agents, can be *pro tanto* wronged in these central ways. Once I have sufficiently defended this claim I will (in Chapter 6) give a subtraction argument of the same form as Jeff Sebo's.⁶ This argument will show that the kind of central treatment that we *pro tanto* owe towards an autonomous individual doesn't change when they lose autonomy. This demonstrates not only that mere agents possess these interests, but that the possession of these interests in autonomous agents, is also grounded in their capacity for agency.

Through this argument over the next two chapters I will not only establish the claim that sentient animals and humans share equal moral status and *pro tanto* entitlements to continued existence, liberty and freedom from pain, but I will also elucidate and defend the Agency Account of Moral Status. I will establish that the Agency Account can plausibly explain the central ways in which a moral status holder can be *pro tanto* wronged: through having pain inflicted upon them, by having their life ended and having their liberty restricted. Further, I will show that the Agency Account can explain the *pro tanto* wrongness of these kinds of actions through reference to agency, the condition for moral status on this account. Thus I will demonstrate that the Agency Account will come closer to the ideal account of moral status than competing accounts.

III. Pain

Pain, or more properly unpleasant affective feelings, are commonsensically thought to be intrinsically bad for an individual in virtue of being experientially unpleasant. As such, many argue we have a duty not to inflict pain on sentient individuals. However the Agency Account offers an alternative explanation of the wrongness of inflicting pain, rooted in the Desire Account of sentience that I argued for in Chapter 3. This explanation makes agency,

⁶ Sebo, "Agency and Moral Status," 18.

not sentience, the ground of the wrongness of inflicting pain. This argument runs like so: Since pain is a mental state that is intrinsically unpleasant and, since unpleasant affective feelings just are the frustration of some final *de re* desires, pain harms one because one's desires are necessarily frustrated. Having one's desires frustrated is a harm because one is being treated in a way that one does not want to be treated. Thus in virtue of being an agent with preferences, one is harmed by the infliction of pain. Without the capacity for final desires or some similar motivating mental states (an essential component of being an agent) one cannot be harmed by the infliction of pain in this way.⁷

Exactly which desire is frustrated by the infliction of pain is dependent upon kind of pain and the situation. In cases of physical pain, the relevant desire might be a desire for the absence of sensations linked to the tearing of one's flesh, or breaking of one's bones, for instance. In the case of emotional or psychological pain, different desires are frustrated. When someone insults me for example I may feel upset. However whether or not I feel upset appears to depend upon whether or not I have a desire not to experience being insulted. For some particularly thick-skinned individuals, who are on the whole indifferent to what others say, an insult might not cause the least bit of unpleasantness, however for other individuals or in a different context, an insult might well result in significant unpleasantness. What is wrong with insulting someone then, is not that it causes them unpleasant affective feelings, but that one treats them contrary to how they want to be treated. To insult them is to fail to respect the preferences they have as an agent.

All this however, is not to deny the unpleasantness of pain, but only to deny that this unpleasantness is at bottom what is wrong with inflicting pain, or causing the experience of unpleasant mental state. Of course pain is unpleasant but what is morally significant is that we experience this unpleasantness because we have a desire frustrated. *This* is what is wrong with inflicting pain. One may also note that other things may make an instance of pain worse (such as the lasting damage to a relationship, a broken bone or the loss of a friend) nonetheless these further harms are not intrinsic to the notion of experiencing pain.

⁷ Unpleasant affective feelings may also be instrumentally harmful for agents because they interfere with their bodily movements, reasoning or motivation, thus obstructing them from satisfying their final desires and engaging in specific courses of action. For instance, if I sprain my ankle, the pain that I suffer when I walk may encourage me to avoid ice skating (something I greatly enjoy). Thus inhibiting me from pursuing a course of action I wish to pursue and undermining my agency. However here I'll focus on the harm of the more direct frustration of desires. Hadley inspired me to think of this point. John Hadley, "Liberty and Valuing Sentient Life," *Ethics & the Environment* 18, no. 1 (2013): 92.

A counterexample for my proposed account of the harm of pain is masochism. Masochists enjoy and seek out painful activities. Plausibly then masochists desire to experience pain. However, if this is the case then my account of the harm of pain must be wrong. I have argued that pain accompanies the frustration of certain desires but in masochists it seems to accompany the satisfaction of desires. If I am to show that my account of the harm of pain is correct then I will have to explain away this problem.

It might simply be the case that the masochist lacks a relevant desire to find certain sensations unpleasant. Thus the masochist's situation is just like that of sufferers of pain asymbolia discussed in Chapter 3. The masochist does not truly experience pain but only experiences sensations that we ordinarily associate with unpleasantness. However, unlike the asymbolic, instead of lacking a desire for the absence of normally painful sensations he has a reversed desire. The masochist desires the experience of ordinarily painful sensations and so, since he desires them, he experiences them as pleasant. As such he doesn't really desire pain but only ordinarily painful sensations. This is a possibility (and may be true for some individuals) however I suspect that in the case of most masochists something slightly more complex is going on.

Perhaps, as most sentient agents do, masochists have a desire to avoid certain sensations that are ordinarily experienced as unpleasant but unusually, in addition to these desires, have a desire to experience unpleasant affective feelings. Thus when they experience ordinarily unpleasant sensations they have their desire to avoid experiencing such sensations frustrated, as most of us do, and feel unpleasant affective feelings. However, they then have their further, higher-order, desire to experience unpleasant affective feelings satisfied. Thus when they feel pain, they also feel pleasant feelings. As such the masochist, genuinely does enjoy the experience of pain. If either of these explanations are accurate then the example of masochism also doesn't present a counterexample to my account of the harm of pain. Pain still frustrates a desire of the masochist and thus is a *pro tanto* harm even if all things considered, satisfying the masochist's desire for unpleasant feelings, outweighs this harm.

Pain then, and unpleasant affective feelings generally, harm us by frustrating final desires for the absence of specific sensations. Thus, causing unpleasant affective feelings is *pro tanto* wrong because it undermines one's agency by treating one in a way in which one does not want to be treated. This view preserves the intuitive claim that pain is intrinsically bad for us, although it offers a different explanation for why this is the case to that which is usually

given. It is desire frustration not unpleasantness itself that makes inflicting pain *pro tanto* wrong.

IV. Liberty

Liberty is considered of vital importance for full moral status holders and having one's liberty restricted in certain ways is an important wrong. In this section I will consider several arguments for the value of liberty for animals. My aim here is to show that liberty has the same kind of value for mere agents as it has for autonomous agents. To do this I firstly will consider two popular arguments for the instrumental value of liberty. I will suggest that these arguments fail because they only explain the value of a rather limited degree of liberty or they fail to provide substantial argument for the degree of liberty they suggest animals are owed.

Once I have done this, I will suggest some reasons to think that in addition to its instrumental value, liberty has some intrinsic or final value in itself, regardless of its impact upon one's well-being. I will argue that the freedom to choose is valuable in itself in the case of animals, as it is in the case of humans. This is because choice grants one the ability to determine the course of one's own life. As such, it is good that non-autonomous agents possess liberty, even if it does not contribute to their well-being. However, unlike in the case of autonomous agents, it is not good that non-autonomous agents are entirely unrestricted in their actions, since non-autonomous agents do not the capacities required to comprehend the full implications of decisions that could result in significant harm. Thus though liberty has the same kind of value for autonomous and non-autonomous agents, it is not good for them to the same degree. Non-autonomous agents should be free to pursue self-regarding courses of action only so long as they are not a risk of significant harm.⁸

i. Kinds of Liberty

Before going any further I should clear up the notion of 'liberty'. Having one's liberty restricted needn't involve being confined to a specific area by physical barriers. One may have one's liberty restricted by not being allowed entry into a place or building for instance, not being allowed to use a service, or having one's property confiscated. One may also have one's

⁸ Following convention I shall refer to 'liberty' and 'freedom' interchangeably.

liberty restricted by being denied opportunities, such as being denied the chance to apply for a job. One shouldn't take the following discussion to be focused on the question of whether animals ought to be able to move freely in the world without physical confines. The debate here is much more fundamental.

There are three main types of liberty that are recognised in the philosophical literature: *negative liberty*, *positive liberty* and *republican liberty*. Most theorists reject one or more of these forms of liberty in favour of another, arguing that one form of liberty is genuine and that the others are derivative of it, and thus redundant or not *bona fide* notions of liberty.

Having liberty in the republican sense means one is in a state where one's interests could not be arbitrarily interfered with. One is deprived of republican liberty when one is vulnerable to having one's interests interfered with, or as it is frequently said, one is 'dominated' by other agents.⁹ I will not be discussing republican liberty in this chapter, however it should be relatively straightforward to see how much of what I say here could be applied to the question of non-autonomous agents' interest in republican liberty.¹⁰

Negative liberty is often summarised as 'the freedom to act without interference or obstruction from another agent'.¹¹ If no external obstructions from other agents stand in the way of one's course of action, then one has negative liberty.¹² Thus for example, I have freedom to eat the cake in the fridge if I am not restrained and the fridge is not padlocked. In contrast to this, one has positive liberty when one is free to act in accordance with what one truly wants, or acts authentically in line with one's highest-order desires.¹³ Thus whilst I might be free in the negative sense to eat the cake because I'm not restrained and there is no padlock on the fridge, I will not be free to eat the cake in the positive sense if I am under hypnosis or brainwashed not to eat the cake.

⁹ Quentin Skinner, "A Third Concept of Liberty," *Proceedings of the British Academy*, 2002, 243; Philip Pettit, "The Republican Ideal of Freedom," in *The Liberty Reader* (London: paradigm publishers, 2006), 224.

¹⁰ For a thorough discussion of non-human animals and their interest in republican liberty see: Valéry Giroux, "Animals Do Have an Interest in Liberty," *Journal of Animal Ethics* 6, no. 1 (2016): 20–43; Andreas T. Schmidt, "Why Animals Have an Interest in Freedom," *Historical Social Research* 40, no. 4 (2015).

¹¹ Isaiah Berlin, "Two Concepts of Liberty," in *Four Essays on Liberty* (London: Oxford Paperbacks, 1969), 122.

¹² Taylor, "What's Wrong with Negative Liberty?" 143–44.

¹³ Berlin, "Two Concepts of Liberty," 131,136.

Positive liberty is usually taken to apply only to autonomous beings because acting freely in a positive sense appears to require higher-order desires or acting authentically in line with one's judgments, capabilities that are tightly bound up with acting autonomously. Because of this close connection to autonomy one may contend that the relevant interest in liberty that full moral status holders possess is at least partially an interest in positive liberty. Thus since non-autonomous agents could at best only possess interests in negative liberty, they cannot be full moral status holders.¹⁴ Contrary to this, I doubt the clear distinction between positive and negative liberty and thus, I take it that an interest in positive liberty specifically is not an integral part of full moral status.

Following MacCallum I think that both positive and negative liberty are instances of the same kind of relation merely with different subjects, obstructions and courses of action. Thus, any instance of liberty (positive or negative) is a case where x is free from y to z .¹⁵ For example, David (x) is free from obstruction by Janet (y) to go out tonight (z). Or, I (x) because of my addiction (y) am not free to stop gambling (z). People distinguish between positive and negative kinds of liberty, MacCallum maintains, when they don't see this relation clearly enough.¹⁶ If this is right then (putting aside republican liberty) there is just one kind of liberty. Thus, showing that non-autonomous agents possess an interest in being free from obstructions (i.e. possessing what is usually referred to as negative liberty) will be sufficient to establish the claim that they have an interest in the same kind of liberty as autonomous agents.

This does not mean that autonomous agents and non-autonomous agents can suffer *exactly* the same liberty restrictions. There will certainly be some specific liberty restrictions that will only be applicable to autonomous agents. Only autonomous agents can have their liberty to act in line with their higher-order desires restricted, for example. Non-autonomous agents could have no interest in being free to act in line with their higher-order desires because they lack higher-order desires. Likewise, there will also be some specific liberty restrictions that will be unique to other groups of individuals. Not all agents can have their liberty to fly restricted since many agents lack an interest in being free to fly (because they lack wings). This

¹⁴ If I am mistaken, and some or all non-autonomous agents *do* possess an interest in positive liberty, this will only make my argument below stronger.

¹⁵ Gerald C. MacCallum, "Negative and Positive Freedom," *The Philosophical Review* 76, no. 3 (1967): 314.

¹⁶ *Ibid.*, 318.

just shows that autonomous agents' specific interests in liberty can have different content, not that they have interests in different kinds of liberty.

What also seems to be important here, is kind of interest agents have in liberty. Possessing liberty might be in one's interest for all kinds of reasons, and I will consider many of these in this section. However, if I am to show that mere agents have full moral status, I will have to show that mere agents are not only interested in the same kind of liberty as full moral status holders, but that they have the same kind of interest in this liberty. This does not mean that I will have to show that liberty is valuable to the same extent for paradigmatic full moral status holders and mere agents, but merely that it is valuable in the same *kind of way*.

I will argue that the kind of interest in liberty that is a hallmark of full moral status, is an interest in being free to make one's own choices without interference from other agents and that this kind of liberty has value because it allows one to determine the course of one's own life. Merely in virtue of being agents, I will argue, individuals have this kind of interest in liberty. I shall proceed by considering two popular arguments for the instrumental value of liberty for non-autonomous agents. Once I have done this I will then consider several arguments for the value of liberty in itself and argue that it is good that non-autonomous agents are able to make their own choices and determine the course of their lives.¹⁷ Thus non-autonomous agents have the same kind of interest, in the same kind of liberty, as paradigmatic full moral status holders.

ii. The Instrumental Value of Liberty

Almost all theorists who recognise the moral status of animals also take the liberty of animals to have some value. However many theorists do not offer a thorough and sustained argument for why they take liberty to have value, or a clear explanation of the kind of value liberty has. This I suspect is because many have taken liberty to be an obvious component of well-being. This being said, two claims that have been defended are that liberty has some

¹⁷ One should note that all the arguments I consider here aim to establish the value of the liberty to engage in self-regarding acts since possessing absolute liberty is not morally defensible. One should not be free to restrict another's liberty, or to harm others for instance (at least not without sufficient justification). What's more, being entitled to freely engage in some actions doesn't mean that one is entitled to engage in these actions *under any circumstances*. Plausibly it seems that one's liberty can be permissibly restricted under certain conditions, when one becomes a danger to others for instance, or as a form of punishment when one fails to abide by the law.

value for animals because it allows them to live lives that are satisfying and liberty has value because it permits them to exercise their natural capacities.¹⁸ I'll here try to charitably interpret these claims in order to give two accounts of the value of liberty for animals: *the Desire-Based Account* and *the Teleological Account*. I will then argue that these accounts appear insufficient in capturing the value of liberty in the case of animals.

iii. The Desire-Based Account

Cochrane argues that the value of liberty is purely instrumental for animals. Liberty is good for them because it allows to pursue their desires.¹⁹ Thus the restriction of liberty is bad for them where, and to the extent that, it obstructs them from pursuing their desires. According to this account animals are not harmed by restrictions of liberty themselves, but by the suffering that these restrictions cause. However, this explanation isn't sufficient to show that liberty is instrumentally valuable. This argument is an attempt to show that liberty is valuable because it provides us with something else of value: the pursuit of desires. But why ought we to consider the pursuit of desires to be something good? The obvious answer is that through pursuing desires we can (at least sometimes) satisfy them and satisfying desires is good.²⁰

However desire satisfaction isn't always valuable. For one, though satisfying a desire may be pleasant, the overall effect on our well-being may be negative, for example where satisfying a desire brings along with it substantial health problems. As Cochrane himself notes, a dog may have a desire to eat as much food as he can. In the long-term, if this desire was satisfied every time it was pursued, it would lead to obesity and related-problems; this certainly wouldn't be good for the dog's well-being.²¹ Further desire satisfaction doesn't always result in pleasure. Where I desire something that, as it turns out, I don't enjoy very much, like eating at a new restaurant which I discover serves incredibly underwhelming food,

¹⁸ David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (Cambridge; New York: Cambridge University Press, 1996), 269.

¹⁹ Cochrane also suggests another argument for the instrumental value of liberty which states that liberty is valuable where, and to the extent that, it leads to pleasure. I will not discuss this argument separately since I take it that this argument, and the Desire-Based Account, rest on the same fundamental claim that liberty is good where it contributes to one's well-being, and thus they both ultimately suffer from the objections I put forward below. Alasdair Cochrane, "Do Animals Have an Interest in Liberty?," *Political Studies* 57, no. 3 (2009): 664.

²⁰ The reason I draw out this, perhaps painfully obvious claim, will become clear below.

²¹ Cochrane, "Do Animals Have an Interest in Liberty?," 670.

desire satisfaction may be accompanied with disappointment, not pleasure. Thus not all desire satisfaction appears to contribute to one's well-being and so liberty cannot be good because it grants one the opportunities to satisfy just any desires.

At best, liberty could be valuable on the Desire-Based Account where it allows one to pursue desires which contribute towards one's well-being when satisfied. Thus perhaps liberty is good where it allows one to pursue the satisfaction of rational and informed desires. Or one might think that there are some objectively important desires, e.g. desires for food or water, the satisfaction of which contributes towards one's well-being, and thus that liberty is valuable when it allows one to pursue these. I'll take no stand on which of these sets of conditions (if either) most accurately pick out the group of desires that contribute towards one's well-being when satisfied. However if the satisfaction of some sub-set of desires contributes towards one's well-being, then being free to pursue these desires would plausibly be good. So one explanation of why liberty is a good for animals is that it grants one the opportunities to pursue well-being-promoting desires.

It certainly seems that animals should be free to pursue well-being-promoting desires. However being free *only* to pursue these desires seems like a very limited degree of freedom. If liberty is valuable only in this instrumental sense, then it seems we could permissibly restrict animals' actions in some considerable ways. For instance, where the overall contribution of the satisfaction of a desire towards an animal's well-being is neutral, they need not be free to pursue this desire. Thus, imagine a dog that has a desire to eat grass every day during his walk and that eating grass occasionally causes him mild digestive discomfort. Let's assume that over the course of his life, the contribution to his well-being from satisfying this desire is exactly equal to the amount that he loses from his well-being through digestive discomfort. In such a case, since satisfying his desire to eat grass does not overall contribute to his well-being, we can permissibly obstruct him from pursuing this desire without any further justification. Likewise, where a desire is misguided, and its satisfaction will not result in any contribution to one's well-being.

If it is the case that the satisfaction of one of these desires will not only fail to contribute towards his well-being but in fact negatively affect his well-being, then not only do we have no obligation to allow him to satisfy this desire but, seemingly we have an obligation to *obstruct him* from pursuing this desire. If liberty is good when, and to the extent that, it allows one to pursue well-being-promoting desires, then presumably it is bad when, and to the extent that, it allows one to pursue desires that diminish one's well-being. Thus, if the dog's mild

digestive discomfort, was detrimental to his well-being to a greater extent than eating grass was beneficial to his well-being, as moral agents we ought not to allow him to pursue his desire of eating grass. This would be the case even if the negative effect on his well-being was very minor.

Some might think that these kinds of liberty restrictions sound benign. However, the moral significance of these restrictions can be more easily seen when we think about the implications of such restrictions on undomesticated animals and animals within national parks and nature reserves. It would seem that we ought not to let wild animals make bad choices, and need not allow them to even make choices that result in a neutral impact on their well-being. Having this degree of control over animals' lives seems at the very least counter-intuitive.

Additionally, this account allows us not just to obstruct animals from choosing to act in ways neutral or negative towards their well-being, but to actually deprive them of choice altogether. Liberty is valuable on this account, as I made clear above, not because it allows the pursuit of well-being-promoting desires *per se*, but because it allows *the satisfaction* of well-being-promoting desires. This means that so long as an animal's well-being-promoting desires are satisfied, they need not be free to pursue these desires. Thus a life in which one's desires were tended to at all times, would be no more objectionable on this account than one in which one was free to pursue these desires oneself.

For instance, consider the life of a horse on a nature reserve with a very attentive and all-knowing ranger. The ranger is so doting on his companion that the horse has a regimented routine to ensure all necessary goods are provided for her. Every minute of her life is planned out according to a schedule to ensure a maximum possible level of well-being. She has organised feeding sessions with a varied diet of food that is good for her and pleasant tasting, different play sessions, time to socialise with other horses, etc. Though she has all her desires satisfied, she is not really free to pursue her desires, all her desires are satisfied for her. By stipulation the horse is very happy with this arrangement and has a high level of well-being, however her liberty is limited, and, as I will argue later, it seems that she is wronged by this restriction. Nonetheless, such treatment appears permissible on this account and might even be morally required if there would be a greater risk of the animal lowering their well-being were they free to pursue their desires themselves, than if a paternalistic ranger satisfied the animal's desires for them.

I suspect that many may fail to see any problem with enacting this kind of paternalism over animals' lives.²² In fact some may see it as morally required.²³ I'll elaborate in more detail below why I think that this rampant paternalism is problematic. For now however I'll merely note that we don't think that such overarching control is justifiable, let alone morally required, in the case of humans. We take humans to be owed considerably more freedom than this, even humans that lack autonomy. As such, unless there is a morally relevant difference to justify this difference in treatment, there is a *prima facie* case to be made against exerting severe paternalistic control over animals' lives.

iv. The Teleological Account

The Teleological Account offers an alternative account of liberty's instrumental value that is rooted in the Aristotelian idea that beings have an intrinsic *telos*, that is a function, purpose or nature. One's function is determined by the kind of thing that one is, which at least within the context of this argument, is usually taken to be the species to which one belongs. This account suggests that functioning in the way natural for one's species is good and that, since liberty is necessary in order to function, liberty is valuable where and to the extent that it allows one to engage in one's natural functions. Further, obstructing one from engaging in the natural functions of her species harms one.²⁴ Thus, according to the Teleological Account, keeping a cat indoors inhibits her natural function of roaming (for instance) because she cannot engage in such behaviour whilst locked inside a house. As such, it would be *pro tanto* wrong to keep a cat confined to indoor space, since doing so would deprive her of a natural species-specific functioning.²⁵

²² Regan is one of the few to recognise the harm of this specific kind of paternalism in animals: Tom Regan, *The Case for Animal Rights* (Berkeley: University of California Press, 1983), 92.

²³ Cochrane seems to accept that such paternalism is acceptable. Cochrane, "Do Animals Have an Interest in Liberty?," 666.

²⁴ Bernard E. Rollin, *Animal Rights and Human Morality* (New York: Prometheus Books, 1981), 35; Martha C. Nussbaum, *Frontiers of Justice Disability, Nationality, Species Membership* (Cambridge: Harvard University Press, 2007), 347; Paul W. Taylor, *Respect for Nature: A Theory of Environmental Ethics* (Princeton: Princeton University Press, 2011), 108-9; Cochrane, "Do Animals Have an Interest in Liberty?," 670-71. Nussbaum, has argued that animals should not be free to pursue merely natural functions, but only those natural functions which are also central and good to one's species. I'll attack the more general formulation of argument, rather than Nussbaum's more limited 'natural, central and good' formulation. This is mainly because I will argue that the Teleological Account isn't inclusive enough in the variety of actions it suggests animals should be free to pursue and the general formulation is the most inclusive of the two versions of the account.

²⁵ Clare Palmer and Peter Sandøe, "For Their Own Good Captive Cats and Routine Confinement," in *The Ethics of Captivity*, ed. Lori Gruen (Oxford: Oxford University Press, 2014), 145-46.

A significant problem for this account is pinning down exactly what is meant by a 'natural function'. Until this is clear it is difficult to determine what exactly this account suggests an animal should be free to do, and therefore difficult to assess whether it grants animals a plausible degree of liberty compared to other rival accounts. I'll consider some different possible understandings of 'natural function' and the implications these understandings have for this argument for the value of liberty.

One could understand natural functions to be merely biological functions. Taylor endorses this position.²⁶ A biological function might charitably be interpreted to be any function of one's body that needs to be performed to ensure minimum health and continued existence e.g. eating, drinking, sleeping, etc. If we interpret natural functions in this way then this account states that liberty is valuable where it allows one to fulfil one's biological functions. Fulfilling these functions is undoubtedly valuable, since this is necessary to have any reasonable level of well-being, thus liberty plausibly has value where it allows one to fulfil these functions. However *only* being free to fulfil these functions is a very limited degree of liberty. The fulfilment of biological functions alone doesn't constitute a good life. Much more is necessary than this for a good standard of well-being. It seems intuitive that animals ought to at least be *pro tanto* free to pursue those goods necessary to have a good standard of well-being. So, if natural functions are best understood in this sense, then the Teleological Account does not offer a convincing explanation of the value of liberty for animals.

An alternative understanding of natural functions, is as the functions that members of one's species perform when in their natural environment, free from human disruption. This understanding of natural functions also suffers from problems however. Firstly there are some problems here concerning what the 'natural environment' of some animals are (particularly domesticated animals) and thus how we can determine their natural functions. More fundamentally however, interpreting natural functions in this way implies that there is something good about living in a natural environment, and by taking animals out of their natural environment we necessarily wrong them or harm them. This position seems distorted.

Animals behave in the way they do because they have adapted their behaviour to their environment in order to secure a good level of well-being. Where they live in a different environment, different behaviour may well be appropriate to get the most out of life. This shows us that fulfilling natural functions is not good because these specific behaviours are

²⁶ Taylor, *Respect for Nature*, 109.

good in themselves, but because fulfilling these functions *in a natural environment* leads to a good standard of well-being. Thus, where one lives in some other environment it isn't obvious why acting in the way that one's conspecifics do, in environments undisturbed by humans, is good for one. For instance, as Donaldson and Kymlicka note, it may be good for a dog that lives in a city to learn how to use the subway and the appropriate places to defecate, however these skills would be of no use to a dog that lives in the countryside.²⁷ It might also be good for a chimp in contact with humans, to learn and use sign-language, as this would allow him to better communicate and get what he wants.

These examples are clearly not natural functions under this description yet *are* actions that are plausibly good for those animals in their environments. As such it seems plausible that they should be free to pursue such courses of action. So, if we interpret natural functions to be the functions that members of one's species perform when in their natural environment, then the Teleological Account again seems to offer a rather restrictive degree of liberty and doesn't provide a particularly robust argument for doing so.

Another way to understand natural functions is as those behaviours that are motivated by 'natural desires' or what many might call 'instincts'. An immediate problem with this interpretation of natural functions is that much animal action cannot be classified as natural functioning. Many animals, especially social animals, learn at least a portion of the skills they possess as adults from observing their elders whilst young. Consider appropriate social actions, as well as specialist skills such as climbing and hunting, for instance. So, when an adult lion hunts, she is not motivated to engage in such actions by natural desires, but learnt desires. Since forming and acting on these desires seem to be important to an animal's well-being it seems that they plausibly ought to be free to act on them.

To retort, one could argue that whilst animals do learn some behaviours, they have a natural desire that motivates them to learn these behaviours. Therefore, whilst some actions are not *directly* motivated by natural desires, all actions are *ultimately* motivated by natural desires. Unfortunately this response makes the notion of natural functioning too broad and suggests that an animal *always* acts naturally, because all action that an animal could possibly engage in is natural functioning. Further this makes unnatural action impossible at least under most normal circumstances. Thus the 'natural' in 'natural functioning' becomes redundant.

²⁷ Sue Donaldson and Will Kymlicka, *Zoopolis: A Political Theory of Animal Rights* (Oxford: Oxford University Press, 2013), 97.

This raises questions about whether this is a plausible understanding of 'natural functioning'. Moreover however, this would also mean that the Teleological Account would effectively claim that it is *pro tanto* good for an animal to engage in *any* action and so animals ought to be *pro tanto* free to engage in any (self-regarding) action. Some do think that something like this claim is true.²⁸ However even if this conclusion is true, it cannot plausibly be the case that all of an animal's possible actions could contribute towards her well-being *because her actions are natural* (and cannot be anything but natural). Since all actions are natural on this view, this explanation would be entirely vacuous. So this third interpretation of natural functioning also fails to give us a plausible rendering of the Teleological Argument.

Putting aside the question of what counts as a natural function, it isn't clear why engaging in natural functions is good. Since the value of liberty on this account is derived from the value of natural functioning, if natural functioning isn't valuable, then liberty cannot be valuable either. An obvious solution may be that natural functioning contributes towards one's well-being because functioning is pleasant, and/or being unable to engage in natural functioning is unpleasant. However, there are counter-examples to both these claims.²⁹ For instance, fighting between adult male walrus for access to females could plausibly be said to be the fulfilment of a natural function, yet being unable to fulfil this function would clearly not be detrimental to a walrus's well-being. In fact, it seems it would significantly contribute towards their well-being to not fulfil this function (other things being equal) considering the brutality of such fights.

One might attempt to avoid counter-examples like this one by arguing that natural functioning is just a fundamental component of well-being, which is not reducible to any other good, such a pleasure or health. However, I for one, am unconvinced by this claim. It isn't clear how exactly natural functioning could affect one's life if not through granting one some other good or depriving one of some other harm. Nussbaum tries to avoid counter-examples such as the one above by arguing that not all natural functioning is good, but only that functioning which is central and good. She offers little explanation of exactly what she means

²⁸ This will be true even on Nussbaum's more restrictive account too. Nussbaum's account will amount to the claim that one should be free to pursue all actions which are central and good. This is, I think, correct but Nussbaum is attempting to spell out what such central and good actions would be. Naturalness does the bulk of the work for Nussbaum in attempting to answer this question so without the 'natural' criteria, her argument does not get her far toward her conclusion.

²⁹ Cochrane, "Do Animals Have an Interest in Liberty?," 671-72.

by 'central and good' however, and she seems to come very close to question-begging by claiming that those natural functions that are good are those that are good for one.

However even assuming that either of these rejoinders succeed, and it can be maintained that fulfilling natural functions is valuable because it contributes towards one's well-being, then it seems that the Teleological Account at bottom, is based on the same fundamental claim as the Desire-Based Account: liberty is valuable where, and to the extent to which, it contributes towards one's well-being.³⁰ As such, the Teleological Account is subject to the same criticisms as the Desire-Based Account. We could still permissibly exert a great deal of control over the lives of animals and severely limit their liberty according to this account. As I argued above, this is out of kilter with the degree of liberty we believe humans, even severely intellectually disabled humans, to be owed.

To attempt to avoid this conclusion one may argue that natural functioning is not valuable because it contributes towards one's well-being, but is valuable *in itself*. Such an account would maintain that natural functioning is valuable in a *non-agent-affecting way*, just as it is often thought that autonomy is valuable for humans.³¹ However this argument is also rather unconvincing. For what reason could natural functioning be good in itself? It is rather mysterious why this might be the case, considering all that has been said about natural functioning in this section. If natural functioning is valuable at all it seems likely it will be valuable instrumentally. If one was to adopt such a position, significant argument will need to be given for it to be taken seriously.

The Teleological Account has some problems then. As it stands it is unclear why exactly natural functioning should be considered valuable and thus unclear how this account can establish the value of liberty. Further, even if natural functioning is valuable, it is not clear what natural functioning is exactly. All of the interpretations I consider offer an unappealing degree of liberty. Either this account allows animals too little liberty or fails to provide a plausible reason to back up the claim that animals are owed absolute liberty to engage in self-regarding acts.

³⁰ Although, assuming that there is more to one's well-being than natural functioning, the Teleological Account doesn't claim that liberty is valuable wherever it allows one to contribute to one's well-being but only where it allows one to contribute towards one's well-being through natural functioning.

³¹ Rollin seems to be committed to such a view. Rollin, *Animal Rights and Human Morality*, 35, 38-40.

v. Liberty as Choice

In contrast to the Desire-Based Account and the Teleological Account I think that animals ought to be free not just to engage in those actions which contribute towards their well-being, but be free in a stronger sense. I think that animals ought to be *pro tanto* free to make their own choices. It is widely accepted that autonomous agents should be free to make their own choices even if these choices do not directly lead to them having some desire satisfied, or achieving some good.³² It is often thought that liberty is something that is valuable for autonomous agents regardless of whether they use it to make good or bad choices or of any effect it has on their well-being. I will consider several arguments that attempt to explain why this kind of liberty is valuable. I will then argue that the best account, the *Self-Determination Account*, can also explain why liberty has this same kind of value for animals too. However I will argue that though liberty has value of the same kind it is not valuable to the same extent for animals.

Having the liberty to pursue courses of action is thought to be good because having choice, or choosing, has value even where we don't choose the option which is all things considered best for us.³³ Thus where we are deprived of the opportunity to make our own choices, we are wronged. Even in cases where what we are being stopped from doing is something that is detrimental to our well-being, we are wronged because we have our capacity for choice overridden by another agent. Being able to make our own choices is good, in fact, it is central to the idea of being an autonomous agent. However despite the widespread intuition that this is the case, it is not so easy to say why exactly having choice is good.

vi. The Instrumental Value of Choice

One plausible account of the value of choice is the *Prudential Account*. According to the Prudential Account, because I know myself better than anybody else, I am best placed to make

³² Of course, as I mentioned earlier, liberty to undertake *any* course of action cannot be good. We clearly should not be free to pursue courses of action which wrong or inflict unjustified harm on others for instance. We can at best only be free to pursue only self-affecting courses of actions. Marina Oshana, "How Much Should We Value Autonomy," *Social Philosophy and Policy* 20, no. 2 (2003): 99–126.

³³ Göran Duus-Otterström, "Freedom of Will and the Value of Choice," *Social Theory and Practice* 37, no. 2 (2011): 262; Robert Young, "The Value of Autonomy," *The Philosophical Quarterly* 32, no. 126 (1982): 39; Regan, *The Case for Animal Rights*, 91; Mark Rowlands, *Animals Like Us* (London: Verso, 2002), 153; Natalie Thomas, *Animal Ethics and the Autonomous Animal Self* (London: Palgrave Macmillan, 2016), 90–91.

choices that will be most beneficial to my well-being. So having choice has value because it will likely grant one a higher level of well-being than one would have otherwise (other things being equal).³⁴ Choice has instrumental value on this account in the sense that choice does not have value in and of itself but only because of what it brings about: a higher level of well-being. It seems that choice likely does have value in this sense, however there are some circumstances in which this account does not recognise the value of choice, and yet choice is still intuitively valuable.

Imagine that you cannot decide what to do with your day off. There are two options you like, but you cannot decide which one you would prefer. However, nearby is a perfectly informed paternalist.³⁵ This individual is committed to only ever choosing the option that best satisfies you. Since they are perfectly informed, they are better informed than yourself and will always make the decision that best satisfies you. Therefore there is less risk in the perfect paternalist choosing for you, than you making the choice yourself. If you value choice only because being free to choose usually leads to you having a higher level of well-being, you should have no objection to the perfect paternalist choosing what you will do on your day off. In fact in the interest of maximising your own well-being, the perfect paternalist should make all your choices for you.

There's more however: imagine you are again faced with the decision between two options to keep you entertained on your day off. This time there is no perfect paternalist around, however there is an imperfect paternalist nearby. The imperfect paternalist, like the perfect paternalist, is committed to choosing the option that best satisfies you when making any decision on your behalf. However, they are not perfectly informed, in fact they know exactly what you know about yourself and any choices you could make. As such the imperfect paternalist is just as likely to pick the option that genuinely satisfies you, as you are yourself. This being the case having the freedom to choose yourself will not make it any more likely that you will end up with the option that best satisfied you. So, if you think that the Prudential Account can fully explain the value of choice, then you ought not to have any qualms with the perfect paternalist or even the imperfect paternalist choosing on your behalf.

³⁴ T. M. Scanlon, *What We Owe to Each Other* (Cambridge: Harvard University Press, 2000), 251–52; John Stuart Mill, "On Liberty," in *On Liberty, Representative Government, the Subjection of Women* (Oxford: Oxford University Press, 1918), 83–84; R. G. Frey, "Autonomy and the Value of Animal Life," *The Monist* 70, no. 1 (1987): 55; Daniel Brudney and John Lantos, "Agency and Authenticity: Which Value Grounds Patient Choice?," *Theoretical Medicine and Bioethics* 32, no. 4 (2011): 218.

³⁵ This example comes from Duus-Otterström, "Freedom of Will and the Value of Choice."

Many of us however, would not be happy with the perfect paternalist making our choices for us, let alone the imperfect paternalist. Of course we might sometimes *allow* this to happen. We might ask the waiter what they recommend and accept their suggestion, presuming them to have superior knowledge of the available dishes than ourselves. However, we wouldn't want individuals to make choices on our behalf uninvited, even if this would be guaranteed to result in us being satisfied with the chosen option. I don't want to walk into a restaurant and be told what food I'm having! Whilst choice likely does have instrumental value of the kind supposed by the Prudential Account, it seems to have some other kind of value. Having choice seems to be of value even when others could make better decisions for us, by our own lights, than we could ourselves.

An account that could perhaps explain the value of choice in these situations is the *Self-Development Account*.³⁶ The Self-Development Account claims that choice is valuable because through making choices ourselves, we can develop our choice-making skills to become better choice-makers. Thus, even when we make choices that do not contribute towards our well-being, choice is still valuable because we can learn from our mistakes. This account maintains that choice is still valuable even if there is a perfect paternalist nearby to make a choice for us. Though they would likely make a choice that better satisfies us, if the perfect paternalist were to make the choice for us then we would miss out on a development opportunity.

However the Self-Development Account doesn't specify why developing as a choice-maker is valuable. One reason to think it has value is because good choice-makers choose the best options more often than bad-choice makers. So becoming a better choice-maker will increase one's likelihood of having a higher standard of well-being. If this is the case then developing as a choice-maker is only instrumentally valuable and so having choice must too be instrumentally valuable. Where developing as a choice-maker does not lead to an increase in well-being, choice itself cannot be said to be valuable on this account. Thus, I will argue it seems that there are at least two kinds of cases in which choice will not be valuable on this account, yet in both cases it intuitively seems that choice does have value.

Whilst the Self-Development Account tells us that it is better for us to make our own choice than to let the paternalist choose, this is only because in the long run it will likely lead to us having a higher standard of well-being. Assuming a best-case scenario where you always

³⁶ *Ibid.*, 262.

learn from your choices, you will still never develop into a perfect choice-maker. You will always make choices in which you do not choose the option which maximises your well-being. If there was an *ever-present perfect paternalist* nearby to make your choices for you at all times, this would result in a higher standard of well-being over your lifetime than if you made your choices yourself, and always learnt from your choices. Since self-development is valuable only when it increases your well-being, the Self-Development Account cannot explain why making choices yourself instead of allowing the ever-present perfect paternalist to make them for you, is valuable in this case.

If one believes choice only to have value because it allows one to develop as a choice-maker, one cannot object to the ever-present perfect paternalist making all of one's decisions. Many of us however, will find this unappealing. There seems to be something valuable about being able to choose for oneself even in this case. A defender of the Self-Development Account attempt to may block this line of argument by suggesting that self-development has some value in itself. However assuming that such an argument is possible (and it is not clear what it would look like) there are still some cases where choice doesn't lead to self-development and yet still appears valuable.

For example consider the case of *the serial bad choice-maker*. Imagine someone who day after day makes the decision to buy their lunch from the work cafeteria, but every day they end up disappointed with their lunch. They would prefer to get their lunch from somewhere else but they never learn from their mistake. Every day they make the same bad choice. Being able to choose where to buy their lunch from cannot be valuable because it allows them to develop as a choice-maker, since they do not develop as a choice-maker through having this choice. However having this choice seems valuable. Many find it deeply intuitive that one should be free to make bad decisions over and over in this way, even if one doesn't develop as a choice-maker. If this is the case then the Self-Development Account cannot fully explain the value we believe choice to have.³⁷

If my intuitions are correct then neither of these accounts can fully explain the value of choice. Whilst choice likely does have value in the previously suggested ways, it seems that

³⁷ Another account of the value of choice is the *Representative Value Account* which claims that through making a choice oneself one gives it meaning that would be lacking if someone else were to make the same choice on one's behalf. However, assuming choice has value in the way this account suggests, this account still does not seem to capture what makes the freedom to choose valuable fundamentally. Scanlon, *What We Owe to Each Other*, 252; John Harris, "Consent and End of Life Decisions," *Journal of Medical Ethics* 29, no. 1 (2003): 11.

choice has some more fundamental value. It seems that choice has some value *in itself*. That is to say, being free to choose is intrinsically valuable and thus it is always *pro tanto* good that one makes one's own choices. It is good to choose for oneself even when doing so results in an overall lower level of well-being than one would have were one to lack choice.

vii. The Self-Determination Account

There have been several attempts to give an account that captures the intrinsic value of choice and many of them skirt around the same idea: self-determination.³⁸ However there does not appear to be a single thoroughly defended account that focuses on self-determination as the source of value. As such I will here draw on various overlapping accounts of the intrinsic value of choice to construct *the Self-Determination Account*. That being said, I have tried to make clear through references, which individuals, make which specific claims.

The Self-Determination Account suggests that choice is valuable in itself because through having the liberty to choose we are able to determine the course of our own lives.³⁹ We are able to live the life we want to lead, and determine our own path, making our lives our own.⁴⁰ We get to be the one who 'calls the shots' and this is a good thing even if a lower standard of well-being results.⁴¹ Further, because we can determine our own lives we have a certain kind of authority over our own choices.⁴² We and we alone are able to decide on 'the shape, content, and nature of [our] own life'.⁴³

This value is perhaps easier to grasp when we look at what is wrong with taking away someone's choice. The Self-Determination Account suggests that when we interfere with other's choices we fail to recognise them as individuals capable of making their own choices. We do not give them the appropriate level of respect they are owed as agents capable of making choices about their own life, and we fail to recognise the authority they have over

³⁸ If one outrightly rejects choice having any value in itself at all and believes that it is only valuable in these extrinsic senses suggested by the previously considered accounts then this will not be a problem for my argument overall. It seems likely that choice will possess the kind of value suggested on the Instrumental Account and the Self-Development Account for animals as well as humans. Thus animals will nonetheless have the same kind of interest in liberty to choose as humans.

³⁹ I take the name of this account from Duus-Otterström's formulation of the account: Duus-Otterström, "Freedom of Will and the Value of Choice," 265.

⁴⁰ Harris, "Consent and End of Life Decisions," 11; Young, "The Value of Autonomy," 43.

⁴¹ Duus-Otterström, "Freedom of Will and the Value of Choice," 266; Mill, "On Liberty."

⁴² Stephen Darwall, "The Value of Autonomy and Autonomy of the Will*," *Ethics* 116, no. 2 (2006): 267.

⁴³ Duus-Otterström, "Freedom of Will and the Value of Choice," 266.

their own choices and the direction of their lives.⁴⁴ Through disrespecting agents in this way we wrong them.

Some have argued that having choice in itself, even if one uses one's freedom to choose to make bad choices, is pleasant or contributes to one's well-being in some way.⁴⁵ This may well be true. I won't say anything against this position here, however the kind of value I take choice to have on the Self-Determination Account doesn't depend on any such claim. On the Self-Determination Account choice is valuable even if it doesn't result in any contribution to one's well-being. Thus, regardless of whether being deprived of choice is a harm in the sense that it is detrimental to one's well-being, my argument here is that it is a wrong. Depriving an agent of liberty is *pro tanto* wrong on this account, though it may not lower one's level of well-being. Thus having the freedom to choose, on this account, is good in a *non-agent affecting way*. It is good that agents have choice, though it is not good *for them*, or at least not good for them in the narrower sense of contributing towards their well-being. As such, strictly speaking, agents do not have an interest in being free to choose. Although we may still continue to talk of liberty in this sense being 'in the interests' of agents, so long as we note that here we mean interest in a wider sense and not that having liberty to choose contributes to one's well-being.⁴⁶

Choice on this account is taken by many to be more valuable than well-being. As such even where making one's own choice leads to catastrophic harm being done to oneself, or the ending of one's life, many believe that choice still has value and that one ought to be able to choose to engage in such actions without the interference of others. Though this is commonly accepted, I will argue that this is not a necessary feature of the Self-Determination Account.

Most theorists who support a form of the Self-Determination Account usually believe choice to be valuable because it is linked closely to autonomy in the sense of being able to reflect upon, and revise one's desires, in line with some higher-order values or goals.⁴⁷ Determining one's own life is usually assumed to mean determining one's own life according

⁴⁴ Darwall, "The Value of Autonomy and Autonomy of the Will*," 268.

⁴⁵ Duus-Otterström, "Freedom of Will and the Value of Choice," 262; Young, "The Value of Autonomy," 39; Regan, *The Case for Animal Rights*, 91; Rowlands, *Animals Like Us*, 153; Thomas, *Animal Ethics and the Autonomous Animal Self*, 90–91.

⁴⁶ I will in places, refer to agents as having an interest in liberty in this way, simply for ease of writing.

⁴⁷ Donaldson and Kymlicka, Regan and Thomas are exceptions here. However Donaldson and Kymlicka provide no sustained argument for this position and Regan and Thomas take the value of choice to be reducible to well-being. Donaldson and Kymlicka, *Zoopolis*, 108–12; Regan, *The Case for Animal Rights*, 91; Thomas, *Animal Ethics and the Autonomous Animal Self*, 90–91.

to one's conception of the good life, or according to one's will or ultimate values.⁴⁸ Young poetically captures the heart of this position when he states that: 'To be content or happy is desirable, but autonomously to have been its architect and builder is better.'⁴⁹ Whilst I accept that determining one's life through using one's capacity for autonomy is valuable, I think most theorists have overlooked the fact that choice has this kind of value even where autonomy is not involved. As such agency, not just autonomy, is valuable because it allows agents to have choice and determine the course of their lives, even though it doesn't allow agents to determine their lives to *the same extent* as autonomous agents. Thus, I will argue, agents, regardless of their other capacities, have the same kind of interest in liberty as autonomous agents. Agents have an interest in liberty in itself because self-determination is good and having one's liberty restricted fails to respect the authority agents have to determine their own lives.

Consider the case of the individual who completely lacks a life plan, they have no thought-out conception of a good life. Every day they get up and just do what they happen to feel like. Most of the time this is watching television or going to see their friends. This individual is perfectly capable of autonomy. If they wanted to, they could form a life plan or conception of a good life and start to bring their first-order desires in line with this plan, making use of their autonomy. However they never do so, in fact, they never even consider doing so. Further, let's say that they have always acted in this way. It seems obvious that this individual ought to be *pro tanto* free to make their own choices.

However if exerting self-determination over one's life means strictly determining the course of one's life in line with one's own conception of the good life or one's will, etc. then this individual's freedom to choose cannot be valuable on the Self-Determination Account. This individual lacks any conception of a good life or higher-order life plan and acts merely on their first-order desires. Nonetheless, their choices still seem intrinsically valuable in much the same way that the Self-Determination Account suggests. Thus it seems that autonomy isn't relevant to the question of whether one should have the freedom to choose. Although autonomy is relevant to the *degree* of freedom one should have, and the available methods one

⁴⁸ Harris, "Consent and End of Life Decisions," 11; Frey, "Autonomy and the Value of Animal Life," 61; Cochrane, "Do Animals Have an Interest in Liberty?," 665; Alasdair Cochrane, *Animal Rights without Liberation: Applied Ethics and Human Obligations* (Columbia University Press, 2012), 11; Duus-Otterström, "Freedom of Will and the Value of Choice," 266; Darwall, "The Value of Autonomy and Autonomy of the Will*," 267; Young, "The Value of Autonomy," 43.

⁴⁹ Young, "The Value of Autonomy," 39.

can make use of to determine the course of one's life, one can determine the course of one's life in the relevant way without autonomy.

One may note that though this individual doesn't *exercise* autonomy they do have the *capacity* for autonomy. Thus, one might think that whilst it isn't necessary to exercise autonomy over one's life in order to engage in self-determination of the relevant kind, it is necessary that one possess the capacity for autonomy in order to engage in self-determination of the relevant kind. Thus freedom of choice is valuable for both individuals that exercise autonomy and individuals that have the capacity for autonomy but fail to exercise it. I doubt that this is the case; consider another example.

A different individual similarly lacks a life plan and a clear thought-out conception of a good life. Just like the individual who does not exercise their capacity for autonomy, this individual makes decisions 'on the fly'. Most days they choose to watch television and see their friends. This individual however has significant intellectual disabilities and lacks a capacity for autonomy. As such, they are unable to form a conception of a good life or any higher-order desires. They can make many everyday choices and, with the help of support workers, friends and family they can make many more. Again it seems that choice has value for this individual. They, like those that do not exercise autonomy, should have the freedom to make choices. Or at least, they should have this freedom to some extent. We may think, since they lack certain intellectual capacities to grasp the significant implications of some choices, they ought not to have absolute freedom with regard to self-affecting choices like autonomous agents. They should at least be limited from ending their own lives and inflicting significant harm upon themselves.⁵⁰

It isn't clear why the mere capacity for autonomy should make the freedom to choose valuable if one neglects to use this capacity. The individual that possesses autonomy but fails to exercise it engages in *exactly the same* actions as the agent who entirely lacks the capacity for autonomy. It seems odd to suggest that the former determines the course of her life in a way that the latter does not and cannot. Both individuals determine the course of their respective lives on the basis of mere first-order desires. Thus, if exercising autonomy is not a pre-requisite for determining the course of one's own life, it doesn't seem that possessing the capacity for autonomy should be either (other things being equal). It seems their choice has value in itself because it allows them to determine their lives and agents lacking the capacity for autonomy,

⁵⁰ I will say a little more about the extent of the freedom such individuals are owed below.

just like those that possess the capacity for autonomy, should still be *pro tanto* free to make a great many choices about their life.

This is an intuitive and widely supported view in the case of humans lacking autonomy. Support workers, carers and nurses try their best to maximise the amount of control that their patients and service users have over their own lives.⁵¹ Moreover, here in the UK, this position is enshrined in law through the Mental Capacity Act 2005. The Mental Capacity Act states that individuals that lack the capacity to make some decisions, even particularly important ones, such as whether to accept medical treatment, still have the legal right to make choices in other areas of their life.⁵² One can only deprive an individual lacking mental capacity of their legal right to liberty, where doing so is necessary to ensure their continued existence or prevent significant harm.⁵³ Even where one lacks mental capacity, one has the legal right to make choices that may not contribute towards one's well-being. What's more, where decisions do have to be made on one's behalf they must still take heed of the individual's views and minimise restrictions to one's liberty.

Seemingly then, choice has value for non-autonomous agents, and it has value because it gives them self-determination over their own lives. If this is correct then what is the appropriate way to understand self-determination? In order to try to give an answer to this question I'll return to Young's analogy of the architect and builder. Young claims that being the builder and architect of one's happiness is better than merely having one's happiness provided for one. Despite being categorised together here, a builder and an architect are roles which are usually occupied by different persons. The architect holds the majority of creative control over the project. They set the overall goals of the project. They decide if they are designing a house, a cathedral, a library, etc. Further they draw out the plans for the project and design how the building will look. We play the role of the architect in our own lives when we exercise our autonomy and form a conception of the kind of life we want to live. We have ultimate creative control over our lives. Builders however perform a different role. They follow the plans given to them by the architect and aim to bring about the completion of the project. They actually build the project. We take on the role of builders in our lives when we make the smaller scale choices in our lives. For instance, when we choose to skip dessert

⁵¹ Donaldson and Kymlicka give a good overview of some of the applied philosophical work in disability studies that supports this claim. Donaldson and Kymlicka, *Zoopolis*, 105-8.

⁵² *The Mental Capacity Act 2005*, <http://www.legislation.gov.uk/ukpga/2005/9>.

⁵³ *Ibid.*, sec. 4B.

because our life plan involves staying healthy and we've already had a sugary snack at lunch time. Thus we fulfil both the role of architect and builder in determining our lives.

Mere agents, lacking the capacity for autonomy, cannot be the architects of their own life. However they can (and do) play the role of builders. They make the smaller scale choices about their life and actually construct a life. This may seem insignificant at first, however one shouldn't be dismissive about the value of this role. Builders don't slavishly follow the plans of architects. They must make various decisions themselves that will affect the outcome of the project. For instance, if you were a builder in the mid to late 19th Century, working on one of the many neo-gothic civil buildings that were being built all over the country, you would have had many choices to make. As a builder you would have the usual decisions associated with interpreting a plan and perhaps choosing fixtures and fittings. However, the neo-gothic architecture tradition, unlike the classical architecture tradition, supposedly recognised the skills of the individuals working on the project and so devolved authority over the design of various elements to the builders themselves. Thus on a neo-gothic project, the design of the fine details of the building such as the patterns found in a frieze, column capitals and any other smaller decorative elements, would be left to the builder.

In this way, being a builder brings with it opportunity to shape the project. Of course one could not have as much input as the architect in the overall look of the project. One could not decide to build a church instead of a town hall, or move the building to a new site, but one would, within the remit of the project, be able to make it one's own. Thus though it is better to be an architect and builder of one's happiness than a mere receiver, it is plausibly better to be a mere builder than a mere receiver. Though not capable of formalising a conception of a good and acting in accordance with it, agents lacking autonomy can still 'build' a good life. They can, within the confines of their project so to speak, make choices on a smaller scale which affect their lives. So choice still has value for those that lack autonomy and thus being free to exercise one's choice is plausibly a good for agents.

Of course, one may remark, builders are useless without architects! Even if they make some choices themselves, without some plan to work towards, builders cannot build. It doesn't make sense to think of mere agents working towards some conception of a good life since they lack a plan to follow. Thus, one might think, this analogy falls apart under scrutiny. It seems one has to be an architect-builder in order for one's freedom to choose to be of value. However this is not so. It is thought that long before the Victorian gothic revival, when the medieval gothic churches were being built across Northern Europe, the role of an architect

simply didn't exist. Builders themselves were responsible for whole building projects from conception to completion. However they didn't draw plans or design the buildings in advance. Builders designed as they went. They did not conceptualise the overall project but formed the project through considering each smaller choice in term. Supposedly such builders even regularly left and returned to sites after long periods away whilst other builders continued work on the project. The returning builders would simply continue construction on the structure that they were presented with as they saw fit, with little to no information from the previous set of builders. Through this method they managed to construct a building without any overarching plan for the project.⁵⁴

This analogy isn't perfect but hopefully it makes my point sufficiently clear. Just as the medieval builders built without a formal plan or any individual fulfilling the role of architect, agents can exercise self-determination over their lives through making choices on a case by case basis, without any formal conception of a good life. Through the choices they make every day, they exercise self-determination over their lives just as autonomous agents do. Of course they do not exercise self-determination *to the same extent* as autonomous agents, they can only choose between acting on various desires, nonetheless they can still build a life, just as autonomy-exercising agents can. That they do not determine their lives through the same process as individuals that exercise autonomy does not matter. Even agents that are capable of autonomy may not determine the direction of their lives through using life plans but may, like agents lacking the capacity for autonomy, rely on determining the direction of their life 'on the fly'.

Importantly however, the extent of one's freedom to choose should be limited by one's capacities to understand the implications of one's choices. Self-determination is one value among others. As such, the value of being free to determine one's life must be weighed against one's overall well-being. Where one realises, or is capable of realising, the implications of decisions that could cause them significant harm, we take them to be free to engage in such actions. However where an individual lacks the capability to realise these implications, they should not be free to pursue a course of action which could cause them significant harm. This distinction, between those that have these capabilities and those that do not, doesn't appear to line up exactly with the capacity for autonomy. We may for instance think that we shouldn't allow someone who is under the influence of drugs to make decisions that will cause them

⁵⁴ For my purposes it doesn't matter whether this actually was the case. Such a case is certainly possible.

significant harm, even if they have the capacity for autonomy, because they are not currently acting autonomously.

Such restrictions of liberty appear to me to be justified. The precise circumstances for justified paternalism are not something I can defend here. However neither is a defence of these conditions necessary to make my point. It is sufficient for my argument to note that it seems that non-autonomous agents ought never to be free to choose to engage in actions which would cause them significant harm. Thus, it seems we should act in line with the Mental Capacity Act. Non-autonomous agents should be free to make poor choices, however where an agent is at risk of inflicting significant harm upon themselves, their liberty should be restricted. Nonetheless, paternalist intervention should still be through the least restrictive means possible so as to minimise the effect of such invention upon the determining of their lives.

There is of course a question of when a poor choice becomes a significant harm. I will not attempt an answer to this question here. There are borderline cases in many moral problems so the fact that there is no straightforward answer to this question shouldn't count against my argument. It is sufficient for my purposes that we can distinguish some cases where agents should be free to make their own choices from others in which we would be justified in acting paternalistically. As such, though there may be a grey area in between, there should be a presumption *against* interfering with mere agents' choices unless we have reason to believe that they will inflict significant harm upon themselves.

One should also note that just because non-autonomous agents are owed the freedom to make choices, including poor choices, this doesn't mean that we are obligated to leave them entirely to their own devices. Where such agents are in our care, it seems we should plausibly help them to realise a good life. Allowing them to make their own choices doesn't mean that it is wrong for us to persuade and encourage them to make choices that will benefit them. There does not seem to be anything wrong in trying to encourage a severely autistic person in my charge to try a new food or encouraging him to walk a different route. So long as at bottom he still has the choice, such treatment seems like precisely the right thing to do.⁵⁵

⁵⁵ Donaldson and Kymlicka touch on this point. Donaldson and Kymlicka, *Zoopolis*, 108–12.

viii. Animals and Humans

Now, if my argument here is correct, and freedom of choice has value for mere agents in virtue of granting them the ability to determine their lives, then choice will have value for animals as well as humans. In virtue of being agents, animals can determine the course of their lives through a piecemeal process, just as non-autonomous humans do. Animals just like non-autonomous humans must be the builders of their own lives. One can only resist this conclusion if one can cite some morally relevant difference which can account for a difference in interests despite a similarity in capacities. To reject that choice has value in animals, yet accept that it has value in non-autonomous humans, without any morally relevant difference, would be inconsistent and speciesist. Hopefully from I have said this argument will not seem unreasonable. Ideally, this argument will be wholly convincing, however I suspect that there will be some resistance to this claim.

One possible parry here is to defend an alternative explanation for why we grant humans choice even when they lack autonomy. Further this explanation does not demand that we owe animals the same treatment. This explanation states that since non-autonomous humans are members of our society, the same principles of dignity that apply to autonomous individuals, apply to them, regardless of whether choice actually has value for them. Thus, even though choice doesn't have value for non-autonomous humans themselves, they are entitled to the same degree of respect and freedom of choice as autonomous humans.

There are several problems with this kind of argument however. Firstly in this objection there is an implicit assumption that animals are not members of our society. Though they lack recognition as legal citizens, Donaldson and Kymlicka have recently argued that at least domestic animals fulfil all three of different conditions for citizenship. They have a right to reside in a particular territory, the state governs on their behalf and they possess political agency (roughly the ability to be affected by the law and 'voice' preferences regarding their treatment - though they need not understand the role of the law in bringing about this treatment).⁵⁶ If animals are members of our society then they too should surely be extended the dignity of freedom of choice.

A bigger problem with this alternative explanation however, is that even if animals aren't members of our society, it just doesn't seem that the freedom to choose is owed to non-autonomous humans merely because of their societal relations to us. We wouldn't think it any

⁵⁶ *Ibid.*, 57-58.

more acceptable to control a non-autonomous individual's life were they not a member of our society. Imagine a nomadic non-autonomous human, free from any society, it doesn't seem that we can permissibly control their life any more than we can control the life of a non-autonomous human in a local care home. It doesn't seem to be a mere courtesy that we owe to non-autonomous agents choice. It seems like something we are morally required to do in virtue of the kind of beings that they are: agents - beings that are capable of making choices. The more plausible, and I think more intuitive option, is that it is good in itself that non-autonomous agents have choice, and good because it gives them control over their lives.

One might find my argument here unconvincing and fail to see the wrong done to non-autonomous individuals through paternalistic control. As such one might be willing to concede that choice has value neither for non-autonomous humans, nor animals, and as such there is no inconsistency in one's views. I however, like many others, have deep intuitions that we wrong non-autonomous humans when we fail to grant them a limited freedom of choice. Further, I think that this is because choice has value in itself. The Self-Determination Account offers the best account of this value and thus I maintain that choice is of value in animals too, in virtue of them being choice-makers in just the same way as non-autonomous humans.

As such both humans and animals have the same kind of interest in liberty. Liberty to choose is valuable for both humans and animals because it allows them determine the direction of their own lives. Humans and animals do not have interests of the same strength in freedom. In virtue of being capable of grasping the implications of their choices which could cause significant harm, many humans have a stronger interest in liberty. Additionally liberty may be valuable for humans in other ways that it is not for animals, nonetheless, the central kind of interest in liberty that autonomous humans possess, is shared by animals.

ix. Practical Implications

One may point out that my argument will have little impact on our treatment of free wild animals or captive wild animals since humans tend not to interfere with their choices anyway. This may be true in some cases (though certainly not all) but it seems likely that this is for convenience and financial reasons, rather than the reasons I suggest here. As such, my argument gives moral reasons in support of non-interference in agents' lives and makes such actions morally required. Moreover, this argument will have significant implications for our

treatment of domestic animals. Assuming the argument is correct, there would be a *pro tanto* obligation to give one's companion dog meal options, feed them when they want to be fed rather than instigating a meal time, allow them to take initiative when walking and let them lead the walk to a place of their choosing. In short, help them have more control over the decisions in their life.

Such agency-promoting conditions may seem incredibly difficult to put into practice, especially for those of us who lack the time, money or knowledge of how to go about increasing independence in our companion animals. Nonetheless, we should do what we can, and where one can promote independent choice making, one should do so. There are at least some cases where allowing such choice isn't going to be particularly burdensome at all, allowing a dog to choose which way to go on walk for example. The reason that promoting independent choice making seems demanding, at least in part, may be because such duties have been neglected for so long. Moreover, the fact that such actions may be demanding of our time or other resources isn't sufficient reason to forgo engaging in them. There are similar significant resource costs to giving non-autonomous humans choices and allowing them to do things themselves, rather than simply providing them with goods. Nonetheless, having the freedom to make their own decisions is considered important enough to warrant these costs. As such, it appears the same should be done of animal agents in our care.

One may question whether many animals are capable of such independence or how we can possibly give animals the choices discussed. I will not here consider the practical methods available which can be utilised to achieve this end but simply note that attending to animals' behaviours can give us significant insight into their preferences. Donaldson and Kymlicka suggest some interesting cases and methods that show such promotion of agency is certainly possible and not all that difficult.⁵⁷ Furthermore, this problem is present in the case of non-verbal, non-autonomous humans yet we manage to encourage independent decision making in these individuals. Thus the practical problems of granting freedom to choose are not insurmountable and they do not give us sufficient reason to withhold such a freedom from agents.

Through this section I have considered various arguments for the value of liberty. I have shown previous arguments for the instrumental value of liberty in virtue of it granting one desire satisfaction, or allowing one to fulfil natural functions, only show the value of a

⁵⁷ Ibid., 108-12.

very limited degree of liberty. I then considered two accounts of the value of choice with the aim of establishing the value of a less restrictive kind of liberty. I argued that though the Prudential Account and the Self-Development Account do capture some of the value of choice, it seems that there is something intrinsically valuable about choice itself. I argued that the Self-Determination Account captures this intrinsic value.

According to the Self-Determination Account agents should be free to make their own choices because this grants them determination over the course of their own lives. I argued that autonomy is not necessary in order to exert this kind of self-determination, and that whilst autonomous agents may exert more determination over their lives, non-autonomous agents still nonetheless exert the *same kind* of self-determination over their lives. Thus, liberty is good in the case of autonomous agents and non-autonomous agents. It is not *good for them* however but good in a non-agent-affecting way. As such, whilst agents are not harmed by having their liberty restricted where their well-being is unaffected, they are nonetheless wronged. It is *pro tanto* wrong to restrict the liberty of agents to engage in self-regarding actions (except in the case of non-autonomous agents where there is a risk of significant harm).

The duty of non-interference that we owe to mere agents does differ in degree from that which we owe to autonomous agents, but it does not differ in kind. Just as it is good that autonomous agents are free to act without interference, it is good for the very same reasons (though to a limited extent), that non-autonomous agents are free to act without interference. Finally, from this it seems that we can conclude that the possession of autonomy is unnecessary, and that the possession of agency is sufficient, to secure the kind of freedom from interference that we ordinarily consider to be exclusively possessed by full moral status holders.

Conclusion

In this chapter I argued that agency is necessary for moral status because one could not possess interests without agency, and without interests one couldn't be morally considered. I then argued that if the central interests of full moral status holders (freedom from pain, liberty and continued existence) are possessed by mere agents, then mere agents must also possess full moral status. I have shown that agents have an interest in being free from pain because pain is the frustration of certain final desires that embodied agents necessarily possess. I also argued that agents have interests in liberty. I argued that it is good

that mere agents are free to make their own choices (so long as their choices do not result in significant harm) because being free to determine the course of one's life is good. Whilst this may not contribute to one's well-being, I suggested that it is good in a non-agent affecting way that one has this kind of liberty. I argued that though liberty is usually only thought to have this kind of value when possessed by autonomous agents, it actually has this same kind of value when possessed by non-autonomous agents. Though non-autonomous agents do not follow a life plan they can still determine the course of their lives in a meaningful way. In the next chapter I will show that agents have an interest in continued existence and conclude my argument that agency is sufficient for full moral status.

Chapter 6

The Agency Account of Moral Status: Continued Existence



Continuing my argument from the previous chapter, I will here argue that mere agents have an interest in continued existence. I will do this by arguing that having certain final desires makes death bad for an individual. Further, agents, in virtue of being the kind of individuals that are able to (and almost always do) possess desires of this kind, are therefore harmed by death and are wronged when they have their lives ended. I will start by giving a brief account of my positive argument in defence of this claim. Then I will defend this argument against several objections that contend that some further psychological capacity or ability is necessary for death to harm one. In defending my view against these objections I will further elucidate and support my account of the harm of death.

Once I have completed this section on continued existence, I will have shown that mere agents possess the three central interests of full moral status holders. Then, I will give an argument to show that the central interests of autonomous agents are also grounded in agency, not autonomy. As such all agents, autonomous or not, possess full moral status in virtue of possessing agency. Therefore, agency is a sufficient and necessary condition for moral status. I will then end this chapter by considering some objections to this claim.

I. Continued Existence

The most widely accepted account of the badness of death is *the Deprivation Account*.¹ According to the Deprivation Account, death is not intrinsically harmful itself (though the process of dying often involves pain and suffering) but is only an extrinsic harm. Death is bad for us (when it is bad for us) because it deprives us of some goods. To use Broome's words:

¹ Thomas Nagel, "Death," *Nous* 4, no. 1 (1970): 73–80; James W. Yeates, "Death Is a Welfare Issue," *Journal of Agricultural and Environmental Ethics* 23, no. 3 (2010): 229–41. As well as those cited below who endorse a (wholly or partly) desire-based form of Deprivation Account.

'The badness of your death for you is the difference between the goodness of the longer life you would have led had you continued living and the goodness of the life you actually do lead.'² The Deprivation Account entails that where one has no goods of which to be deprived, death does not harm one. As such, as well as plausibly explaining the badness of death, this position can accommodate the permissibility of voluntary euthanasia and rational suicide under certain circumstances.

Whilst there is general agreement regarding death being an instrumental harm in light of the goods it deprives us of, there is of course disagreement regarding what goods death robs from us. One's preferred account of well-being will, for the most part, determine one's position here. However, one interpretation of the Deprivation Account that is particularly popular is a *Desire-Based Deprivation Account*.³ Such accounts are also the most widely used form of Deprivation Account to defend the claim that animals have an interest in continued existence.⁴

The Desire-Based Deprivation Account is as follows:

- 1) Satisfying some subset of desires is good for us.
- 2) It is bad for us to be deprived of goods.
- 3) Death stops us satisfying these desires (and thus possessing goods).

² John Broome, "The Badness of Death and the Goodness of Life," in *The Oxford Handbook of the Philosophy of Death*, ed. Ben Bradley, Fred Feldman, and Jens Johansson (Oxford: Oxford University Press, 2013), 225.

³ Bernard Williams, "The Makropulos Case Reflections on the Tedium of Immortality," in *Problems of the Self* (Cambridge: Cambridge University Press, 1973); Jeff McMahan, "The Comparative Badness for Animals of Suffering and Death," in *The Ethics of Killing Animals* (Oxford: Oxford University Press, 2016), 70; Stephan Blatti, "Death's Distinctive Harm," *American Philosophical Quarterly* 49, no. 4 (2012): 317–330; Fred Feldman, *Confrontations with the Reaper: A Philosophical Study of the Nature and Value of Death* (Oxford University Press, 1992); Susan Luper-Foy, "Annihilation," *The Philosophical Quarterly* 37, no. 183 (1987): 233–52.; and many others.

⁴ Elizabeth Harman, "The Moral Significance of Animal Pain and Animal Death," in *The Oxford Handbook of Animal Ethics*, ed. Tom L. Beauchamp and R. G. Frey (Oxford: Oxford University Press, 2014), 730; Aaron Simmons, "Do Animals Have an Interest in Continued Life?," *Environmental Ethics* 31, no. 4 (2009): 389; David DeGrazia, *Taking Animals Seriously: Mental Life and Moral Status* (Cambridge; New York: Cambridge University Press, 1996), 231–32; Tom Regan, *The Case for Animal Rights* (Berkeley: University of California Press, 1983), 99–100; Bernard E. Rollin, *Animal Rights and Human Morality* (New York: Prometheus Books, 1981), 48; Rosemary Rodd, *Biology, Ethics, and Animals* (Oxford: Clarendon Press, 1992), 126; Alasdair Cochrane, *Animal Rights without Liberation: Applied Ethics and Human Obligations* (Columbia University Press, 2012), 65; Peter Singer, *Practical Ethics* (Cambridge: Cambridge University Press, 1993), 125–26; McMahan, "The Comparative Badness for Animals of Suffering and Death," 70; Alastair Norcross, "The Significance of Death for Animals," in *The Oxford Handbook of Philosophy of Death*, ed. Ben Bradley, Fred Feldman, and Jens Johansson (Oxford: Oxford University Press, 2013), 466; Mark Rowlands, *Animals Like Us* (London: Verso, 2002), 76–79.

C) Death is bad for us because it deprives us of valuable desire satisfaction.

Since this account is fairly widely accepted and it has significant intuitive appeal, I will adopt the Desire-Based Deprivation Account of the badness of death. I maintain that death is bad for an individual where it deprives one of the satisfaction of certain final desires.⁵ Thus I will demonstrate that agents can have an interest in continued existence because they are able to possess the relevant final desires, the satisfaction of which, death deprives one. This being said one should not assume that my argument here entails that desire satisfaction is the *only* good available in life, or that this is all that death deprives one of. I aim to delineate the minimally sufficient conditions for death being bad for an individual. As such, one may be harmed in other ways by death apart from being deprived of relevant final desire satisfaction. However I take it that unless one is capable of being deprived of relevant final desire satisfaction one cannot be deprived in any other way.⁶ As discussed elsewhere, I take it that unless one has desires then one does not have interests, though if one does possess desires one may have further interests that are not based on these desires e.g. an interest in remaining healthy.⁷ My reasons for defending this particular formation of the Deprivation Account will become clear as I argue against other rival accounts but I shall give a brief defence of the view here.

Firstly, as Bradley rightly notes, it is intrinsic desires that are usually considered to be important to one's well-being according to those that emphasise the importance of desire satisfaction on well-being⁸ thus plausibly, it is this kind of desires that would be relevant to the badness of death. Bradley takes 'intrinsic desires' to be what I refer to as 'final desires' in Chapter 3. These are desires that one has for an object in itself and *for no further reason*.⁹ These are the kind of desires that I have argued are necessary for the possession of sentience. To recapitulate, where final desires are unfulfilled (and one remains alive) they result in

⁵ Broadly I agree with Bradley's view on the relevant kind of desires needed to make death bad for one. Ben Bradley, "Is Death Bad for a Cow?," in *The Ethics of Killing Animals* (Oxford: Oxford University Press, 2016).

⁶ An exception to this might be if one can be harmed merely because one is a mortal being and is always faced with the threat of death as Blatti suggests. Although Blatti's argument anchors on the notion of autonomy it is needless to say that it seems one could make his argument regarding agency more generally. Nonetheless I set aside this form of harm here. Blatti, "Death's Distinctive Harm," 324.

⁷ Chapter 1 Sec.I

⁸ *Ibid.*, 55; Richard B. Brandt, *A Theory of the Good and the Right* (Oxford : New York: Oxford University Press, 1979), 111; Chris Heathwood, "The Reduction of Sensory Pleasure to Desire," *Philosophical Studies* 133, no. 1 (2007): 23–44. See Chapter 3.

⁹ For clarity and consistency I will thus refer to the desires relevant to being harmed by death as 'final desires'.

unpleasantness. However other desires may go unfulfilled without it resulting in any unpleasantness, if one's guiding final desire is still satisfied.¹⁰ For instance if I have a final desire to have something to eat, and an instrumental desire for pizza, where it is the case that pizza was unavailable but burritos were available, I could frustrate my instrumental desire whilst still satisfying my final desire. This would not result in any unpleasant feelings and would in fact result in pleasant feelings since my final desire is satisfied.

Secondly, the relevant final desires required for death to harm one, are final desires that cannot be frustrated by death. It cannot be that the deprivation of any old final desire satisfaction will harm one. If one has a final desire to die, clearly death would not deprive one of the satisfaction of this desire.¹¹ Further if one had a final desire to never experience war, death would clearly not frustrate this desire either. If I have never experienced war and I die, then my final desire to never experience war is satisfied. Further, if it was the case that had I not died, I would have gone to war, then not only is my final desire satisfied, but it is satisfied by death. Death does not deprive one of the satisfaction of all final desires then. Death cannot deprive one of the satisfaction of a final desire for death, or a desire for something that death will bring about. In cases where death doesn't deprive one of the satisfaction of one's final desires, death could not plausibly be said to harm one. Therefore being harmed by death requires that one has final desire(s) that do not have content such that they may be satisfied by one's own death.

An implication of Desire-Based Deprivation Accounts is that death harms different individuals differently (or not at all in some cases). The harm one suffers through death will depend upon what one hopes for, aims for and achieves through one's life. This I think is correct and intuitive; death doesn't seem to harm us all equally. Thus this account shows us that we should consider the harm of death for some individual to be determined by facts about the individual themselves, rather than relying on assumptions regarding species as wholes. Not only is this intuitive, but it is in line with moral individualism, which as I have argued, we should endorse.¹²

Hence, on this account it is in principle possible that an animal could be harmed by death as much as a human (and perhaps even more than a human) whilst still accepting that in general, humans have more to lose from death than other animals. Whilst my main aim

¹⁰ I discuss this further below.

¹¹ Bradley, "Is Death Bad for a Cow?," 55.

¹² Chapter 1 Sec.IV.

here is to give an account of the badness of death that can explain how humans and animals are harmed by death *in the same way*, I will spend some time considering the comparative badness of the ending of animals' lives and ending humans' lives. Whilst I take it that determining the badness of death for an individual is a question for applied ethics, I will give some argument for the purpose of showing that there is no great gulf in the badness of death for humans and animals. Though death may be worse for humans on the whole, death is by no means insignificant for animals either. Without further ado I will now consider some counter-arguments to this account.

i. Future-Directed Desires

One kind of desire that some have argued is necessary in order to be harmed by death is a future-directed desire. Death deprives us of our future lives, thus without the satisfaction of our desires in our future, it is suggested, we cannot be harmed by death. Further it has been argued that unlike humans, animals live 'moment to moment', lacking any ability to think about the future and having only an awareness of the here and now. Thus, they must be incapable of possessing future-directed desires and so have nothing to lose were their lives to end at any given moment. Therefore death cannot harm them through the deprivation of future desire satisfaction.¹³

This argument rests upon two principal claims:

- 1) One must have future-directed desires in order to be harmed by death.
- 2) Agential animals lack the capacities to possess future-directed desires.

If either of these claims are false then this argument fails. I will argue that in fact both of these claims are false. I will start with the latter.

¹³ Christopher Belshaw, "Death, Pain and Animal Life," in *The Ethics of Killing Animals* (Oxford: Oxford University Press, 2016), 38-39; Singer, *Practical Ethics*, 118-19; McMahan, "The Comparative Badness of Animals of Suffering and Death." It is worth noting that those who argue in support of this claim often note that this is not the case for *all* animals. It is often conceded that apes and cetacea, for instance, may possess future-directed desires. A modification of this view which I discuss below is that animals do not hold future-directed desires that are sufficiently long-term for them to be significantly harmed by death. As I will show, if one wants to avoid implausibility one must argue for this modified view rather than the more hard-line view that animals lack *any* future-directed desires. Whilst Belshaw, at times, appears to argue for the more hard-line thesis, McMahan clearly argues for the weaker view and Singer seems to argue for both at different times.

Belshaw believes that animals must lack future-directed desires because they are *psychologically unconnected* to their past and future selves.¹⁴ By this Belshaw means that the psychological being that exists in the body of an animal at one moment in time is completely psychologically separate from the psychological being that will exist in the same body in the next moment of time.¹⁵ There is no continuation of desires, beliefs or mental states of any kind. In his own words, Belshaw describes the lives of animals as ‘more or less discrete moments’ and goes on to suggest that one may think of the life of an animal as ‘in effect, a series of discrete lives’.¹⁶ For Belshaw, most animals are not biological beings with psychological identities that exist over the lifetime of the organism but rather are psychological individuals that come in and out of existence momentarily, a single biological organism housing many such individuals over its lifetime. I will refer to this view as the *Momentary Psychology Account* or MPA.

If we interpret MPA as meaning that an individual animal only ever exists for a single instance in time, and doesn’t persist through time *at all*, it would plausibly rule out the possibility of such beings having *any* desires. Acting on a desire and even coming to possess a desire are events that happen *over time*. Thus if animals have any desires at all, then it seems they must persist through some period of time.¹⁷ Since I have shown that all sentient beings have some desires, this account cannot be right. Thus such a literal understanding of MPA must be false.

A more charitable interpretation of MPA is that animals do not exist only for mere instances of time, but nonetheless that they exist only over short periods of time. It is unclear which position Belshaw holds since he notes that animals have ‘some rudimentary future-directed desires’ and that ‘A dog for instance wants its master to come home and take it for a walk’¹⁸ but elsewhere suggests that a cow couldn’t have a future-directed desire to ‘eat grass over there’.¹⁹ Others however (specifically Singer and McMahan at least) endorse this latter

¹⁴ Belshaw, “Death, Pain and Animal Life,” 37.

¹⁵ I’ll discuss the notion of psychological connectedness further below.

¹⁶ Belshaw, “Death, Pain and Animal Life,” 43.

¹⁷ One might also note that seeing as brain activity occurs over time (even if it is very short periods of time), if brain activity plays some role in mental activity, which seems a fairly safe assumption, then mental activity must also occur over some period of time. As such, one couldn’t possibly possess desires if one’s life was truly instantaneous.

¹⁸ Belshaw, “Death, Pain and Animal Life,” 37.

¹⁹ *Ibid.*, 39. Belshaw’s point here seems to confuse his argument. He accepts in this passage that a cow could desire to eat grass over there *now*, be motivated by this desire and set off to satisfy this desire, yet denies that a cow could desire that she eats the grass over there at some point in the future. Whilst there is clearly a difference in desire content between these two examples it doesn’t seem that the first

formulation of MPA, accepting that animals at least have psychological identities extending over hours.²⁰ Thus it is worth considering. I will show that this formulation of MPA is also false and that many animals possess psychological identities extending much longer.

As I will argue in the next sub-section, many animals demonstrate memories of significant lengths. Not only are many animals able to remember what happened days ago but also important information from months or years past. If animals are connected not only to their immediately past states, but also to long past states, it seems foolish to rule out their connectedness to more long-term future states in principle. If animals can have long-term memories perhaps they can also possess non-immediate-future-directed desires. One should not misunderstand my point here. Evidence of long-term memories doesn't itself give us evidence for non-immediate-future-directed desires but evidence of long-term memories does show that animals are not merely 'living in the moment'. This is a necessary condition for possessing future-directed desires. However in order to show that any animals *actually* possess such non-immediate-future-directed desires we will need to look at empirical evidence.

Many animals engage in actions such as hunting, tracking, migrating. These actions require one to set oneself to the task at hand (or paw) for hours to days. Therefore one must, if one is capable of action, be capable of future-directed desires. As Regan notes, wolves run for hours to get to a specific location where they wait for prey.²¹ If during a hunt a wolf were to move from one discrete life to another they would clearly be unable to complete their activity, or at least complete it with the efficiency that wolves do. Bradley gives us the example of a cow jumping a fence and running quite some-distance from a slaughterhouse.²² Squirrels and scrub-jays (among other birds) cache food with the plan to retrieve them at a later date. Weaver birds build complex nest structures that plausibly require desires or plans that are

formulation would be a non-future-directed desire. The cow must realise to some extent that she cannot fulfil her desire for the grass over there instantaneously, and that she can only fulfil it at some point in the future. She understands on some level that satisfying this desire requires her to act in some way so that some future will come about (though obviously not in such a conceptual or abstract way; probably not even in a conscious way). If she doesn't understand this, then it is mysterious why she acts at all rather than merely standing motionless, surprised that her desires are not satisfied. Heeding this point we could rephrase Belshaw's example of the dog's future-directed desire; the desire that 'its master to come home and take it for a walk' as a desire that 'its master come home and walk it *now*'. This would still plausibly be a future-directed desire despite the grammatical reformulation of the content expression (and Belshaw presumably recognises this).

²⁰ McMahan, "The Comparative Badness for Animals of Suffering and Death," 71; Singer, *Practical Ethics*, 110-19.

²¹ Regan, *The Case for Animal Rights*, 101.

²² Bradley, "Is Death Bad for a Cow?," 53.

directed toward the completion of the nest at some point in the future. Komodo dragons bite large prey, infecting them with the bacteria in their saliva and a slow-working venom. They then go and retrieve the corpse of the prey several days later. Such actions seem to imply a plan to eat their prey at some future point. Similarly the dog that dislikes bathing anticipates what will occur when his human companion gets the shampoo out and reacts in the way he does, plausibly because he desires not to have a bath.

These examples all give us desires directed at points of time in the future, ranging from minutes to months ahead of the present. Thus animals *can* possess future-directed desires and many animals are capable of non-immediate-future-directed desires. Of course how far into the future a being's desires stretch will depend upon one's individual mental capacities. Of course most animals are likely incapable of having future-directed desires as far into the future as us. Nonetheless at best, this means that humans can be harmed more by death but clearly animals and humans can both be harmed by the frustration of long-term desires. However, as I will argue now, it does not seem that future-directed desires are even necessary in order to be harmed by death.

Whilst it is the case that one needs to be capable of possessing desires in order to be harmed by death, one doesn't need to possess future-directed desires specifically or indeed any desires at all. Death robs us of future life and one's future life is good to some extent if it contains the satisfaction of some relevant final desires. However what will contribute to the goodness of my future life, will not only be the satisfaction of some of my current desires, but also the satisfaction of the desires I am yet to possess. It is for this reason that a seriously depressed person who lacks any desire except the desire to end their own life could still be harmed by death. If this individual would form some relevant final desires in the future, were they not to end their life, and would then go on to satisfy these desires, death *now* would deprive them of the satisfaction of these future desires. Thus death now, would harm them.²³ Therefore, the fact that I presently lack desires does not entail the death does not harm me.

What's more, this view has widespread intuitive appeal. Consider the case of an adolescent driver dying in a car accident. You would likely hear friends and family saying things such as 'It's such a shame, he didn't have a chance to raise a family or fall in love'. You

²³ Bradley, "Is Death Bad for a Cow?," 56; Ben Bradley and Kris McDaniel, "Death and Desires," in *The Metaphysics and Ethics of Death*, ed. James Stacey Taylor (Oxford: Oxford University Press, n.d.), 21; Norcross, "The Significance of Death for Animals," 486; Harman, "The Moral Significance of Animal Pain and Animal Death," 730.

might hear others say how it was a great misfortune that he never experienced career success or went to university. We tend to think that failing to achieve these things harms the adolescent whether or not he desired such things before he died. That he had never really considered whether he wanted a family, or whether he wanted to go to university, seems to be relatively unimportant here (so long as he would have come to desire such things). Death harmed him in part because it robbed him of the satisfaction of desires he had yet to come to possess.

Of course one can never know what desire satisfaction one's future life will contain. Thus in light of the barriers stopping us from knowing whether a future life contains desire satisfaction, it seems we should consider death harmful to an individual if it deprives one of *opportunities* for relevant desire satisfaction. Perhaps if the industrious workaholic woman hadn't died at the time she had, she would have had a stress related breakdown and resided herself to a reclusive life neglecting all her ambitions and dreams. Perhaps the apathetic man who wasted his younger years would be motivated to make something of himself, and have a greatly improved life, had he survived the car accident and not died. As such, we can at best consider the *opportunities* for desires to be satisfied rather than desire satisfaction directly. Taking heed of this point, we should take the adolescent driver to be harmed by death because he is deprived of *the opportunities* to form and satisfy the desire to raise a family, fall in love, go to university, etc.²⁴

However we should not take the deprivation of *opportunities* for any final desire satisfaction to be a harm since one almost always misses out on the satisfaction of some final desire through death.²⁵ Now I cannot outline a full method by which we can determine how much death harms one and in any case this is not necessary for my argument. I am here merely trying to specify what it means to have an interest in continued existence. However to shed some light on the topic, I suspect that probability is going to play a role in determining how much harm a given lost opportunity causes. In this way not all lost opportunities are equally harmful and some would not be harmful at all. The more likely it is that one would have acted on that opportunity, and satisfied the relevant desire, the more harmful deprivation of that opportunity is going to be for a given individual. Though I admit I do not have a procedure in mind to determine how likely possession and satisfaction of any given desire would be and

²⁴ Others too write of 'opportunities', 'possibilities' and 'the likelihood' of goods rather than goods themselves: Singer, *Practical Ethics*, 102; Regan, *The Case for Animal Rights*, 100; DeGrazia, *Taking Animals Seriously*, 236.

²⁵ Blatti, "Death's Distinctive Harm," 321.

rely mostly on intuition to this end. However a different, though related, approach might be to consider one's continuing life in the nearest possible world where one survives, as Luper-Foy suggests.²⁶

Thus, I have shown through this section that many animals do in fact have future-directed desires and that many of these go further than merely their immediate future of the next few moments. More importantly however I have shown that one needn't actually possess any future-directed desires in order for death to be bad for one. Death is a harm to one if there is final desire satisfaction in one's future, even if one doesn't currently possess the desires that will be satisfied in one's future.

ii. Psychological Connectedness and Time-Relative Interests

McMahan suggests another problem with animals being harmed by death. Whilst McMahan accepts that animals can be harmed by death, he argues that death doesn't harm animals nearly as much as humans.²⁷ This, he thinks, is because animals lack a comparable level of psychological connectedness over time. He suggests that the badness of death for an individual is relative to the psychological connectedness between the individual at the time of their death, and the individual at the future time when they would satisfy the desires that death deprives them of.²⁸ McMahan calls this the *Time Relative Interests Account* (TRIA) and argues that as well as explaining why animals are harmed less by death than humans, this can explain our intuition that death is worse for an adult than it is for an infant. I'll show that even if one accepts TRIA there is not sufficient evidence to show that animals are substantially more weakly psychologically connected than humans. Thus death is not significantly worse for humans according to TRIA. Further, I'll show that one needn't accept TRIA in order to explain why death is worse for adults than infants, undermining the intuitive force of TRIA. Moreover I'll argue that accepting it means endorsing some rather counter-intuitive implications. As such, we should reject the TRIA and the claim that psychological connectedness over time makes a difference to the harm inflicted by death.

²⁶ Luper-Foy, "Mortal Harm," 240.

²⁷ McMahan, "The Comparative Badness for Animals of Suffering and Death," 70.

²⁸ *Ibid.*

McMahan argues that if all of one's future goods are relevant to the badness of one's death, then the younger one is, the more one has to lose through death (in normal circumstances). Thus it should be that the younger one is, the more harm death causes one. This conclusion is implausible McMahan suggests. If this were the case then death would be worse for a newborn than it would for an adult, yet this is the opposite of what we usually think. It seems that an adult is harmed much more by death than an infant, other things being equal. Thus calculating the harm of death is not as simple as weighing up the amount of future goods individuals are deprived of.

McMahan suggests that we need to take into account one's psychological connectedness over time. Psychological connectedness is the relation between an individual at one point in time, and an individual at another point in time. Such relations could include (for example) an experience and the memory of the experience, the formation and the satisfaction (or frustration) of a desire, maintaining the same beliefs, values, intentions or habits.²⁹ Thus some individual A, is psychologically connected to A from three months ago, if A now continues to possess some mental content (a belief for instance) that A from three months ago possessed. Psychological connectedness also comes in degrees, so I in the present, am psychologically connected to my five year old self to the extent that I remember my experiences of myself at five years old, and share the same memories, desires, beliefs, concerns, thoughts, interests, etc. The fewer relations that hold, the less psychological connectedness we share and vice versa. Psychological connectedness is important because of the role it plays in personal identity debates. Having continuously overlapping periods of strong psychological connectedness is called *psychological continuity* and it is psychological continuity many suggest is a necessary condition for personal identity³⁰ or at least it is 'what really matters' when we are considering questions of personal identity.³¹

What is important then is not just what I am deprived of, but whether the individual who is deprived of desire satisfaction is psychologically connected with the person I am, at the time of my death. If for instance, one dies as a newborn, the majority of the desire satisfaction of which one is deprived would be possessed by an individual who would possess little psychological connectedness to the individual that died. After all, the person who would have got married, had a fulfilling career, fell in love, made strong friendships, etc. would be

²⁹ Jeff McMahan, *The Ethics of Killing: Problems at the Margins of Life* (Oxford: Oxford University Press, 2003), 39.

³⁰ *Ibid.*, 39–40.

³¹ Derek Parfit, "Personal Identity," *The Philosophical Review* 80, no. 1 (1971): 12.

psychologically very different to a newborn baby. The individual who would be deprived of these desires is only remotely psychologically connected with the individual that died. As such, McMahan thinks, this seems good reason to think that the newborn is harmed relatively little by the deprivation of this desire satisfaction. The only desire satisfaction that the newborn will be deprived of that will really harm her, will be the desire satisfaction that belongs to her very-near-future self to whom she is closely psychologically connected. For instance, the desire satisfaction she would have experienced in the following week to month. Unlike newborns, adults can be harmed by the deprivation of desire satisfaction in their remote futures because they have, since their own infancy, become more psychologically connected with their future selves.

The more psychologically connected one is with one's future self who would possess the goods of which death deprives one, the more one is harmed by death. Thus, as children mature, and the amount of good life ahead of them lessens, death in fact becomes a greater harm not a lesser harm.³² More importantly for my argument here, McMahan argues that whilst most animals are sufficiently psychologically connected to be harmed by deprivation of near-future goods (desire satisfaction in the immediate future for instance), they are not sufficiently psychologically connected to be harmed by being deprived of more distant goods (desire satisfaction over a year or more away).³³ If McMahan's claims are correct, animals can be harmed by death, but death is significantly less harmful for them than it is for normal adult humans.

I am not convinced that McMahan's argument adequately shows that animals are harmed less by death than humans. There are four arguments I will make in support of my claim. Firstly, I will suggest that there is insufficient evidence to show that animals *are* weakly psychologically connected. Thus if TRIA is true, we cannot assume that this demonstrates that animals are harmed less by death than humans. Second, I will argue that TRIA does not appear to be true since there doesn't seem to be good grounds to believe that weaker psychological connectedness entails that one is harmed less by future events (so long as one maintains personal identity).³⁴ Thirdly, I will attempt to explain how an adult is harmed by

³² McMahan, "The Comparative Badness for Animals of Suffering and Death," 71.

³³ *Ibid.*, 77, 79.

³⁴ Importantly, one many take psychological connectedness to be a necessary (and sufficient) condition for personal identity. Indeed I assume that this is the case below. Assuming this is correct, my claim should be understood to mean: weaker psychological connectedness does not entail that one is harmed less by future events so long as there is *some* psychological connectedness between you now and you in the future that is the subject of the harm.

death more than an infant without appealing to psychological connectedness, thus providing a competing explanation to the intuition that motivates TRIA and undermining the motivation to adopt TRIA. Finally, I will give some counterintuitive implications of TRIA that seem deeply problematic.

Firstly do animals really have weak psychological connectedness? This question is difficult to answer but most agential animals certainly have more psychological connectedness than McMahan describes a human newborn possessing: an individual with no future-directed desires or intentions, nor capable of having any memory of the present in the future.³⁵ The ability to learn and retain information is a good test of psychological connectedness and a great many animals are able to do this. Cognitive ethnological evidence as well as personal experience can tell us that many animals have memories that last over several years. Getting the dog's shampoo out of the cupboard where it has sat for the last year will still make the dog bark and jump around. This is plausibly because she remembers what happened last time the shampoo came out of the cupboard; it was immediately followed by an experience she didn't enjoy. Lions, after years of in the wild, can remember specific human companions they were once friendly with, and treat them with affection whilst still treating other humans with caution and suspicion.³⁶ Squirrels and birds will find nuts and other food stuffs they had hidden months before. Salmon can remember and recognise the scent of an area of water after several years, using these scents to navigate back to breeding grounds.³⁷ Many animals remember abilities that they learn at a young age from their parents, how to hunt or stalk for example.

It is clear that many animals can possess psychological relations over long periods of time (at least over a year) then. However, perhaps animals do not possess a sufficient number or variety of psychological relations for any substantial amount of time. For instance, they may possess memories but lack long-term retention of beliefs, intentions, desires or other mental characteristics. This seems like a relevant concern but it is not clear to me why one would assume that it is the case. It would be fallacious to assume that a lack of evidence for the retention of various mental states is evidence for a lack of those mental states. There may

³⁵ Ibid., 71.

³⁶ Asad Khan, *"The Lion Man" - Shocking Real Story*, 2010, <https://www.youtube.com/watch?v=rxPaUUaxGIM>.

³⁷ Victoria Braithwaite, *Do Fish Feel Pain?* (Oxford: Oxford University Press, 2010), 14-15; "How Do Spawning Fish Navigate back to the Very Same Stream Where They Were Born?," *Scientific American*, 2008, <https://www.scientificamerican.com/article/how-do-spawning-fish-navigate-back/>.

be all kinds of reasons that can explain why we don't have empirical evidence for long-term beliefs existing in animals. For the most part providing evidence for the presence or absence of the relevant psychological relations in animals is an incredibly difficult process. It would be very difficult to attempt to determine whether a sheep still held the specific beliefs she held four years ago for instance, not least because of the philosophical problems surrounding animals' beliefs discussed in Chapter 4 Sec. I. To draw out just one point, since animal minds are (for the most part) non-linguistic, it may be difficult for us to recognise the relevant processes that are hallmarks of particular psychological relations in humans. This could easily lead to inaccurate data.

Perhaps though, there is more to this argument. One might think that because animals lack language (or some other capacity) they are unable to engage in abstract thought and that this is an upper limit on the kind of mental states available to them. As such, they simply won't be able to possess many of the mental relations that we are able to possess and so they must *at best* be only weakly psychologically connected over time and could never hope to be as strongly psychologically connected as humans. We should be careful with this argument. McMahan claims that one possesses strong psychological connectedness when the number of one's day to day connections is over half the number that hold over each day for a normal person. However it is not clear to me why the number of connections a normal person has is relevant to the strength of one's own psychological connectedness.

TRIA (and other arguments that make use of psychological connectedness) use psychological connectedness as a notion to explain how similar some individual at time $t1$ is to some individual at $t2$. It is an attempt to show, to what degree, two individuals have overlapping psychological architectures. However, the absolute number of psychological connections doesn't appear to be particularly relevant to providing an answer to this question. What seems much more relevant is the percentage of one's connections that persist over time.

Consider for instance two individuals Tina and Tracey. Tina has 100 mental states and Tracey has 10.³⁸ A week later Tina has retained only 60 mental states however Tracey has retained 8. According to McMahan's criteria for strength of psychological connectedness Tina, having with the higher number of connections to her past self, should be taken to have the stronger degree of psychological connectedness. However of the two of them, it is Tracey who

³⁸ I will talk in terms of mental states for ease of comprehension but some things that are relevant connections cannot rightly be called mental states (habits for example).

has retained the largest proportion of her past self's mental states. She has retained 80% whereas Tina has only retained 60%. Claiming that Tina has stronger psychological connectedness merely because she had a larger quantity of mental states to start with seems odd. Since a higher proportion of Tracey's mental states have persisted through time, it seems to me that Tracey has stronger psychological connectedness than Tina.

Of course what might be relevant are the kinds of connections one has over time. One might be considered more or less psychologically connected depending on whether the connections one maintains over time are central to one's identity. Two individuals with the same percentage of connections over time, might be considered to have different levels of connectedness, if one has kept more trivial connections through time, whereas the other has maintained connections more central to one's identity. Nonetheless, the overall amount of mental states one has doesn't appear to be important, so long as one has a sufficient number and they are of a sufficient complexity for the possessor to be considered to have a mind.

This being the case, can we really claim that animals are more weakly psychologically connected than humans? This is a difficult question and a properly informed answer will rely heavily on the relevant empirical research (much of which, for reasons stated above is and will remain, at least for the foreseeable future, unavailable). However, it seems that the wrong answer here is to assume that animals *are* more weakly psychologically connected over time merely because they lack as great a variety of connections as humans. Further, we know that many animals are at least capable of being psychologically connected with regard to some of their capacities, memory for example. Therefore at best I think we should remain agnostic regarding the claim that animals are less psychologically connected than humans at least until we have some evidence the points towards this conclusion.

Secondly, I will consider why one might think that having weak psychological connectedness would mean that one is harmed less by death than one with strong psychological connectedness. It is not clear why McMahan thinks that this is the case (other than that he finds it intuitive) so it is worth considering if there are independent reasons to shore up (or destabilise) this idea.

Assuming that psychological continuity is necessary for personal identity, one might think that an individual that is weakly psychologically connected to its future and past selves, won't be harmed as much by death, because they lack overlapping periods of strong psychological connectedness. Thus they lack psychological continuity with their intermediate

past and future selves, and so lack identity over time. Their lack of identity over substantial periods of time will ensure that 'their' future life will belong to a different psychological individual). As such death won't deprive them of *their* future life, but merely the life of some individual existing in the future (albeit one that is tied to the same organism). So for instance, if a hawk is only strongly psychologically connected to itself two days into the future, the hawk's death *now* would deprive her of the goods she had the opportunity to pursue in the next two days. However death would not deprive her of the lives of the other psychological individuals that would have existed in her body after this time, since these lives are not hers to lose.³⁹

However it doesn't seem that an argument based on the claim that weakly psychologically connected beings lack psychological identity over their lifetime will succeed (as McMahan himself states). As McMahan argues elsewhere, why should we think one needs overlapping periods of *strong* psychological connectedness in order to possess psychological continuity at all? Surely overlapping periods of *any* degree of psychological connectedness matter. It seems arbitrary to maintain that this isn't the case.⁴⁰ Overlapping periods of psychological connectedness (of any strength) McMahan maintains, are sufficient for *Broad Psychological Continuity* and it is this which is a necessary condition for identity over time.⁴¹

If we fail to recognise this claim then we must accept that newborns are not the same individuals as the adults they will become, and those with late-stage Alzheimer's (for example) are not the same persons they once were. These claims seem counterintuitive. Therefore I suggest that we should take only broad psychological continuity to be a necessary condition for identity rather than the stricter for of psychological continuity. Thus so long as one maintains an broad psychological continuity with the individual that is deprived of goods through one's death, one is still harmed by death whether one is strongly psychologically connected to that individual or not.

So long as McMahan accepts that having broad psychological continuity is necessary for having identity over time, it is somewhat mysterious why being more weakly psychologically connected would entail being harmed less by death.⁴² That I in twenty years

³⁹ McMahan, "The Comparative Badness for Animals of Suffering and Death," 75-76.

⁴⁰ McMahan, *The Ethics of Killing*, 50.

⁴¹ McMahan's condition is a little more complicated, specifying that this is only a necessary condition in non-branching cases of personal identity but I leave this concern aside here since it is tangential to the current debate. *Ibid.*

⁴² McMahan, "The Comparative Badness for Animals of Suffering and Death," 73.

will be wholly different save for my continued strong desire in avoiding pain, shouldn't plausibly make this future interest of mine any weaker to me now than if I retained all the rest of my psychological content with it. Why should it matter whether an individual is more or less psychologically connected with one's future self, so long as one is sufficiently psychologically connected so that an identity relation is maintained? If the only reason to accept TRIA is its intuitive appeal then this is problematic. Especially since, as I will now argue, TRIA isn't all that intuitive.

Firstly there are alternative explanations for why we might think the death harms adults to a greater extent than it harms infants. For instance, one might be tempted to think that newborns are so close to lacking any psychological connectedness at all, and therefore close to having no identity over time, that they have less of a claim to identity than more psychologically connected beings.⁴³ This is of course false. Whilst psychological connectedness comes in degrees, identity relations do not. One either does or does not share an identity relation with some future self.

Alternatively (and/or additionally) it may be the case that death *is* just as bad for an infant as an adult. However our intuitions are distorted because of the various side-effects that are common among deaths of adults, the states of affairs in which an adult dies, is usually worse than the state of affairs in which an infant dies. For instance, adults usually have many relationships of a deep level and play some role in society. However infants have no two-way emotional relationships and play no role in society at all. Thus an adult's death will generally result in a worse state-of-affairs than an infant's death. This may skew our intuitions and lead us to feel the intuitive pull of the claim that an adult's death is worse *for them* than an infant's (even though we are consciously aware that their death being a worse state of affairs in itself, does not make it any worse for them specifically).

One might also think resource investment towards satisfying desires is relevant. Where infants won't even possess most of their most important desires for decades, adults already possess them (or are much near possessing them) and have likely invested more resources in satisfying them. As such the death of an adult in most circumstances appears tragic. They have gone so far towards satisfying their desires only to fall at the last hurdle. Compare a sprinter irreparably breaking their legs the night before their Olympic games

⁴³ McMahan states that when we come into existence we are 'only weakly psychologically related to ourselves in the future. We are unaware of having a future in prospect and thus have no future-directed desires or intentions; nor will we later have any memory of our present experience.' *Ibid.*, 71.

debut versus someone who would have otherwise have gone on to become a sprinter, suffering the same injury as a young child. The former case seems all the more tragic because of the athlete's wasted investment in the satisfaction of their desire.⁴⁴

Considering these replies, it seems plausible to me that psychological connectedness does not need to be employed to explain our intuition that the death of an adult is worse than the death of an infant. Most plausibly I think this can be accounted for by the degree of 'wasted investment' that occurs in cases where individuals narrowly miss satisfying their desires. Further, this intuition may in part be accounted for by our mistaken beliefs regarding the variety of negative side effects involved in the adult case, and beliefs about the identity relations of infants over time. This gives us reason to doubt that a strong psychological connectedness between the time that one dies and the time at which one would satisfy one's desires, makes a difference to the degree of harm one suffers from death.

As well as competing explanations for the intuitive appeal of McMahan's infant death case, there are some counter-examples which show that TRIA gives us very unintuitive answers in some other cases. Consider my first counter-example: Constantine had always been a materialistic man with strong opinions. He was a hardened atheist and overly rationalistic to the point of cold-heartedness. One morning however he was visited by a travelling preacher. During this meeting Constantine has a religious experience, immediately renounces his previous beliefs and wholeheartedly embraces a new faith. This meeting so profoundly affects him that he not only changes his beliefs, and the way he sees the world and his past experiences, but it makes him feel different. He instantly feels deep compassion for all other living beings and immediately begins to live life in a radically different way. He throws away all his material goods and becomes a missionary for his new religion.

Now if McMahan is correct, were Constantine to have a fatal heart attack just after his epiphany, Constantine would be harmed more by death than if he suffered a fatal heart attack immediately before his epiphany. This is because if Constantine was to die moments before his epiphany, all of his future possible goods would be attained by a version of himself to which he is only weakly psychologically connected. Whereas if Constantine was to die moments after his epiphany, all his future possible goods would be attained by a version of himself that he is much more strongly psychologically connected to. This I maintain is

⁴⁴ I came to this conclusion independently of Rowlands who makes the same point and, coincidentally, uses a very similar example. Rowlands, *Animals Like Us*, 82-83.

counter-intuitive. It isn't at all clear to me why we should think that this change (radical though it may be) would make his death more or less harmful. The future goods that Constantine misses out on in either case, are the same goods and would still be attained by him (he maintains his identity through the change). The goods in each case seem to be just as much his and so it seems he should be harmed to the same degree by the loss of the opportunities to pursue them in either case.

Consider another counter-example.⁴⁵ Bobby, a newborn baby, has a life-threatening illness. If the illness isn't treated now, Bobby will live a normal life for the next fifteen years then suffer horribly for several months and die. If the illness is treated now, Bobby will undergo surgery (under anaesthetic) tomorrow. Over the next two weeks Bobby will suffer whilst his body heals but not nearly as much as he would, fifteen years from now, if he wasn't treated. He will then be cured and will live a normal healthy life.

Assuming newborns have little psychological connectedness to their future adult selves (as McMahan himself seems to think), it appears that there is no obligation to treat Bobby's illness. In fact, worse than this, it seems that Bobby's interest in not suffering over the next two weeks is going to be significantly stronger than not suffering and dying in fifteen years' time. As such, according to TRIA, one ought not to treat Bobby's illness and let Bobby suffer horribly in fifteen years' time. Though TRIA would tell us that the short-term suffering is justified by the long-term benefit in the case of a psychologically connected adult, or where the suffering is less remote, it cannot accommodate the case of Bobby as described here. This seems deeply counter-intuitive. Further TRIA's suggested course of action is contrary to the widely accepted ethical principle that a small amount of harm now can be justified by significantly greater benefit later.

These examples severely weaken McMahan's claim that TRIA has intuitive appeal. This, combined with the lack of an independent argument to show why less psychological connectedness results in less harm being inflicted by death, gives us good reason to treat TRIA with suspicion. Even assuming that one shares McMahan's intuition that death is worse for an adult than an infant, TRIA has some counter-intuitive implications, so counter-intuitive, I suggest, that we should reject TRIA. These bullets are simply too big to bite.

⁴⁵ This example is based on similar cases of Billy and Tommy that Harman uses: Harman, "The Moral Significance of Animal Pain and Animal Death," 734-35.

Overall then it appears to be the case that we do not have good reason to assume that animals lack strong psychological connectedness (although neither can we assume that they have strong psychological connectedness). Moreover, even if animals do have weaker psychological connectedness than humans, this would not show that they are be harmed less by death than humans, since TRIA fails to establish the connection between one's degree of psychological connectedness and the severity of the harm of death to oneself.

iii. Categorical Desires

Bernard Williams famously argued that death is bad for one only when it frustrates one's *categorical desires*.⁴⁶ Categorical desires, Williams explains, are desires that are not conditional upon one being alive. If for instance I have a final desire to travel to Istanbul, I will desire this unconditionally, that is to say, it is something I desire to remain alive to do. In contrast to such desires are conditional desires; desires I possess only so long as I am alive to satisfy them. For instance, my desire to eat healthy. I desire to eat healthy only because I want to remain alive and avoid health-problems. Were I mere hours from death, I would not feel regret at being unable to satisfy this desire and I would not want to remain alive merely so I could eat healthily.⁴⁷

Williams suspected that categorical desires are unique to adult humans. More recently this position has been defended by Christopher Belshaw who has argued more forcefully that animals are unable to possess categorical desires. Belshaw argues that this is due to the fact that (most) animals lack the ability to think about the future, a necessary condition for possessing these desires.⁴⁸ As such, Belshaw argues, death cannot be bad for animals.⁴⁹ From what I have said already about future-directed desires, it will be clear that I will not accept Belshaw's argument, although I do have reservations about the possibility of animals possessing categorical desires. Much more importantly however, I will argue that categorical desires are not necessary in order to be harmed by death. Thus if animals lack categorical desires, they can still be harmed by death.

⁴⁶ Williams, "The Makropulos Case Reflections on the Tedium of Immortality."

⁴⁷ *Ibid.*, 4.

⁴⁸ Belshaw, "Death, Pain and Animal Life," 37, 39.

⁴⁹ Belshaw does note that some animals may be able to think about the future and thus may be harmed by death, chimpanzees, elephants, some birds and cetaceans. *Ibid.*, 38.

Bradley has recently argued that in fact we should take many animals (roughly all animals that would be considered agents on my account) to be capable of possessing categorical desires. Bradley argues for this conclusion by suggesting different possible interpretations of categorical desires, and showing that of the possible interpretations, the closest to Williams' own formulation of categorical desires, is easily within the capacities of many animals.

Categorical desires are supposed to be desires that provide us with a reason to continue living.⁵⁰ Thus if we die before we satisfy a categorical desire, that desire is frustrated, just as if we had failed to satisfy it were we to continue living. For instance, if I die before I am able to visit Istanbul, this desire is frustrated just as it would be if Istanbul was destroyed by war before I was able to visit. In either case, I am harmed by the frustration of my desire. Non-categorical desires however are supposed to be conditional upon me remaining alive. Thus if I die without satisfying a non-categorical desire, that desire is not frustrated, but merely *cancelled*.⁵¹ Unlike a categorical desire, it is almost as if I had never had this desire in the first place. Since I only desire to eat healthily in order to avoid health problems, once I die, this desire is no longer relevant since the threat of health problems no longer stands. Just like how, for instance, I only desire to wear sunglasses when it is sunny outside. If the sunny weather ends before I find my sunglasses (as it often does during the British summer), my desire to wear them disappears with it. However my desire is merely cancelled rather than frustrated and as such I suffer no negative impact from failing to satisfy this desire.

Since the cancellation of a desire doesn't harm us, conditional desires don't provide us with any reason to continue living. If all our desires are of the conditional kind then death would merely cancel them all and we wouldn't be harmed by death at all. The frustration of desires does harm us however. As such, if we have desires of the categorical kind then we are harmed by death since these desires are frustrated by the ending of our lives.

In we understand categorical desires in this way then it seems that there is little problem attributing categorical desires to many animals. As Bradley argues, even a cow's desire for eating grass for instance seems to be categorical.⁵² The cow doesn't desire to eat grass merely on the condition of being alive. The cow appears to want to eat the grass for no further reason than her wanting to eat the grass. In fact as I argued in Chapter 3 Sec.IV all

⁵⁰ Ruth Cigman, "Death, Misfortune and Species Inequality," *Philosophy & Public Affairs*, 1981, 58.

⁵¹ Bradley and McDaniel, "Death and Desires," 6.

⁵² Bradley, "Is Death Bad for a Cow?," 55.

sentient beings must possess *de re* final desires for sensory experiences. These are desires that one has for sensory experiences for *no further reason*. Thus, it seems that all sentient beings must possess some categorical desires.

Further, as Bradley notes, possessing a desire conditional upon being alive if anything, seems to require *more* psychological sophistication than a categorical one. Those who are sceptical about the mental capacities of animals, he continues, should be more sceptical about them possessing desires conditional upon them being alive than categorical desires.⁵³ Bradley's reason for thinking this is presumably that a desire conditional upon being alive would require one to be aware of several facts including: that one is alive, that one will die and that *x*-ing is a means to continue living and so avoiding death, as well as the complex abstract concepts of life and death, that these claims entail.

Furthermore, understanding animals as acting purely on desires that they possess on the condition that they are alive, seems an excessively teleological understanding of animal action. We would not be happy to accept that all human desires are conditional upon one being alive. After all, consider the explanation we might give in a comparable human case:

I desire to eat pizza tonight. Let us assume that I desire to eat pizza for no further reason other than for the sensations that come along with eating pizza (the usual reason for which I choose to eat pizza). I do not desire to eat pizza merely on the condition that I am hungry and pizza is the means I choose to fulfil this end. If I was to get home this evening to find that there was no pizza available for me to eat, but I could have something else equally nutritious for dinner instead, I would not be satisfied with the situation. This is because I want pizza specifically. As such my desire for pizza (in this instance) must be a categorical desire.⁵⁴ If one accepts that I possess a categorical desire in this case then one should have no problem in accepting that animals have categorical desires.

⁵³ Ibid.; Bradley and McDaniel, "Death and Desires," 18.

⁵⁴ One may point out that the cow's desire to eat grass is present because it is necessary for their continued existence and thus the cow only desires grass on the condition that they are alive. However, in order for a desire to be conditional in the relevant sense here it must be the case that the desire is rationally/motivationally conditional, i.e. not final. The evolutionary claim that cows possess the desire for grass because it is necessary to keep them alive may be true without it being the case that their motivation for this desire is conditional. Since cows plausibly have a motivating reason to eat grass which is formed of a final desire for the sensations that eating grass brings about, their desire is not conditional in the relevant sense here. See Chapter 3 Sec.IV.ii for more discussion of the difference between evolutionary and rational/motivational explanations of reasons.

One may reply to this argument by arguing that there must be more to categorical desires than them not being conditional upon one being alive. Cigman suggests that categorical desires require an understanding of death that mere animals cannot possess for instance.⁵⁵ Others have suggested that in order for a desire to be categorical it must be a final desire for some object that is independently valuable.⁵⁶ Thus the mere final desires for grass and pizza wouldn't count as categorical desire, since eating grass or pizza in themselves aren't intrinsically valuable activities. However a desire to visit Istanbul, as something which might plausibly be considered valuable in itself, could be a categorical desire.

However if either of these interpretations are true then categorical desires cannot plausibly be necessary for death to be harmful. It cannot be that the object of one's desires must be independently valuable for death to harm one, since there are individuals who have idiosyncratic (but final) desires and are plausibly harmed by death.⁵⁷ Further it cannot be that one needs to possess an understanding of death for death to be harm, since this cannot account for the harm of the deaths of infants, the intellectually atypical, etc.⁵⁸

Perhaps, though he doesn't say so, Williams understood categorical desires to be the kind of desires that we focus our lives around; what we might think of as life-goals. If this is correct then whilst categorical desires are final desires, not all final desires are categorical desires. Categorical desires mark out a very limited set of final desires that one has. Categorical desires in this sense are some of the most important desires that we have and are exactly the kind of thing that would motivate one to continue living, and make death bad for one. In fact, these life-centring desires not only make death bad for us but they seem to give our lives a deep purpose. In other words, they don't just make our lives worth living, they give them meaning. I am somewhat persuaded by this interpretation of categorical desires; it doesn't seem that Williams would accept that final desires for pizza or grass are categorical in the sense he had in mind. However, if final desires have to be sufficiently important to us that we centre our lives around them in order for them to count as categorical, it seems that categorical desires just aren't necessary to be harmed by death. One's life does not seemingly need to have meaning in order for it to be a harm to lose it.

⁵⁵ Cigman, "Death, Misfortune and Species Inequality," 57-59.

⁵⁶ Bradley and McDaniel, "Death and Desires," 8; Mikel Burley, "Immortality and Meaning: Reflections on the Makropulos Debate," *Philosophy* 84, no. 4 (2009): 79.

⁵⁷ Bradley, "Is Death Bad for a Cow?," 57.

⁵⁸ There is also the risk of bald anthropocentrism in the present context; how do we determine what is valuable without merely determining what is valuable *for us*?

It is not difficult to find an example to show that this is the case. Consider a human that lacks any ambitions or desires that give their life a purpose. Even if they have merely the dull everyday desires that one might think a cow possesses, they are still intuitively harmed by death. Assuming we understand categorical desires in this deeper way then, though the frustration of these desires will certainly make death a significant harm, the absence of them cannot plausibly make death harmless.

From what has been said it appears that conditional desires are not sufficient to be harmed by death. These desires merely seem to be cancelled by death. However, categorical desires (if they are understood in the deeper sense) are not necessary in order to be harmed by death. It seems that in order to be harmed by death one need only possess final desires. Death does not merely cancel these desires but frustrates them, just as with categorical desires. However as noted at the start of this section on continued existence, the possession of just any final desire will not be sufficient to make death bad for one. Some final desires, can be satisfied by the ending of one's life, the desire the absence of pain for example. As such possessing only this desire would not make death bad for one. Therefore, death is *pro tanto* bad for one only if one has final desires which are cannot be satisfied by death.

All agents, in virtue of being agents possess some final desires and, so long as these desires are not satisfied by death, must be capable to be harmed by death. Further as I argued in Chapter 3, all sentient beings possess some final desires for the experience of some sensations. These are final desires that are not satisfied by death. Therefore all sentient beings can be *pro tanto* harmed, to some extent, by death. Of course this does mean that death is always a harm *overall*. In some cases this *pro tanto* harm may be outweighed by other considerations and death may be, all things considered, good for one.

Through this section I have argued that non-human agents are harmed by death in the same way as typical humans. Further I have presented arguments against the claim that there is a significant difference between the extent to which humans, and the extent to which animals, are harmed by death. I argued firstly that animal agents, just like humans, possess future-directed desires, but also that such desires are not necessary in order to be harmed by death. Secondly, I argued that animals are psychologically-connected over time and thus seemingly possess identities over time. Furthermore, I argued that there are various independent reasons to reject McMahan's Time-Relative-Interests Account and thus one's psychological connectedness does not appear to be relevant to the degree of harm one suffers at some point at time. I next argued that non-human agents are capable of possessing

categorical desires on some understanding of categorical desires. Further, I showed that if one adheres to a deeper understanding of categorical desires, which animals may not possess, then such desires are not plausibly necessary to be harmed by death. In fact I argue that final desires that are not satisfied by death, seem to be all that is required in order to be harmed by death. Further these are possessable by all agents, and possessed by all agents of which we have knowledge (that is all sentient beings).

As such it seems that an individual will be *pro tanto* harmed by death, if death deprives them of the satisfaction of final desires that would not be satisfied by death (even if they do not, at the time of their death, possess the desires which will be frustrated by death). As such since agents necessarily possess final desires, and it is practically always the case that at least some them are not satisfied by death (and with sentient agents necessarily so), merely in virtue of their agency, agents must have an interest in continued existence. Further this must mean that ending the life of an agent is *pro tanto* wrong (though it may of course be all things considered permissible and even good).

II. The Loss of Autonomy Argument

So far I have shown that the possession of agency is sufficient for interests in freedom from pain, continued existence and liberty. I have supported these claims by arguing that agents have an interest in avoiding pain because pain is a result of the frustration of important desires. Agents have an interest in liberty (in the wider sense of the term that it is good that they have liberty but does not necessarily contribute towards their well-being) because it allows them to determine the course of their own lives.⁵⁹ And finally, agents have an interest in continued existence because death deprives them of the satisfaction of certain final desires.

In the argument up to this point then I have shown that agents possess the interests we consider to be essential to the notion of full moral status. However this alone does not establish that agency is a sufficient condition for full moral status. In order to show that this is the case I need to demonstrate that these interests, when possessed by paradigmatic full moral status holders (i.e. autonomous agents), are grounded by the very same capacity. In this section, I will attempt to establish this conclusion, showing that in fact autonomous agents

⁵⁹ I will throughout the rest of this work refer to agents having an interest in liberty. I use this term only for convenience. As mentioned in the previous chapter, more properly I mean that liberty is good for them in a non-agent-affecting way.

have interests in continued existence, liberty and freedom from pain, in virtue of their agency (and not because of any other capacity, such as autonomy). In doing this, I will show that agency is the shared basis for these interests among autonomous and non-autonomous agents and thereby that all agents, autonomous or not, should be taken to possess full moral status. In order to do this I will use of a thought experiment based on an argument by Jeff Sebo.⁶⁰

Consider a regular human autonomous agent: Jane. Jane has full moral status. One night during her sleep Jane is drugged by evil philosophers of mind. The drug permanently strips her of her autonomy, reducing her to a non-autonomous agent. Whilst she can still act in accordance with the principles she previously held as an autonomous agent, she no longer *acts upon* those general rules or principles (since she cannot cognise them) nor does she endorse or repudiate any of her desires. Jane merely acts upon here first-order desires. Other than this, Jane is exactly the same as she was before the drugs were administered. She is still an agent in the sense that her actions are rationalisable in light of her desires and beliefs, and she is certainly still causally responsible for her own actions. This being so, has Jane plausibly lost her full moral status? I suggest not.

It seems that Jane would still be *pro tanto* wronged by having pain inflicted upon her, having her life ended or her liberty restricted. Moreover she would be wronged by such actions *in the same way* and *for the same reason* as she would have been, had she been a victim of these wrongs before she lost her autonomy. She still plausibly has interests of the same kind in continued existence, liberty and freedom from pain, as she had before she was robbed of her capacity for autonomy. She will not seemingly have *the very same interests* however. Since she has lost some of her capacities, at least some of her interests will have changed in strength. Thus the degree of harm she will suffer from having these interests frustrated, and the extent to which it will be wrong to frustrate these interests, will change too.

For example, she will not be wronged to the same extent were she deprived of her liberty now than she would have been, had she been deprived of her liberty before she lost her autonomy. Since there are a large amount of choices she can no longer make, specifically higher-order choices. She will no longer be wronged by being obstructed from acting

⁶⁰ Sebo suggests that if an autonomous agent was drugged and reduced to a non-autonomous agent they would not lose their interests in property, liberty or continued existence. My argument is slightly different in that I discuss the interest in the absence of pain rather than the interest in property and because I suggest that one would maintain the same *kind* of interests, through not necessarily interests of the same strength. Jeff Sebo, "Agency and Moral Status," *Journal of Moral Philosophy* 14, no. 1 (2015): 18.

authentically in line with what she truly desires. She will no longer be wronged by being put into an autonomy-undermining state. Since she has lost her capacity for autonomy, she also won't be able to give informed consent.⁶¹ So she couldn't be wronged by not being given the choice to opt out of a medical procedure. Neither will she be morally responsible for her actions any longer since she has ceased to be a moral agent, and will therefore not be due punishment in the case of apparent wrongdoing.⁶²

Jane will certainly be entitled to different treatment then. Nonetheless, she will still maintain interests in the three central goods that characterise full moral status. Further, she will maintain the same *kind* of interests in these goods. Whilst her interests have changed (where they have changed) with respect to these goods, they have merely weakened in strength, not changed in their format. It is still impermissible to kill her or confine her or inflict pain upon her, for the same reason that it was impermissible to do these things when she was autonomous. Furthermore, the interests she retains are due equal consideration with like interests, just as they were before she lost her autonomy. This shows that autonomy is not necessary for the possession of these interests in freedom from pain, continued existence and liberty, and that these interests are grounded in the possession of mere agency. Further, one cannot object that we only believe Jane to continue to possess these interests after the loss of her capacity for autonomy merely because she used to be an autonomous agent. The sections on freedom from pain, liberty and continued existence over this chapter and the last, have established on independent grounds that ordinary non-autonomous agents possess these interests, even though they have never, and will never, possess autonomy.

Now my argument for agency as a condition sufficient for (full) moral status is complete. The kind of interests we take paradigmatic full moral status holders to have in liberty, continued existence and freedom from pain can be shown to be grounded in the mere capacity for agency. So even in autonomous agents, it is agency and not autonomy, which bequeaths them these interests and thus their moral status. Hence, where individuals lack autonomy, and are merely non-autonomous agents, they too must possess these same interests. Therefore the capacity for agency is sufficient for the possession of moral status and as such all agents, autonomous or not, must possess moral status.

⁶¹ Ibid.

⁶² Ibid., 19.

III. Objections to the Agency Account of Moral Status

One might think that, even accepting all the previous arguments, there is motivation to view autonomous agents as worthy of a higher moral status than non-autonomous agents in light of the fact that they have a more sophisticated kind of agency. If agency is morally significant, the argument goes, then autonomous agency must be *more* morally significant. Is there anything that can be said to resist such an argument? Firstly, as I have already argued in Chapter 1 Sec. III moral hierarchies are problematic. As I argue there, it is difficult to see how one could justify granting extra moral weight to any individual's interests. More importantly, one should note that even if a coherent justifiable account of moral status could be defended, one has misunderstood the argument contained in the this chapter, and the previous chapter, as well as the notion of agency in play, if one thinks that having autonomous agency would grant one a higher level of moral status than mere non-autonomous agency.

The relevant notion of agency for my purposes is the *capacity* to act for reasons. It is the possession of this capacity that grants one moral status. Agency as I have argued for it here then, is a binary concept that one either possesses or lacks. As such, one cannot be more or less of an agent than any other agent. One can only be an agent or a non-agent. Further autonomous and non-autonomous agency are not different 'kinds' of intentional agency. If one can intentionally act, then one is an agent in the sense I have defended in this thesis, whether one possesses autonomy or not. Possessing the capacity for autonomy merely means that one is able to apply one's agency to different domains, to consider different and often additional reasons, due to one's self-awareness and ability to reflect upon and scrutinise one's reasons. Autonomous agents may have a greater variety of opportunities to act, reasons to act on and awareness of their capacity for agency, but they don't have any more or less agency than any other agent. One is still an agent in the sense that one is able to act on motivating reasons and one's actions can be rationalised in light of one's desires and beliefs.

Thus whilst autonomy confers further interests, duties and capabilities upon agents, and these will in turn mean autonomous agents are due some different treatment, this doesn't make a difference to one's moral status. Consider a point I made back in the Introduction: women may have an interest in being free to have an abortion, however since men are unable to have abortions, they plausibly don't have such an interest. This doesn't mean that men and women must have different status though. Though men and women have different interests in some respects, and are owed different treatment in line with these interests, they have the

same *central* interests and thus equal moral status. This is similar to how we should view the difference between autonomous and non-autonomous agents. Though they have differences, and can be harmed and wronged in unique ways, they are all agents that can be harmed and wronged in the same central ways. Thus they share equal moral status.

A relevant difference between the cases of autonomous and non-autonomous agents, and men and women however, is that men and women have interests of equal strength regarding continued existence and liberty for example. This is not the case with autonomous and non-autonomous agents. As such autonomy does make a relevant difference; it strengthens one's interests in certain goods and deepens the ways in which one can be harmed and wronged in some respects. However this is still insufficient to establish autonomous agents as possessing a different degree of moral status to non-autonomous agents. Two individuals with vastly different levels of intelligence will plausibly have interests of different strength in being able to attend to university. Whilst someone with a genius level intellect, and an ambition to revolutionise artificial intelligence, will have a significant interest in being able to attend a university, someone with below average intelligence, and no intention of studying, will have a very weak interest in being able to attend university. However whilst we take this difference in strength of interest to justify different treatment, we do not take it to be an indicator, or a justification for a difference in status.

What is important is not whether agents with autonomy are different from non-autonomous agents (they clearly are in some respects) but whether these differences can justify a difference in status. Since the central ways in which they are harmed and wronged are the same, I suggest that the differences between them cannot justify a difference in status. Non-autonomous agents possess the ability to act on motivating reasons just as autonomous agents do, and it is this capacity which grounds their status. All agents, regardless of the reasons on which they can act, or the domains over which they can exercise their agency, are agents in the very same sense and thus possess equal moral status.

Conclusion

In this chapter I completed my argument for the claim that agency is a sufficient condition for moral status. I argued that future-directed desires, strong psychological connectedness and categorical desires are not necessary in order to be harmed by death. Death is a harm I argued where it deprives one of the satisfaction of certain final desires (specifically

final desires that are not satisfied by death). Therefore mere agents, as beings capable of possessing these desires (and all sentient beings as beings that necessarily possess some such desires) can be *pro tanto* harmed and therefore *pro tanto* wronged by having their lives ended. Thus mere agents possess the three central interests we consider to belong to full moral status holders; interests in liberty, continued existence and being free from pain. Further, it is their capacity for agency that grants them these interests.

I then argued that an autonomous agent could still be wronged in all three central ways were they stripped of autonomy. Thus it is plausibly this same capacity for agency in autonomous agents (paradigmatic full moral status holders) that grants them these same interests and moral status. Therefore, through this final chapter I have shown that the possession of agency underpins moral status. Further I demonstrated the virtues of the Agency Account. I have shown that the Agency Account can plausibly accommodate the moral status of humans and animals on the same account, for the same reason, and that it gives us an explanation of the *pro tanto* wrongness of central ways in which moral status holders can be wronged. Thus I have shown that the Agency Account comes closer to the ideal account of moral status than rival accounts.

Concluding Remarks



Through this thesis I have argued in defence of three main claims: that sentient beings are agents, that agency is the necessary and sufficient condition for moral status and that moral status holders are owed strong *pro tanto* duties to be free to make their own choices, to not have pain inflicted upon them and to not have their lives ended. I have attempted to provide a philosophically robust account of moral status that shows that humans and sentient non-human animals possess equal moral status, where this is understood as them being entitled to have similar interests equally considered. I have further argued that humans and animals possess moral status for the same reason, because of their shared capacity for agency understood as the capacity for intentional action, and that this capacity for agency entitles them to the same central treatment.

I have argued that the Agency Account of Moral Status overcomes the problems of the next best competitors (the Sentience Account, the Subject-of-a-Life Account and the Kantian Account). Unlike these rival accounts, it can give a plausible and defensible condition for moral status, and can explain why those that lack this condition, lack moral status. Since individuals that lack agency must lack interests, there is nothing to consider in the case of non-agential entities. Thus they cannot be owed any duties (though we may be obligated to fulfil duties regarding them).

Furthermore, the Agency Account can plausibly give an account of the fundamental duties owed to moral status holders and justify them with reference to that property which grants them moral status: agency. In light of their capacity for agency inflicting pain on them wrongs them because it frustrates certain final desires. Further ending a moral status holder's life wrongs them because it deprives them the satisfaction of certain final desires. And finally depriving them of liberty wrongs them because it undermines the authority they have to determine the course of their own lives. Broadly then, it is wrong to frustrate the interests of moral status holders because doing so undermines their agency. Thus the agency account not only provides a plausible and robust argument for why we owe this treatment to agents, but also does so by appealing to the property which grounds moral status on the Agency Account. As such agency explains not just why individuals matter, but in which ways being an agent

makes one matter. By doing this the Agency Account gives us a cohesive account of the *pro tanto* central treatment that we owe to moral status holders which can feed into a normative ethical theory.

I have shown that sentient animals meet the conditions for moral status on the Agency Account since sentience entails agency. Thus the Agency Account unites humans and animals under a single account of moral status, recognising their similarities instead of emphasising their differences. It recognises all beings that matter are, at bottom, the same kind of entities: agents. Thus they are entitled to the same basic treatment and ought to have similar interests, similarly considered. This helps to dissolve the us/them divide that has been constructed between humans and animals, and instead proposes that we view the differences between humans and animals are merely variations among fundamentally similar beings. Thus, though we recognise that all humans are different, there is no assumed chasm of moral value between a severely disabled intellectually human and myself, the same is true of humans and animals.

Despite this however, the account can still accommodate plausible differences in treatment of humans and animals. Humans and animals may be entitled to certain treatment in virtue of the various relations they possess and the positions they hold in cultures, societies, etc. as well as their differing mental and physical capacities. Many of these attributes may well have significant moral importance and establish duties. The Agency Account merely claims that these differences are not sufficient to establish different kinds of moral status or radically different treatment.

Related, the Agency Account also provides a plausible explanation for the equal moral status of humans, despite the vast range of cognitive capacities present among us. Moreover the Agency Account can do this without relying on discriminatory or unjustifiable claims about the nature of our species. Further, it does not claim that beings with similar mental capacities should be treated radically different because of their membership to specific biological or social groups.

The Agency Account will seemingly have little impact upon our treatment of fellow humans (seeing as we already recognise the agency of humans) however this will have a considerable impact upon the way we see and interact with animals. Of course a moral status account alone cannot lay out a whole moral theory (nor should it). Nonetheless this account informs us that through restricting their liberty, ending their lives or inflicting pain upon

them, we *pro tanto* wrong animals. Therefore, any action which has such effects requires justification. Further, since we must equally consider similar interests of all moral status holders, it seems that many of our standard practices involving animals are going to be called into question. Animal farming, animal experimentation and other industries involve actions that will be impermissible. Further zoos and the keeping of wild animals in captivity will (depending upon the animals which are kept) be impermissible or practically and/or economically untenable.

My thesis opens several doors for further research. Particularly research regarding the agency of animals. If animals are agents in the same way as humans, then this will require us to re-evaluate many of our philosophical views regarding the mind. For instance, how should we understand the role of the will within the mind, and specifically strength and weakness of will cases, in light of the claim that animals are agents? We might consider only autonomous agents to possess a will, since only they can form higher-order desires, yet there seems to be something odd about saying that one is an agent yet lacks a will. Further, since non-autonomous agents lack autonomy can we justifiably genetically engineer them to possess certain desires even if these would be detrimental to their well-being?

There are also questions on the more empirical side of things. It seems that non-autonomous agents that belong to a species of which typical members are autonomous, may act in different ways to non-autonomous agents that belong to a species of which the typical members are non-autonomous. They may possess different kinds of mental states and different psychological systems for action. For instance, non-autonomous humans may struggle to use cognitive maps, but some birds may use cognitive maps to great effect to satisfy their desires. The different systems that non-autonomous agents use to exercise their agency will likely have implications for the kind of duties we owe to different agents.

There is also considerable work to do putting the Agency Account of Moral Status into practice. In particular the Agency Account may give us interesting answers regarding how we ought to treat wild animals. Since they are agents, it may be better to allow them to struggle than to provide them with food or shelter, however since we tend to think that this isn't the case with non-autonomous human agents, perhaps we have obligations to ensure that they have a minimum level of well-being throughout their lives. Similarly the Agency Account implies that we should either reconsider our views on euthanising animals or euthanising non-autonomous humans. If animals are agents then it seems it might be wrong to end their lives even if they have little to lose in death.

Bibliography

- Anderson, Elizabeth S. "Animal Rights and the Value of Non-Human Life." In *Animal Rights: Current Debates and New Directions*, by Cass R Sunstein and Martha C. Nussbaum. Oxford: Oxford University Press, 2004.
- Andrews, Kristin. *The Animal Mind: An Introduction to the Philosophy of Animal Cognition*. New York: Routledge, 2014.
- Anscombe, G. E. M. *Intention*. Basil Blackwell, 1966.
- Arpaly, Nomy. *Unprincipled Virtue : An Inquiry Into Moral Agency*. Oxford: Oxford University Press, 2004.
- Asad Khan. "The Lion Man" - *Shocking Real Story*, 2010.
<https://www.youtube.com/watch?v=rxPaUUaxGIM>.
- Audi, Robert. "Acting for Reasons." *The Philosophical Review* 95, no. 4 (1986): 511–46.
- Aydede, Murat. "A Contemporary Account of Sensory Pleasure." In *Pleasure: A History*, edited by Lisa Shapiro, Vol. 1. Oxford: Oxford University Press, 2014.
- — —. "How to Unify Theories of Sensory Pleasure: An Adverbialist Proposal." *Review of Philosophy and Psychology* 5, no. 1 (2014): 119–133.
- Balcombe, Jonathan. *Pleasurable Kingdom: Animals and the Nature of Feeling Good*. London: Palgrave Macmillan, 2006.
- Beck, J. "The Generality Constraint and the Structure of Thought." *Mind* 121, no. 483 (2012): 563–600.
- Beck, Jacob. "Why We Can't Say What Animals Think." *Philosophical Psychology* 26, no. 4 (2013): 520–46.
- Belshaw, Christopher. "Death, Pain and Animal Life." In *The Ethics of Killing Animals*. Oxford: Oxford University Press, 2016.
- Bentham, Jeremy. *An Introduction to the Principles of Morals and Legislation*. White Dog Publishing, 2010.
- Berlin, Isaiah. *Four Essays on Liberty*. London: Oxford Paperbacks, 1969.
- — —. "Two Concepts of Liberty." In *Four Essays on Liberty*. London: Oxford Paperbacks, 1969.
- Bermudez, Jose Luis. "The Distinction between Conceptual and Nonconceptual Content." In *Oxford Companion to the Mind*. Oxford: Oxford University Press, 1998.
- Blatti, Stephan. "Death's Distinctive Harm." *American Philosophical Quarterly* 49, no. 4 (2012): 317–330.
- Bradley, Ben. "Is Death Bad for a Cow?" In *The Ethics of Killing Animals*. Oxford: Oxford University Press, 2016.
- Bradley, Ben, and Kris McDaniel. "Death and Desires." In *The Metaphysics and Ethics of Death*, edited by James Stacey Taylor. Oxford: Oxford University Press, 2013.
- Braithwaite, Victoria. *Do Fish Feel Pain?* Oxford: Oxford University Press, 2010.
- Bramble, Ben. "The Distinctive Feeling Theory of Pleasure." *Philosophical Studies* 162, no. 2 (2013): 201–17.
- Brandt, Richard B. *A Theory of the Good and the Right*. Oxford : New York: Oxford University Press, 1979.
- Broglio, C., I. Martín-Monzón, F. M. Ocaña, A. Gómez, E. Durán, C. Salas, and F. Rodríguez. "Hippocampal Pallium and Map-Like Memories through Vertebrate Evolution." *Journal of Behavioral and Brain Science* 05, no. 03 (2015): 109.
- Broome, John. "The Badness of Death and the Goodness of Life." In *The Oxford Handbook of the Philosophy of Death*, edited by Ben Bradley, Fred Feldman, and Jens Johansson. Oxford: Oxford University Press, 2013.

- Brudney, Daniel, and John Lantos. "Agency and Authenticity: Which Value Grounds Patient Choice?" *Theoretical Medicine and Bioethics* 32, no. 4 (2011): 217–27.
- Burley, Mikel. "Immortality and Meaning: Reflections on the Makropulos Debate." *Philosophy* 84, no. 4 (2009): 529–47.
- Cabanac, Michel. "Physiological Role of Pleasure." *Science* 173 (1971).
- Camp, Elisabeth. "Thinking with Maps." *Philosophical Perspectives* 21, no. 1 (2007): 145–182.
- Chater, Nick, and Cecilia Heyes. "Animal Concepts: Content and Discontent." *Mind & Language* 9, no. 3 (1994): 209–46.
- Cigman, Ruth. "Death, Misfortune and Species Inequality." *Philosophy & Public Affairs*, 1981, 47–64.
- Cochrane, Alasdair. *Animal Rights without Liberation: Applied Ethics and Human Obligations*. Columbia University Press, 2012.
- — —. "Do Animals Have an Interest in Liberty?" *Political Studies* 57, no. 3 (2009): 660–79.
- Dally, Joanna M, Nathan J Emery, and Nicola S Clayton. "Cache Protection Strategies by Western Scrub-Jays (*Aphelocoma Californica*): Hiding Food in the Shade." *Proceedings of the Royal Society B: Biological Sciences* 271, no. Suppl 6 (2004): 387–90.
- Dancy, Jonathan. *Practical Reality*. Oxford: Oxford University Press, 2002.
- Darwall, Stephen. "The Value of Autonomy and Autonomy of the Will*." *Ethics* 116, no. 2 (2006): 263–284.
- Darwin, Charles. *The Origin of Species and the Voyage of the Beagle*. Edited by Ruth Padel. London: Vintage Classics, 2009.
- Davidson, Donald. "Rational Animals." *Dialectica* 36, no. 4 (1982): 317–27.
- Dawkins, Richard. *The Ancestor's Tale*. London: Orion Publishing Group, Limited, 2005.
- DeGrazia, David. "Moral Status as a Matter of Degree?" *The Southern Journal of Philosophy* 46, no. 2 (2008): 181–198.
- — —. *Taking Animals Seriously: Mental Life and Moral Status*. Cambridge; New York: Cambridge University Press, 1996.
- Diamond, Cora. "Eating Meat and Eating People." *Philosophy* 53, no. 206 (1978): 465–479.
- Donaldson, Sue, and Will Kymlicka. *Zoopolis: A Political Theory of Animal Rights*. Oxford: Oxford University Press, 2013.
- Duus-Otterström, Göran. "Freedom of Will and the Value of Choice." *Social Theory and Practice* 37, no. 2 (2011): 256–284.
- Dworkin, Gerald. *The Theory and Practice of Autonomy*. Cambridge: Cambridge University Press, 1988.
- Evans, Gareth. *The Varieties of Reference*. Oxford: Oxford University Press, 1989.
- Feinberg, Joel. "The Rights of Animals and Unborn Generations." In *Ethical Theory an Anthology*, edited by Russ Shafer-Landau, 372–80. Oxford: Wiley-Blackwell, 1974.
- Feldman, Fred. *Confrontations with the Reaper: A Philosophical Study of the Nature and Value of Death*. Oxford University Press, 1992.
- Fletcher, Mark. "The Mating Game." BBC, 2013.
- Fodor, Jerry. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. London: MIT Press, 1987.
- Fodor, Jerry, and Zenon Pylyshyn. "Connectionism and Cognitive Architecture." *Cognition*, 1988.
- Francione, Gary L. "Animals - Property of Persons?" In *Animals as Persons: Essays on the Abolition of Animal Exploitation*. New York; Chichester: Columbia University Press, 2009.
- Francione, Gary L. *Animals as Persons: Essays on the Abolition of Animal Exploitation*. New York: Columbia University Press, 2008.
- Frey, R. G. "Autonomy and the Value of Animal Life." *The Monist* 70, no. 1 (1987): 50–63.
- Gallup, Gordon G. "Chimpanzees: Self-Recognition." *Science* 167, no. 3914 (1970): 86–87.

- Giroux, Valéry. "Animals Do Have an Interest in Liberty." *Journal of Animal Ethics* 6, no. 1 (2016): 20–43.
- Glock, Hans-Johann. "Animals, Thoughts and Concepts." *Synthese* 123, no. 1 (2000): 35–64.
 — — —. "Can Animals Judge?" *Dialectica* 64, no. 1 (2010): 11–33.
- Goodpaster, Kenneth E. "On Being Morally Considerable." *The Journal of Philosophy* 75, no. 6 (1978): 308.
- Hadley, John. "Liberty and Valuing Sentient Life." *Ethics & the Environment* 18, no. 1 (2013): 87–103.
- Hare, Brian, Josep Call, and Michael Tomasello. "Chimpanzees Deceive a Human Competitor by Hiding." *Cognition* 101 3 (2006): 495–514.
- Harman, Elizabeth. "The Moral Significance of Animal Pain and Animal Death." In *The Oxford Handbook of Animal Ethics*, edited by Tom L. Beauchamp and R. G. Frey. Oxford: Oxford University Press, 2014.
 — — —. "The Potentiality Problem." *Philosophical Studies* 114, no. 1 (2003): 173–198.
- Harris, John. "Consent and End of Life Decisions." *Journal of Medical Ethics* 29, no. 1 (2003): 10–15.
- Heathwood, Chris. "The Reduction of Sensory Pleasure to Desire." *Philosophical Studies* 133, no. 1 (2007): 23–44.
- Heck, Richard. "Are There Different Kinds of Content?" In *Contemporary Debates in Philosophy of Mind*, edited by Brian P. McLaughlin and Jonathon D. Cohen. Oxford: Blackwell, 2007.
- Horta, Oscar. "The Scope of the Argument from Species Overlap." *Journal of Applied Philosophy* 31, no. 2 (2014): 142–54.
 "How Do Spawning Fish Navigate back to the Very Same Stream Where They Were Born?" *Scientific American*, 2008. <https://www.scientificamerican.com/article/how-do-spawning-fish-navigate-back/>.
- Kagan, Shelly. "The Limits of Well-Being." *Social Philosophy and Policy* 9 (1992): 169–89.
- Kamm, F. M. *Intricate Ethics: Rights, Responsibilities, and Permissible Harm*. Oxford: Oxford University Press, 2006.
- Kant, Immanuel. *The Metaphysics of Morals*. Translated by Mary J. Gregor. Cambridge: Cambridge University Press, 1991.
 — — —. *The Moral Law: Groundwork of the Metaphysics of Morals*. Translated by Paton, H. J. London: Routledge, 1989.
- Keijzer, Fred. "The Sphex Story: How the Cognitive Sciences Kept Repeating an Old and Questionable Anecdote." *Philosophical Psychology* 26, no. 4 (2013): 502–19.
- Klein, Colin. "What Pain Asymbolia Really Shows." *Mind* 124 (2015).
- Korsgaard, Christine M. *Christine Korsgaard on the Status of Animals*. Philosophy Bites, n.d. <http://philosophybites.com/2015/02/christine-korsgaard-on-the-status-of-animals.html>.
 — — —. "Fellow Creatures: Kantian Ethics and Our Duties to Animals." *The Tanner Lectures on Human Values* 25 (2004): 26.
 — — —. "Kantian Ethics, Animals, and the Law." *Oxford Journal of Legal Studies* 33, no. 4 (2013): 629–48.
- Kuhse, Helga. "Interests." *Journal of Medical Ethics* 11, no. 3 (1985): 146.
- Kundey, Shannon, Andres De Los Reyes, Chelsea Taglang, Rebecca Allen, Sabrina Molina, Erica Royer, and Rebecca German. "Domesticated Dogs (*Canis Familiaris*) React to What Others Can and Cannot Hear." *Applied Animal Behaviour Science* 126, no. 1 (2010): 45–50.
- Lawlor, Rob. "The Ethical Treatment of Animals: An Example of the Moral Significance of Darwin's Theory." In *Evolution 2.0: Implications of Darwinism in Philosophy and the*

- Social and Natural Sciences*, edited by Martin Brinkworth and Friedel Weinert. New York: Springer, 2011.
- Leiblum, Sandra, Candace Brown, Jim Wan, and Leslie Rawlinson. "Persistent Sexual Arousal Syndrome: A Descriptive Study." *The Journal of Sexual Medicine* 2, no. 3 (2005): 331–37.
- Liao, S. Matthew. "The Basis of Human Moral Status." *Journal of Moral Philosophy* 7, no. 2 (2010): 159–179.
- Luper-Foy, Susan. "Annihilation." *The Philosophical Quarterly* 37, no. 183 (1987): 233–52.
 — — —. "Mortal Harm." *Philosophical Quarterly* 57 (2007): 239–51.
- MacCallum, Gerald C. "Negative and Positive Freedom." *The Philosophical Review* 76, no. 3 (1967): 312.
- McMahan, Jeff. "Death and the Value of Life." *Ethics* 9, no. 1 (1988): 32–61.
 — — —. "'Our Fellow Creatures.'" *The Journal of Ethics* 9, no. 3–4 (2005): 353–80.
 — — —. "The Comparative Badness for Animals of Suffering and Death." In *The Ethics of Killing Animals*. Oxford: Oxford University Press, 2016.
 — — —. *The Ethics of Killing: Problems at the Margins of Life*. Oxford: Oxford University Press, 2003.
- Melis, Alicia P., Josep Call, and Michael Tomasello. "Chimpanzees (Pan Troglodytes) Conceal Visual and Auditory Information From Others." *Journal of Comparative Psychology* 120 (2006): 154–62.
- Menzel, E. W. "A Group of Young Chimpanzees in a One-Acre Field." *Behavior of Nonhuman Primates* 5 (1974): 83–153.
- Mill, John Stuart. "On Liberty." In *On Liberty, Representative Government, the Subjection of Women*. Oxford: Oxford University Press, 1918.
- Morris, R. "Developments of a Water-Maze Procedure for Studying Spatial Learning in the Rat." *Journal of Neuroscience Methods* 11, no. 1 (1984): 47–60.
- Nagel, Thomas. "Death." *Nous* 4, no. 1 (1970): 73–80.
- Norcross, Alastair. "The Significance of Death for Animals." In *The Oxford Handbook of Philosophy of Death*, edited by Ben Bradley, Fred Feldman, and Jens Johansson. Oxford: Oxford University Press, 2013.
- Nussbaum, Martha C. *Frontiers of Justice Disability, Nationality, Species Membership*. Cambridge: Harvard University Press, 2007.
- Oderberg, David S. *Applied Ethics: A Non-Consequentialist Approach*. Oxford: Wiley-Blackwell, 2000.
- Oshana, Marina. "How Much Should We Value Autonomy." *Social Philosophy and Policy* 20, no. 2 (2003): 99–126.
- Palmer, Clare, and Peter Sandøe. "For Their Own Good Captive Cats and Routine Confinement." In *The Ethics of Captivity*, edited by Lori Gruen. Oxford: Oxford University Press, 2014.
- Parfit, Derek. "Personal Identity." *The Philosophical Review* 80, no. 1 (1971): 3–27.
- Pettit, Philip. "The Republican Ideal of Freedom." In *The Liberty Reader*. London: paradigm publishers, 2006.
- Pluhar, Evelyn B. *Beyond Prejudice: The Moral Significance of Human and Nonhuman Animals*. Duke: Duke University Press, 2012.
- Poucet, Bruno. "Animal Navigation and Cognitive Maps," 2006.
 — — —. "Spatial Behaviour of Cats in Cue-Controlled Environments." *The Quarterly Journal of Experimental Psychology*, 1985.
- Proust, Joëlle. "A Plea for Mental Acts." *Synthese* 129, no. 1 (2001): 105–128.
 — — —. "Mental Acts." In *A Companion to the Philosophy of Action*, edited by Timothy O'Connor and Constantine Sandis, 209–17. Wiley-Blackwell, 2010.

- — —. "Rationality and Metacognition in Non-Human Animals." *Rational Animals?*, 2006, 247–274.
- Rachels, James. *Created from Animals: Moral Implications of Darwinism*. Oxford: Oxford University Press, 1999.
- — —. "Do Animals Have a Right to Liberty?" In *Animal Rights and Human Obligations*, edited by Peter Singer and Tom Regan. Englewood Cliffs, N.J.: Pearson, 1989.
- Regan, Donald H. "The Value of Rational Nature." *Ethics* 112, no. 2 (2002): 267–91.
- Regan, Tom. *The Case for Animal Rights*. Berkeley: University of California Press, 1983.
- Rescorla, Michael. "Cognitive Maps and the Language of Thought." *The British Journal for the Philosophy of Science* 60, no. 2 (2009): 377–407.
- Rodd, Rosemary. *Biology, Ethics, and Animals*. Oxford: Clarendon Press, 1992.
- Rollin, Bernard E. *Animal Rights and Human Morality*. New York: Prometheus Books, 1981.
- Rollin, Bernard E. *The Unheeded Cry: Animal Consciousness, Animal Pain and Science*. Oxford: Oxford University Press, 1990.
- Romanes, George John, and Charles Darwin. *Mental Evolution in Animals: With a Posthumous Essay on Instinct by Charles Darwin*. Cambridge University Press, 2011.
- Roskies, Adina L. "A New Argument for Nonconceptual Content." *Philosophy and Phenomenological Research* 76, no. 3 (2008): 633–59.
- Rowlands, Mark. *Animals Like Us*. London: Verso, 2002.
- — —. *Can Animals Be Moral?* Oxford: Oxford University Press, 2012.
- — —. "Moral Subjects." In *The Routledge Handbook of the Philosophy of Animal Minds*, edited by Kristin Andrews and Jacob Beck. London: Routledge, 2017.
- Sachs, Benjamin. "The Status of Moral Status." *Pacific Philosophical Quarterly* 92, no. 1 (2011): 87–104.
- Salt, Henry. *Animal's Rights Considered in Relation to Social Progress*. Clarks Summit: Society for Animal Rights, Inc., 1980.
- Sapontzis, S. F. *Morals, Reason, and Animals*. Philadelphia: Temple University Press, 1987.
- Scanlon, T. M. *What We Owe to Each Other*. Cambridge: Harvard University Press, 2000.
- Schmidt, Andreas T. "Why Animals Have an Interest in Freedom." *Historical Social Research* 40, no. 4 (2015).
- Schroeder, Timothy. "Pleasure, Displeasure, and Representation." *Canadian Journal of Philosophy* 31, no. 4 (2001): 507–530.
- Sebo, Jeff. "Agency and Moral Status." *Journal of Moral Philosophy* 14, no. 1 (2017): 1–22.
- Shettleworth, Sara J., and Jennifer E. Sutton. "Do Animals Know What They Know." *Rational Animals*, 2006, 235–246.
- Sidgwick, Henry. *Methods of Ethics*. 7th ed. Indianapolis: Hackett Publishing Co, Inc, 1981.
- Simmons, Aaron. "Do Animals Have an Interest in Continued Life?" *Environmental Ethics* 31, no. 4 (2009): 375–392.
- Singer, Peter. *Animal Liberation: Towards an End to Man's Inhumanity to Animals*. London: Granada Publishing Ltd, 1977.
- — —. *Practical Ethics*. Cambridge: Cambridge University Press, 1993.
- Skinner, Quentin. "A Third Concept of Liberty." *Proceedings of the British Academy*, 2002.
- Smith, Micheal. *Moral Problem*. Oxford: John Wiley & Sons, 1994.
- Smith, Peter. "On Animal Beliefs." *The Southern Journal of Philosophy* 20, no. 4 (1982): 503–12.
- Steward, Helen. "Animal Agency." *Inquiry* 52, no. 3 (2009): 217–31.
- Stich, Stephen P. "Do Animals Have Beliefs?" *Australasian Journal of Philosophy* 57, no. 1 (1979): 15–28.
- Stout, Rowland. *Action*. Routledge, 2005.
- Strawson, Galen. *Mental Reality*. Cambridge, Massachusetts: MIT Press, 1994.
- Street, Sharon. "A Darwinian Dilemma for Realist Theories of Value." *Philosophical Studies* 127, no. 1 (2006): 109–66.

- Tanyi, Attila. "Pure Cognitivism and Beyond." *Acta Analytica* 29, no. 3 (2014): 331–48.
- Taylor, Charles. "What's Wrong with Negative Liberty?" In *The Liberty Reader*, by David Miller. London: Paradigm Publishers, 2006.
- Taylor, Paul W. *Respect for Nature: A Theory of Environmental Ethics*. Princeton: Princeton University Press, 2011.
- The Mental Capacity Act 2005*, <http://www.legislation.gov.uk/ukpga/2005/9>.
- Thomas, Natalie. *Animal Ethics and the Autonomous Animal Self*. London: Palgrave Macmillan, 2016.
- Tomasello, Michael, Josep Call, and Brian Hare. "Chimpanzees Understand Psychological States – the Question Is Which Ones and to What Extent." *Trends in Cognitive Sciences* 7, no. 4 (2003): 153–56.
- Tooley, Michael. "Abortion and Infanticide." *Philosophy & Public Affairs*, 1972, 37–65.
- Williams, Bernard. "The Makropulos Case Reflections on the Tedium of Immortality." In *Problems of the Self*. Cambridge: Cambridge University Press, 1973.
- Wooldridge, Dean E. *Machinery of the Brain*. New York: McGraw-Hill Inc., 1963.
- Yeates, James W. "Death Is a Welfare Issue." *Journal of Agricultural and Environmental Ethics* 23, no. 3 (2010): 229–41.
- Young, Robert. "The Value of Autonomy." *The Philosophical Quarterly* 32, no. 126 (1982): 35–44.