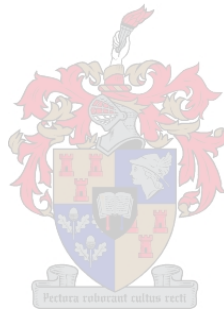


**THE MORAL COMMUNITY AND MORAL CONSIDERATION: A PRAGMATIC
APPROACH**

by
Christopher Stephens

*Thesis presented in partial fulfilment of the requirements for the degree of Master of Arts in the Faculty
of Arts at Stellenbosch University*



Supervisor: Dr Tanya de Villiers-Botha

March 2015

DECLARATION

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: November 2014

Abstract

The aim of this thesis is to argue for a new metric for determining the moral status of another being. Determining this status is of foundational importance in a number of legal, political, and ethical concerns, including but not limited to animal rights, the treatment of criminals, and the treatment of the psychologically afflicted. This metric will be based upon one's capacity to morally consider others. In other words, in order to have full moral status, one must be able to have moral concern for others and act upon this concern to even a minimal degree. In doing so, one will be considered to belong to a "moral community", which affords the member a certain set of rights, privileges, and duties towards other community members. Arguing for the existence of such a community achieves the pragmatic aspect of this thesis. I argue that morality is geared towards group-survival strategies which have been evolutionarily selected for, and thus by organizing societal structures towards the tools which nature has armed us with, we may maximize the powers and capacities of the community members.

In order to achieve these aims, I defend a concept of morality as based in emotion, requiring certain neurological structures, which gives the first set of criteria for identifying potential members of the moral community. I then discuss the issue of identifying the capacity for morality in non-human minds, arguing that we may infer moral capacities from behaviourism.

In summary, the findings of this paper are that first, morality is essentially emotional in nature and is a product of the nature of our neurological system, although rational processes and enculturation shape particular moral sensitivities and priorities. Second, one can infer the existence of moral capacities in animals from their behaviour, and, at risk of engaging in anthropomorphism, to deny these capacities completely entails solipsism. Thirdly, and most importantly, those who are capable of morally considering others ought to be afforded full moral status themselves and be brought into a "moral community" wherein special rights, freedoms, and privileges allow the members to most efficiently contribute to the community, maximizing the powers and benefits of the community.

Abstrak

Die doel van hierdie tesis is om 'n nuwe maatstaf voor te hou waarvolgens die morele status van 'n ander wese bepaal kan word. Die bepaling van hierdie status is van fundamentele belang vir 'n hele aantal regs-, politiese en etiese aangeleenthede. Dit sluit, onder andere, diereregte, die behandeling van misdadigers en die behandeling van diegene met sielkundige probleme in. Hierdie maatstaf sal gebaseer word op die vermoë van die individuele wese om ander moreel in ag te neem. Met ander woorde om volle morele status te hê, moet 'n wese daartoe in staat wees om moreel besorg te wees oor ander en om, ten minste tot 'n minimale mate, na gelang van hierdie besorgheid op te tree. Op grond hiervan kan daar aanvaar word dat daardie wese tot 'n "morele gemeenskap" behoort, wat ook aan hom 'n stel regte, voordele en pligte teenoor ander gemeenskapslede sal besorg. Om 'n argument vir die bestaan van só 'n gemeenskap te maak sal die pragmatiese doelwit van hierdie tesis bereik. Ek argumenteer dat moraliteit ingestel is op groepsoorlewingstrategieë wat evolusionêr geselekteer is. Dit wil sê deur samelewingstrukture op só 'n wyse te organiseer dat dit gebruik maak van die gereedskap waarmee die natuur ons bewapen het, sal ons die bevoegdheid en die vermoëns van gemeenskapslede kan maksimaliseer.

Om hierdie doelwitte te bereik, verdedig ek 'n verstaan van moraliteit as gebaseer in emosies wat sekere neurologiese strukture benodig. Dít verskaf die eerste stel kriteria waarvolgens potensiële lede van die morele gemeenskap geïdentifiseer kan word. Ek bespreek vervolgens die moontlikheid om die vermoë tot moraliteit in nie-menslike verstande te identifiseer en argumenteer dat morele vermoëns vanuit gedragsleer afgelei kan word.

Ter opsomming is die bevindinge van hierdie tesis, eerstens, dat moraliteit wesenlik emosioneel van aard en 'n produk van ons neurologiese sisteem is, alhoewel rasonale prosesse en verkulturering spesifieke morele sensitiwiteite en prioriteite vorm. Tweedens kan die bestaan van morele vermoëns in diere afgelei word vanuit hulle optrede, en, alhoewel ons hier die risiko van antropomorfisme loop, behels die ontkenning van hierdie vermoëns solipsisme. Derdens, en die belangrikste, diegene wat daartoe in staat is om ander moreel in ag te neem behoort self volledig morele status toegeken te word. Hulle word sodoende in die "morele gemeenskap" betrek waar spesiale regte, vryhede en voordele gemeenskapslede sal toelaat om op die mees effektiewe wyse tot die gemeenskap by te dra om sodoende die bevoegdheid en voordele van die gemeenskap te maksimaliseer.

Table of Contents

1. Introduction	1
2. Defining the Moral Domain.....	12
2.1 Introduction	12
2.2 The duel between emotionism and rationalism.....	13
2.3 Veneer vs. core theories.....	23
2.4 Conclusion	32
3. Morality beyond Humanity.....	34
3.1 Introduction	34
3.2 Anthropomorphism and anthropodenialism.....	35
3.3 The computing ape and the kamikaze honeybee	39
3.4 Behaviourism, Batman, and Solipsism	48
3.5 Conclusion	51
4. The Metric and Moral Communities	54
4.1 Introduction	54
4.2 A claim, three arguments, and an observation walk into a bar.....	55
4.3 Means, ends, and rights	63
4.4 Clearing the air: responding to possible objections	68
4.5 Conclusion	75
5. Denouement	77
5.1 Areas for future research.....	77
5.2 Conclusions	79
References	83

1. Introduction

On what grounds do we consider another subject to be worthy of moral consideration? The term “moral consideration” will receive more in-depth discussion in section 3.3 in this thesis. For the moment, let us simply say that it entails “the avoidance of harm and the promotion of care”. When, then, ought we invest time and resources into caring for others and avoid causing them harm? This thesis is aimed at answering this question. I will argue that, on pragmatic grounds, the capacity for the moral consideration of others is itself the primary justification for incorporating others into a community of moral concern. In doing so, I will argue, the moral community will be able to reap considerable pragmatic rewards via maximizing the members’ usefulness to society.

While an idealist may disagree with this position and prefer to answer that subjects are always worthy of moral consideration, throughout the natural world we often observe a necessary life/death dependency between two or more species where such concern is not possible. Indeed, death appears to be as much a part of life as birth, and one can hardly begrudge a lion for eating an antelope; it is a necessary consequence of the nature of the lion and the nature of the antelope. While this is a statement of that which *is*, some would argue that this is not what *ought* to be.

The question of how one ought to treat others has been the subject of much debate over the centuries, and has formed and informed many pertinent issues such as the raising and eating of animals, the treatment of criminals, the extent to which abortion is legal and so forth. This debate has come by many names such as the “moral standing” of an individual, their “moral status”, and “moral considerability” (Jaworska & Tannenbaum, 2013). I will be using the latter terminology in this thesis.

The common-sense notion of moral considerability suffers from numerous contradictions. Jaworska and Tennebaum point out that while there is general consensus of affording full moral consideration to human beings, whether cognitively impaired or not,

there is a great deal of disagreement when it comes to non-humans. This is so even though providing an acceptable theory that unambiguously accounts for all humans without also incorporating many animals has been a problematic task (*ibid.*)¹

To be morally considerable, or to have moral status, bears with it the three suppositions. The first is the presumption against interference, which is akin to the avoidance of harm although it has the broader implication that you may not interfere with another even if there would be no harm. However, this presumption may be overridden in special cases, especially where interference is necessary to prevent harm to others (*ibid.*). The second presumption is that of rendering aid to those who are morally considerable – to promote care. This presumption is not as stringently held as the first, possibly because rendering aid usually comes at a cost to oneself, while non-interference merely requires restraint. The third supposition for those of full moral status is that they ought to be treated fairly. This is not the same as saying they ought to be treated equally, although given equal need and no other differentiating considerations, the treatment ought to be so (Jaworska & Tannenbaum, 2013). In my model of moral consideration, I hope to show that we have a real, pragmatic motivation for ensuring that these three suppositions for those who are worthy of moral consideration are adhered to.

Of course, before we can go about promoting these suppositions, we need a clear understanding of what criterion (or criteria) is needed for one to be morally considerable. We can categorize the justifications for moral status into two camps: it is either a matter of degrees, or an all-or-nothing account. Jaworska & Tannenbaum note that while some theorists acknowledge that moral considerability may be a matter of degree, none have managed to develop a detailed account of what each degree might require or provide. In the earliest drafts of this thesis, I initially intended to provide such an account, however it

¹The reason as to *why* we intuitively favour the human species is thought to be an evolutionarily encouraged practice of kin-selection which may be overgeneralized through anthropomorphism. This will be discussed further in chapter 3. For a detailed discussion of this selection-bias, refer to Rushton et al.'s excellent 1984 research article *Gene Similarity Theory: Beyond Kin Selection*.

proved to be a larger and more complex model than could be presented in thesis.

Nonetheless, in the concluding chapter I leave some suggestions in how one might build such a model.

Many models of moral considerability are reliant upon the subjects' cognitive capacities for their justification (*ibid*). For example, we might say "I can eat this beef because cows are stupid", and let us assume that they are indeed so. However, where might this leave newborns, or the mentally handicapped? May we then eat them too? To defend against this, some will make what is known as the "potential account", where the future potential cognitive capacities of the subject factor into its considerability. This potential account is one which I will also advocate. Since my argument is a pragmatic one, it is concerned with future benefits which potential capacities will eventually feed into. However, at what point would one not be smart enough to be morally considered? We might have some scalar quantification – the more intelligent a being is, the greater their worthiness – or we might hold a threshold view, that as soon as one demonstrates some intelligent capacities one is awarded full moral status (*ibid*). In this thesis I will endorse a threshold view, although I will also offer a possible method to adjust the model to a scalar model.

Let us take a look at the modern vegan movement as an example of some the above ideas. It argues that human beings are not the only animal worthy of moral consideration. This is communicated in various ways, most overtly through the PETA slogan "animals are people too". It is also communicated more subtly by anthropomorphising animals; in one series of advertisements by PETA, a calf is shown walking through a meadow, while a child-like narrator's voice states "one day I will be big and strong, just like my dad" before cutting to scenes of a slaughterhouse.

Putting aside the assumption of the exact nature of a cow's consciousness, this suggests that certain characteristics are required for a subject to be considered worthy of

moral consideration. As mentioned above, intelligence has been the metric by which we often measure this worthiness (Singer, 1997: 17, 20, 22; 2006: 18), and arguing for the intelligence of a given species is the strategy by which they are afforded certain privileges above that of other non-human animals. An example of this would be the Great Ape Project, a campaign that has found political traction in both England and Spain to grant chimpanzees and other higher primates access to legal and moral rights on par with those of humans (Cavlieri & Singer, 1993: 104-106).

This move to incorporate animals into an expanding sphere of moral concern is a modern departure from the traditional paradigm of humans-separate-from-nature (Vining, et al., 2008: 1, 8). This traditional paradigm has been deeply ingrained into Western thinking, particularly in the Abrahamic religions as is evidenced from the Christian Bible:

Then God said, "Let us make mankind in our image, in our likeness, so that they may rule over the fish in the sea and the birds in the sky, over the livestock and all the wild animals, and over all the creatures that move along the ground.

(NIV, Genesis 1:26)

Here it is apparent that the human being is thought of as having an element of the divine because of being made "in His image" – that is, to have certain God-like qualities (Botterweck & Ringgren, 1997: 259), as opposed to the non-divine animals. This divinity was bestowed upon humankind for the purpose of having mastery over the other forms of life. This paradigm of separation is not unique to western thought either. Teffo and Roux describe the hierarchy of African metaphysics as "God at the apex and extra-human beings and forces, humans, the lower animals, vegetation and the inanimate world, in this order, as integral parts of one single totality of existence" (2003: 167). But it is not only on religious grounds that the paradigm has come about. Richard Dawkins states:

It isn't just zoological classification that is saved from awkward ambiguity only by the convenient fact that most [intermediate species] are now extinct. The same is

true of human ethics and law. Our legal and moral systems are deeply species-bound.

(1996: 262)

This humans-separate-from-nature paradigm is termed “speciesism”. The philosopher and advocate for the veganism movement Peter Singer defined speciesism as “a prejudice or attitude of bias in favour of the interests of members of one's own species and against those of members of other species” (2002: 6).

In contemporary discourse, words like “prejudice” have become loaded with negative connotations. However, the true concern, I believe, is not with the prejudice itself, but rather the unjustifiable discrimination. Thus, for the purposes of this thesis, we will use the more nuanced definition of speciesism offered by Fjellstrom:

A normative opinion O, held by a person P, is speciesist if, and only if

- (i) O favours at least the normal members of P' s own species, against the normal members of all other species;
- (ii) O is, in P's intellectual milieu, not grounded –or grounded on validational tools that assure the favouring of P's own species; and
- (iii) No high level meta-ethical justification is offered in P' s milieu of the criteria used that is not implausible to philosophical common sense, in particular by being non-partial.

(2002: 71)

This definition is preferred to Singer's, as under Singer's stipulation there is no possibility for dissent for any reason. That is, it automatically assumes that a speciesist act is at all times an arbitrary act that cannot be justified. Fjellstrom's definition allows for a dialog of ideas to take place, it does not presume its own truth.

The anti-speciesism of Dawkins and Singer does not stand unopposed. It is countered by Daniel Dennett, who argues that consciousness is not a binary opposite as is

commonly understood; that is to say, an organism either has the capacity for sapient, self-aware consciousness or it does not (1995: 694-696). Rather, he argues, consciousness is a construct brought about by the formation of data—a structuralist approach that suggests human culture is a key factor in the understanding and awareness of the self (*ibid*). The implication of this view for Singer, whose position on animal rights is chiefly concerned with suffering (Singer, 2002: xxi, 7, 8), is that it is nonsensical to argue for the equal consideration of suffering when the experience of suffering itself is unequal. In Dennett's own words: "Such body-protecting states of the nervous system [reflexive pain] might be called 'sentient' states without thereby implying that they were the experiences of any self, any ego, any subject. For such states to matter -- whether or not we call them pains or conscious states or experiences -- there must be an enduring, complex subject to whom they matter because they are a source of suffering." (1995: 711) Dennett further points out that any descriptions of pain as being intrinsically awful is "circular and question begging" (Newman, 2014: 7).

However Dennett's position only provides a potential counter to arbitrary or undifferentiated speciesism where Fjellstrom's point (iii) is not taken into account, such as the belief that an ant is equal in all respects to a chimpanzee and ignores aspects like intelligence, social mechanics, or emotional capacities, and only deals with the nature and capacity to suffer.

It is possible that there exists a species other than human beings which satisfies Dennett's requirements² for consciousness. It is thus my aim to provide a pragmatically normative framework that can be applied to any conscious subject and is not necessarily limited to human beings.

² Dennett has been criticized for not providing concise requirements for consciousness (Newman, 2014: 6), however he has described it as being necessarily a massively parallel process of sensory input and interpretation which creates "something *rather like* a narrative stream or sequence" (1991: 135, original emphasis).

The concern over how to measure a subject's worthiness of moral consideration extends beyond the realm of animal rights. Robert Freitas, Carl Sagan, and others involved in astronomy and space exploration have written extensively on the problem of contact with extraterrestrials; not merely on the issue of communication, but on how two wildly different species would view, consider, and understand one another. Furthermore, this issue can also be applied to artificial intelligences as evidenced by various works of speculative fiction such as Isaac Asimov's *The Bicentennial Man*, and Philip K. Dick's *Do Androids Dream of Electric Sheep?*, the basis for the 1982 movie *Blade Runner*. In these works the question of personhood, moral responsibility, freedom, and access to moral resources is applied to machine intelligences as well. While we may be some decades away from such an electronic facsimile of human consciousness—and sceptics such as Searle believe that we may never achieve it (Cole, 2014)—it does nonetheless broaden the question to include even potentially non-biological forms of consciousness.

The examples and case studies presented in this thesis will come from the field of animal behaviourism and the animal rights debate, since we have not yet come into communication with alien or machine intelligences. Despite that, I would like to stress that it is not my intention for this to be a thesis on the issue of animal rights, but rather on the question of both ethical and practical treatment and communication between different consciousnesses, even within the human species. Is it morally permissible to use a psychopathic human being – who has not committed a crime – as a means to end, for example? If we create a sentient computer program are we morally compelled to keep it powered? If we ever come in contact with an alien race who, through the very biological structure of their brain cannot take the perspective of another, would we be able to establish a social community with them that goes beyond economic concerns?

In this thesis I aim to provide the meta-ethical grounds from which we may reap some answers for the above questions. Before I do so, I will begin by answering two

questions upon which my thesis is based, namely, what is morality, and who owns it? To answer the first question, in Chapter 2 I will discuss the metaphysics of morals: those particular features which identify a given property or argument as belonging to the moral sphere. This discussion will mostly rely upon recent neurobiological research, as well as work in psychology and sociology. I will argue that morality is neither purely rational nor purely emotional, but is rather an emergent property of these two mental processes. I will also argue that morality is chiefly a biological artefact, subject to shaping through sociological forces, but primarily a matter of biology and a product of evolutionary forces.

The second question, “Who owns morality?”, will be answered in Chapter 3. While it may seem a strange question to ask, it will become apparent in the second chapter that it is commonly held belief that morality is exclusively a human trait, and that to see it in animals is grave error of judgement. Thus Chapter 3 will necessarily also deal with the problem of anthropomorphism. In this chapter I will argue that morality exists beyond humanity, that it can be found in birds, bees, and chimpanzees (among others). However, I will also argue that not all morality is equal. I will argue that the “morality” of bees, known as kin-altruism, is self-serving. I will also argue that while birds are capable of a “true” altruism, its scope is so narrow as to make it effectively useless to a moral community.

It is in this discussion of the birds and bees that another important capacity will be discovered: that of targeted helping. Targeted helping is in itself not a moral trait; it is the capacity to identify the cause of another’s suffering, create a strategy to alleviate the suffering, and hold the ability to execute the strategy. While this is the bedrock of intentional altruistic behaviour, it is also the bedrock of trade and certain forms of manipulation. Nevertheless, in order for the moral community to achieve its maximum pragmatic and utilitarian potential, targeted helping will be a prerequisite to for membership to such a community.

With these three concepts in hand – what morality is, who might have it, and how one uses it – I will present my main argument in Chapter 4. Ultimately I wish to show that, in describing and classifying our actions and interactions, there exists a special community, a “moral community”. Those who fall within this moral community are those who we see as fellow beings whose freedoms, ambitions, fears, and hopes are treated with a similar degree of respect as our own. With the broadest strokes of the brush we can say that beings in this moral community may not be treated as a means to an end, but only as ends in themselves (Kant, 2002: 46-47).

In ancient times moral communities and cultural communities were indistinguishable: the Romans raided and enslaved barbarians just as the barbarian tribes raided and enslaved the Romans. However, with the progression of technology the world has become increasingly cosmopolitan where cultural influence has less and less to do with national borders, and the French and Italians may happily rub shoulders in Shanghai. In short, the gamut of humanity has been largely incorporated into a single moral community.

Of course one might point out that people still engage in slavery, in genocide, in rape, and many other activities which undermine the value of beings inside the moral community. Certainly these terrible things occur, but we do not regard those who perpetrate such crimes as belonging to the moral community. As such, we engage these threats to the moral community in a markedly different manner to how we would deal with a momentary transgression from a member of the moral community. A psychopath, for example, who commits multiple murders may be given the death penalty, while a spouse who has been cheated on and kills in a fit of jealous rage may be given ten years of imprisonment³.

³ One could read this as an endorsement of capital punishment, and indeed the idea of a moral community may lend weight to such an endorsement. However, in the interests of full disclosure I would like to point out

This point highlights why I decided on the particular term “community” over other possible terms, such as “class” or “group”. Firstly, communities are not necessarily groupings of equals. In the economic community of a town – an analogy I will return to often in this thesis – the doctor may produce more valuable economic goods than the janitor, but both are capable of the give/take actions which are fundamental to the existence of the economic community itself. Likewise in a moral community a child suffering from Down’s syndrome might not be able to comprehend the complexities of Kant’s or Hume’s moral theories, but is still able to perform the essential avoidance of harm/promotion of care behaviours required for a moral community to exist. Secondly, communities promote themselves as a whole while either utilizing elements outside of themselves to their own advantage or removing outside threats. Thirdly, a community is often better defined by its interactions within itself than its interactions with elements outside itself, and since this is a question of moral behaviour, the term is most fitting.

Communities, however, are not ever-inclusive. As with my above analogy of the town, one cannot belong to the moral community without being able to contribute to it, even if only at a minimal level. Thus the boundaries of the moral community are set at those beings that are able to morally consider others. This is a point that distinguishes my thesis from the position of many animal rights activists – the issue of animal rights being inextricably tied to this discussion, as I mentioned earlier.

Philosophers such as Singer would argue that we ought to promote care and avoid harming any being capable of forming preferences (1993: 94-97; 1997: 206, 275; 2002: 9), preferences which are based on the being’s ability to suffer, regardless of its moral capacities. However this I believe to be erroneous on two points: firstly, in terms of utility it makes no sense to invest in something which will never reciprocate; and secondly, in that an over-commitment of finite resources provides sub-optimal benefit and even has the

that there are numerous arguments against capital punishment which I believe considerably outweigh this apparent endorsement.

potential to harm as I will argue in Chapter 4. There is a considerable difference between giving ten dollars to one beggar, or one dollar each to ten beggars. Both acts cost me equally, yet only in one scenario can anyone afford a meal.

So long as supply falls short of need, we must make decisions as to how to invest limited resources. My thesis is that it is ultimately to the maximum benefit to most sentient beings to invest in the moral community. Further, since members would be of the sort who seek to improve the world for others, the moral community may even be in the interests of beings outside the community in the long term.

2. Defining the Moral Domain

2.1 Introduction

According to Gert (2011), the term “morality” can be used either

- 1) Descriptively, to refer to certain codes of conduct endorsed by a culture, society, or by an individual for their own behaviour, or
- 2) Normatively, to refer to certain codes of conduct that, in certain contexts, would be put forward by all rational persons.

While this thesis posits a pragmatic normative claim, that we ought to only have moral concern for subjects with moral capacities, it is a claim that rests on descriptive moral properties. To that end, this chapter is dedicated to defining the moral domain.

“Certain codes of conduct” which Gert offers is a very vague definition for what constitutes moral behaviour. As we will see in this chapter, moral behaviour exhibits certain properties that set it apart from legal, religious, or social codes of conduct. It is the elucidation of these properties that constitutes “defining the moral domain”. In this way, if we encounter some belief of a culture that we do not hold as a moral belief, we can still be able to recognize it as being so from the perspective of that culture.

The analysis of the moral domain will be broken down into four sections. Some of the issues raised are centuries old, while others are recent additions from the fields of neuroscience and animal behaviourism. We will first consider whether moral beliefs are primarily constituted by rational beliefs, as Kant believed, or emotional beliefs, as Hume argued. We will then consider whether morality is a “cultural veneer” that can be washed away as T.H. Huxley and Jesse Prinz believe, or if it is a core biological process as Darwin and F.W. de Waal believe. Once these structural elements of morality have been explored, we will then look at the broad categories or “dynamics” which define the form that moral beliefs take.

2.2 The duel between emotionism and rationalism

Whether morality is derived from rational or emotional processes is a debate well-known to philosophy. Two exemplars of this debate are Kant and Hume, with Kant supporting rationalist theories and Hume widely considered to be the forbearer of the emotivist⁴ movement. Until recently both moral philosophy and moral psychology have been dominated by some permutation of the rationalist model, with less attention being given to emotivism (Greene et al., 2004: 389; Haidt, 2001: 814; Kohlberg, 1969: 512, 530; Turiel, 1983: 42, 168). This question is of particular importance to this thesis as it will answer the question as to what can be said to be capable of morality, and thus belong to the moral community. For example, should a true artificial intelligence be created, but be incapable of emotions, could it be capable of moral agency?

A rationalist metaphysics of morals makes the claim that morality is constituted by claims that are supported by rational premises. These premises are in a sense publicly observable, in that they are not subjective mental events like emotions, and are open to scrutiny for truth-value. Moral rationalism thus usually entails moral objectivism, where a moral claim can have a definite Boolean value as either true or false.

Emotivist moral theories are based on the idea that moral judgments do not entail statements about the state of the world; rather they are expressions of our attitudes. Thus to make declaration “stealing is wrong” is no different to cry “stealing!” in revulsion (Prinz, 2007: 17). Strictly speaking then, “morality” is a special case of emotional reactions and not an independent concept. For Hume, one of the earliest and perhaps most iconic emotivists, the nature of these reactions are predicated on the basis that ideas, including moral ideas, are “stored copies of sensory impressions” (2007: vii). It follows then that the

⁴ I will use the terms “emotivism” and “emotivist” to generally refer to the schools of thought and their advocates which identify the emotions as the primary property of moral beliefs, even where some schools may refer to themselves with their own nomenclature, such as Prinz’s “Constructive Sentimentalism” (2007: 167-169). This discussion is more concerned with the primacy of the emotions or of rationalism in morality, and not the nuances involved.

ideas of “goodness” and “badness” are the result of the sentiments of approbation and disapprobation – the positive and negative impressions respectively – we experience when performing the relevant acts.

More recently, emotivist theories have garnered further support from moral psychology with the development of the CAD Triad Hypothesis.. The aim of the study was to map three principle emotions - contempt, anger, and disgust - to three principle moral paradigms - community, autonomy, and divinity respectively (Rozin et al., 1999: 574). That is to say, a threat to a person’s autonomy such as personal insults, theft, or assault, would elicit anger in the victim. Seeing a transgression of communal codes, including social hierarchy – such as seeing a teenager begin eating at the dinner table before everyone else is served – arouses contempt for the transgressor. The principle of divinity is also referred to as the principle of purity, and is theorized to have had evolutionary origins in disgust or distaste for disease or rotten food (578). An example of a transgression of divinity or purity would be an act of incest. Contempt, anger, and disgust, the authors argue, are the primary colours from which we paint our varied pictures of morality across cultures (575).

In the study, conducted separately in Japan and the United States, participants were asked to match a variety of scenarios, including moral transgressions, to appropriate facial expressions (576-577, 579). For example, the “anger” facial expression featured narrowed eyes, a furrowed brow, and pursed lips.⁵ The authors found a great degree of similarity between the American and Japanese responses. In the given situation of a teenager eating at a dinner table before the rest of the family is served, 75% of American interviewees identified “contempt” as the appropriate reaction, as did 73% of Japanese interviewees (578).

⁵ While the study also asked the participants to match situations to emotional words, the authors admit that certain subtleties between English and Japanese may have skewed results (Rozin et al. 1999: 581).

The results showed that the “moral emotions map rather cleanly onto three different moral codes”, even across the cultural divide between the US and Japan. Had these emotional responses been a matter of cultural conditioning, the predicted result would be divergence between cultural borders, not similarity. Rozin et al.'s research thus indicates the existence of an apparently culturally-independent, emotionally-based moral structure (585).

While emotivist theories have been enjoying empirical support in recent years, rationalist theories have been unable to account for a number of behavioural anomalies found in anthropology, primatology, and biological psychology (Haidt, 2001: 814). Among these anomalies is the ubiquity of morality and moral behaviours. If morality is a product of rationalization, Andrew Sneddon (2007: 735) argues, then only those who have engaged in the necessary thinking and discourse would be capable of moral behaviour, and yet moral behaviour is so commonplace. Further, there has been research to show that violent and sociopathic behaviours are greatly increased when a child is raised in an “emotionally malicious” home, regardless of education or propensity towards rationality (Poulshock, 2006: 31).

In an attempt to establish a new descriptive moral paradigm to account for these anomalies, the Social Intuitionist Model (SIM) was developed as an affective, intuitionist-based model of moral judgment which minimizes, although does not eliminate, the function of rational processes in moral judgements. Here rational processes are effectively reduced to secondary, post-hoc processes of rationalization. The SIM thus predicts rational thinking only occurs after the moral judgment has been rendered as post-hoc reason-giving. Jonathan Haidt, the developer of the SIM, argues that the issue of morally dumbfounding a person serves as evidence for an intuition-based judgment system (Haidt, 2007: 998; Prinz, 2007: 30). When a person is morally dumbfounded they claim that a given action or situation is morally impermissible, while being unable to give reasons why.

In an experiment, Haidt (2001:814) asked participants if it was morally permissible for two siblings on holiday in France to commit an act of incest, even when every form of protection was taken, and the act was kept a secret which actually strengthened their relationship. Eighty percent of respondents claimed it was morally impermissible, but when pressed for reasons, could not provide any. In a similar study, Prinz (2011: 3) asked participants if it was morally permissible for a paedophile to fondle an infant, even if the infant was not physically or psychologically harmed in any way, and would not even remember the incident. One hundred percent of respondents stated that it was impermissible. Prinz, unfortunately, did not stipulate whether he asked for reasons. However, I strongly believe very few people would be able to respond with anything other than "it just is!" This implies that there is some given set of moral beliefs that are emotionally rooted and not initially motivated by rationality. The question arises, however, do the emotivist explanations of morality given by Haidt and Prinz describe the complete set of all moral beliefs?

If the answer to this question is negative, then we ought to be able to demonstrate that moral dilemmas are, at least at times, resolved rationally, as this will fall outside of the scope of predictions of Haidt's and Prinz's theories. In order to achieve this we need to first identify a category of responses to moral dilemmas that can be reliably classed as rationalist, then we need to find one or more moral dilemmas where respondents apply this category of solution.

A good candidate for a rationalist moral system seems to be utilitarianism. The commitment to doing what creates greatest benefit for the greatest number of subjects will eventually call for an adherent to resist his or her own emotional urges and perform the calculated optimum action. There are two empirical studies which demonstrate utilitarianism to be a rationalist mode of moralizing. The first is Greene et al.'s 2004 functional Magnetic Resonance Imaging (fMRI) scans of brain activity during moral

deliberations (2004: 393-395). In this study, the timing and activity levels in the anterior dorsolateral prefrontal cortex (anterior DLPFC, the region of the brain involved with abstract thinking) of patients making utilitarian responses and non-utilitarian responses to given moral dilemmas were compared. Volunteers were presented with moral dilemmas such as the trolley dilemma and the crying baby dilemma⁶. Higher activity in the anterior DLPFC was detected with the utilitarian responses, and higher activity in the emotional and social centres of the brain for the non-utilitarian responses (Greene et al., 2004:393-394). This suggests that utilitarian morality is fundamentally rational in nature, and deontological morality is situated in a social-emotional matrix⁷.

With regards to the timing of activity in the anterior DLPFC, in utilitarian responses activity was relatively high—up to ten seconds before a response was given—and was the highest at the point of response before falling back to normal readings (*ibid.*) For non-utilitarian responses, however, anterior DLPFC activity was below normal and in fact dropped sharply at the point of response, only to increase to above-average levels eight to ten seconds after the response. This late increase of rational activity complies very well with Haidt's SIM, which predicts intense rational processes only after a moral judgment is made as post-hoc defence and rationalization, something which is not necessary if the response was achieved by rational processes, as is the case with the utilitarian responses. This suggests that the intuitive responses here followed a model similar to what is predicted in the SIM. Contrary to the SIM, however, the increase in activity in the anterior

⁶ In the crying baby dilemma, you and several others are hiding in a house which enemy soldiers are searching. They will kill you if they do find you. Your baby begins to cry, and the option is to smother your baby, saving your group but killing the infant (the utilitarian option), or through inaction allowing the soldiers to find your group (the non-utilitarian, deontological option). In the standard trolley dilemma originally formulated by Foot (1967) – there are many important variants – an out of control railway trolley is speeding towards five railway workers. You are standing near a control switch that will shift the course of the trolley onto tracks that will hit one railway worker. To switch the tracks reduces the number of deaths, entailing a positive utilitarian choice, but you become an active agent in the death of another, entailing a negative deontological choice.

⁷ One might argue that the Kantian categorical Imperative is both rational and deontological in nature. I do not believe this to be entirely true. Consider the first maxim of universality. If we are deliberating whether or not to engage in an extramarital affair, and imagine a world where affairs are universal, it may very well be our emotional reaction to that imagined world that informs our reply to the first maxim. Such an interpretation is probably not what Kant intended, and I wish to be cautious of accusing Kant of being in error, but the possibility does exist.

DLPFC prior to the utilitarian responses suggests that these moral judgements were a product rational processes without post-hoc rationalization.

The second study which demonstrates the rationalist nature of utilitarianism was conducted by Fumagalli et al. (2010). The experiment involved transcranial direct current stimulation as a non-invasive method for inducing stimulation in the ventral prefrontal cortex, which houses the anterior DLPFC, which performs high-order rational thought, in an attempt to alter the proportion of utilitarian judgments made by participants (Fumagalli et al., 2010). This research was inspired by a comment made by Greene et al. who stated that it is possible that the nature of their fMRI tests may have caused participants to try and “push through” to a rational response, thus distorting the results (2008: 1151). Greene et al. hypothesized that a less intrusive technique such as transcranial direct current stimulation⁸ – the selective excitation or impairment of specific groups of neurons in the brain – would avoid such a problem and provide more accurate results.

It was found that area of speciality in education – life science vs. human science – had no significant effect in the favouring of utilitarian decisions, nor did religious upbringing (Fumagalli et al.: 4)⁹. However, all of the volunteers were educated to at least tertiary level, so this does not answer any questions on the relation between quantity of education and utilitarian judgment. What did play a major role was the sex of the volunteer: transcranial direct current stimulation, whether anodal or cathodal, had no effect in altering the proportion of utilitarian to non-utilitarian judgments in men. In women, when tDCS was applied in the excitatory, anodal mode over the prefrontal cortex, utilitarian decisions became significantly more common and were made more quickly (*ibid.*) Unlike the male

⁸ Transcranial direct current stimulation (tDCS), or the related technique of Transcranial Magnetic Stimulation (TMS), is a relatively safe method of altering functionality of a particular region of the brain using weak electric currents or magnetic waves, respectively. The alteration could be inhibitory in the cathodal mode: by hyperpolarizing neurons into a positively charged state, it makes it impossible for the neurons to fire properly, effectively a temporary brain lesion (Gazzaniga, et al. 2014: 89). tDCS or TMS can also be done in the anodal state, bringing the resting state of the neurons closer to the voltaic threshold needed to fire, and thus improving performance in the affected area (*ibid.*)

⁹ This might suggest that our moral conduct is innate, and not so much a taught cultural veneer. This question is looked at more closely in the next section.

subjects, who were unaffected, cathodal tDCS caused far fewer utilitarian decisions, and those that were made took much longer to make (3)¹⁰.

These results are puzzling to say the least. One might predict, on the hypothesis that utilitarianism is a rational process, a decreased response time for utilitarian thinking as it was the rational region of the brain that was primed. And yet, while this prediction was confirmed, there was also a reduced response time for non-utilitarian judgment which was not anticipated. The only speculative explanation for this is that the stimulation of the rational centres of the brain pushed borderline or ambiguous cases over to a rational response, allowing for only the most definitive emotional responses to overpower rational responses, which would allow for more rapid reaction times. As for the discrepancy between men and women, the authors do not attempt to explain it beyond pointing out that it exists, and stating that it may explain why men tend to be more involved in criminal activity than women (5)^{11 12}. This is quite a telling conclusion if we accept their theory. In the experiment, women were more prone to emotional responses than men, even after the transcranial stimulation. The implication here being that feeling about morality is better than thinking about morality when it comes to actually acting morally. Regardless, taken in themselves, the results on the women test subjects shows utilitarian thinking to be rationalist in nature.

We have seen that to be a utilitarian is to be a rationalist. But is utilitarianism a normative code of conduct only, or can it be a descriptive theory of how we usually engage with moral dilemmas?

¹⁰ That a matter of biology would have such a glaring effect on our moral conduct further suggests morality as an innate process. The fact that other social influences such as type of education or religious background does not change one's responses is also indicative of an innate morality, rather than a socially instituted one.

¹¹ I can only hazard a guess that either the male brain does not respond to the exact techniques and settings used in the experiment, or that utilitarian deliberation occurs in a different region for men.

¹² Their statement on the relation between utilitarianism and criminality is based on what I infer is an unstated assumption that calculating utility is not necessarily an other-directed task. In other words, the ability or propensity to calculate the optimal outcome in a given situation is likely to be selfish or egotistical action unless we have some motivation – such as empathy – to care for the outcome of others.

Studies by Greene et al. (2001; 2004; 2007; 2008; 2009; also Paxton & Greene, 2010) incorporating fMRI scans, thinking about moral dilemmas under a cognitive load¹³ and reaction tests, seem to suggest a dual-process model of moral judgment. The two processes are those of moral reasoning and emotional responses, which are managed by a “conflict monitor” (Paxton & Greene, 2010: 514). This conflict monitor is hypothesized to be the anterior cingulate cortex, whose role is to detect conflicts in thoughts and beliefs – such as conflicting emotional and rational judgments – and to select or mediate an appropriate response (*ibid.*) Greene’s model holds rational processes to be vital and fundamental in making moral judgments. If this is the case, then Haidt’s and Prinz’s models do not describe the complete set of all moral beliefs.

In the 2001 study, Greene et al. compared responses between two variants of the trolley dilemma. In the standard dilemma an out-of-control railway trolley is speeding towards five workers who are bound to the tracks. You are near a control switch that would divert the trolley onto another path, where only one worker is tied to the tracks. When presented with this scenario, most people approve of the utilitarian action, pulling the control switch to alter the course of the trolley, condemning one person to death to save five (Greene et al. 2001: 2105). In the footbridge variant, both you and a worker wearing a heavy backpack are on top of a footbridge that spans the tracks. The combined weight of the worker and the pack are enough to stop the trolley, but there is not enough time for you to take the pack from him. Despite the apparent popularity of utilitarian judgments in the standard version, in this instance very few people chose to push the worker off the footbridge in order to save five lives (*ibid.*)

Greene et al. conducted a second experiment, modelled on the first, testing three factors in the trolley/footbridge dilemma: physical contact, spatial proximity, and personal

¹³ Put simply, a cognitive load is a distracting task that requires concentration, such as the Stroop task. In the Stroop task subjects must say aloud the colour of the ink a word is written in, and not the word itself. For example, “RED” may be written in green ink, making the correct response “green”.

force. Greene denotes personal force as follows: “The force that directly impacts the other is generated by the agent’s muscles, as when one pushes another with one’s hands or with a rigid object” (Greene et al., 2009: 365). For example, in one scenario used in the experiment “Joe” uses a pole to push a worker off of the footbridge, which is negative for physical contact, but positive for personal force.

The experiment revealed that spatial proximity played no statistically significant role in judging the moral appropriateness of an action (366-367). In fact, there was a small increase in the number of utilitarian responses when Joe uses a switch to drop the worker on to the tracks, but is standing close by when he does so (negative for contact and force, positive for proximity) (*ibid.*) Of the three factors tested, the most significant was personal force, followed by physical contact, with spatial proximity playing no significant role.

The lesson to be taken from Greene’s experiments so far is that utilitarianism is a descriptive model of our moral judgments in certain contexts¹⁴. In the scenario where Joe is standing next to the backpack-wearing worker when he flicks the switch, killing the worker and saving five others, it was a highly approved-of action (Greene et al., 2009: 366). Approval dropped only once Joe started to invest large amounts of effort to accomplish the same goal. Arguably, the respondents were not attempting to comply with an artificial, normative ethical theory. They were presented with a dilemma and responded accordingly. This demonstrates that utilitarianism is a natural method of approaching moral dilemmas, which is what a descriptive model of ethics needs to account for.

By themselves, rationalist and emotivist moral theories have always suffered a blind spot in their descriptive powers. As we have seen, utilitarianism is not a normative theory

¹⁴ It is worth noting that this may not be true, as it is something of a matter of perspective. In non-moral dilemmas we are also utilitarians. If I am deliberating whether to go to gym tonight or to go and watch a movie, I am weighing up the costs and consequences of my choice in the same manner one approaches the footbridge dilemma. Nakamura (2013: 806-807) showed that by adjusting the numbers of workers in the footbridge dilemma, even by small amounts, one can easily convert people to and from utilitarianism. Changing the lone worker to a pair of workers halved the number of utilitarian responses, even if it meant saving five lives. It may be more accurate to think of utilitarianism as our “default” stance, a stance which is eventually overcome by moral emotions.

only, but rather describes how people spontaneously choose to resolve some moral dilemmas. And yet, when we engage in high levels of effort, an emotive approach becomes the more accurate model. It is only in combining both emotional and rational models that can we explain both types of behaviour, yet we are still stuck with the discrepancies in the variants of the footbridge dilemma. In order to explain our flip-flopping between utilitarian and emotive responses a final thread needs to be added to tie the two processes together, a thread that will explain why and how we choose between the two.

I believe this thread to be that of emotional responsibility. As our effort towards a goal increases, so too does the perceived joy of success or the pain of failure; as our effort invested in solving a moral dilemma increases, so too does our emotional investment in the dilemma. This thread explains Greene et al.'s findings regarding the application of force and the propensity to pursue deontological solutions to a dilemma. When we are emotionally charged through the investment of time, energy, and attention, we are less likely to act rationally. The net result of this is that normative moral theories – at least those that entail a purist approach to moral deliberation such a strict deontological or utilitarian approach – will have an additional burden: they will need to elucidate *how* to maintain the frame of mind required for the mode of moralizing. This is the inevitable outcome of the precept “ought implies can”; it makes no sense to say that I am morally obliged to pick up the boulder that is crushing a child unless I am in some way able to do so. Likewise, I can only obey a normative dictum to always make utilitarian judgments if I am able to divorce myself from any powerful moral-emotional forces that compel me to act otherwise. For example, if we took the standard trolley dilemma and exchanged the lone trapped worker and replaced it with a toddler, I expect that a steadfast utilitarian would be exceedingly hard to find.

The dual process model is an adaptable and accurate model of moral deliberation. When coupled with emotional responsibility it explains our inconsistencies when faced with near-identical moral dilemmas in ways purist theories simply cannot.

One possible objection one might raise is that the DPM suffers from a lack of parsimony. It is indeed a more cumbersome model than those before it; the entirety of the SIM lives within a single node of the larger DPM, and who knows what maze of flowcharts might be found to be living in the others? Ockham's razor looms large here. However, it can only be applied to unnecessary extravagance or desperate ad-hoc defences. Every node of the model is necessary to explain the apparent moral inconsistencies that people engage in when faced with moral dilemmas. To cut even one in the interests of parsimony will result in anomalies and the breakdown of the descriptive model. Perhaps this structural extravagance is to be expected; as I stated in the introduction, this thesis considers evolution to be the de facto case, and we will see in both the next section and Chapter 3 evidence that morality is derived from evolutionary processes. Evolution is hardly a clean process. Vestigial organs and genetic throwbacks are often observed, leading to a degree of encumbrance one would not see in a perfectly engineered design. Perhaps one of the two processes is a fading biological relic, in the process of being replaced by a more efficient method of obtaining social cohesion. In any case, the argument from parsimony does not seem to mount much of a threat to my findings.

2.3 Veneer vs. core theories

So far we have seen how that which constitutes the moral domain incorporates both emotional responses and rational arbitration. Haidt and Prinz both declare their positions as naturalists and evolutionists, while Greene makes no mention of the issue. They also both state that they believed that evolutionary processes furnished human beings with the essential neurological and somatic hardware necessary for a person to engage in moral deliberation (Prinz, 2007: 71, 246; Haidt, 2001: 821, 826; 2004: 283). What is not clearly

stated by either Haidt or Prinz is to what degree this hardware is responsible in generating our moral understandings. This is the essential question which I wish to explore in this section. My position is that ethics and morality have taken their essential form from our biology, and other factors such as socialization, debate, and private reflection affect the finer details and precise articulations of our morality. To provide an illustrative analogy: biology is not the lump of clay from which the bowl of morality is formed, rather the bowl exists as a matter of biology, but its exact shape, size, and ornamentations are formed through socialization.

This conception of the moral hardware is not a new one. Charles Darwin alluded to a biologically-based source of morality, what he called the “moral sense” (2009: 50, 68, 71-72). It is a position also championed by the neuro-ethicist Patricia Churchland (2011: 7-8, 46), and by David Hume (2007: 161).

Churchland argues that the mechanisms which drive our moral world are the same mechanisms which drive our own causal world only expanded to incorporate others (2011: 7-8). In other words the same neurobiological hardware that causes me to avoid electrical shocks and acquire caramel sundaes is used in our moral deliberations. This is of course an oversimplification. Consider this specific example which Churchland uses: the insula is a small region of the brain tucked under the frontal lobe and is responsible for processing pain signals received from around the body (2011: 37). However when patients exhibit frontotemporal dementia, reducing the functionality of the insula, the patients not only demonstrate reduced sensitivity to pain, but also a loss of empathic responses (*ibid.*) The insula, a brain structure geared towards concern-for-me is also intimately involved in the concern-for-you. This multiplicity of functions for individual brain structures is what one might expect from an evolutionist’s paradigm, as Churchland points out: “Biological evolution does not achieve adaptations by designing a whole new mechanism from scratch, but modifies what is already in place... social emotions, values, and behaviour are

not the result of a wholly new engineering plan, but rather, an adaptation of existing arrangements that are intimately linked with the self-preserving circuitry” (2011: 46).

Psychopaths are also a prime example of when neurological hardware failure leads to behavioural moral failure. Psychopaths do not experience remorse, guilt, or shame – which as discussed in section 2.2 are important moral emotions – and exhibit brain structures markedly different from control populations (Churchland, 2011: 40-41). Specifically the paralimbic regions of the brain, subcortical structures related to emotional and social responses, are both smaller and function at lower levels of activity to control groups. The high levels of heritability of psychopathic tendencies among twins and families (about 70%) also provides evidence that essential moral structures are carried in the DNA and are essentially biologically sourced (*ibid.*) The two biological systems which Churchland identifies as the most important are the oxytocin and arginine vasopressin (2011: 50-53, 63). These two hormones spike during both maternal care – nursing young for example – and during acts of altruism, conversely, when the neurotransmitters for these two hormones are blocked, the test subject shows a marked disinterest in the suffering of others (*ibid.*)

Studies in infant and toddler behaviour supplement these statements, since new born children have not yet acquired significant socialization or conditioning. For example, infant humans will orientate themselves towards the distress of others, regularly adding their own cries (Preston & de Waal, 2002: 1). After the first year, children start to perform helping behaviour towards those who are distressed, even if they are themselves distressed (*ibid;* de Waal, 1996: 45). Further, Prinz cites several studies which show that two year old children are be able to differentiate between social rules and moral rules as the first is dependent on convention and the latter on harms inflicted on others (2007: 35-36). This further reinforces the claim that moral understandings are grounded in instinct.

Due to the central role which socialization plays in the formation of our moral beings for both Haidt and Prinz, it is reasonable to assume that they would disagree with my and Churchland's position. Instead they would perhaps state that the biological hardware provides room in which a moral character may be created and nurtured, and is not automatically inherited¹⁵. Ayala (1987: 235) holds such a position, and so rather than wrestle with ghosts, I will direct the debate over this question towards his work.

Ayala's thesis is that the moral domain is exclusive to human beings, based on the two arguments that the biological structures needed for morality to form were evolved through millions of years of evolution resulting in a neurological base, and through human socialization and enculturation. I aim to show that according to Ayala's own requirements for ethical thinking, we will find ethical behaviours and thus presumably ethical thinking in animals. From that, I may further substantiate my earlier claim that biology forms the essential moral processes, demonstrating a reduced role of social factors in the formation of basic ethical convictions.

Ayala describes three necessary and sufficient conditions for ethical deliberation to take place (1987: 237-239). The most fundamental of which is to anticipate the consequences of one's own actions. Ayala provides the example of a person pulling the trigger of a firearm: if that person has no knowledge of a gun is, or what pulling the trigger would result in, they cannot be held morally accountable for the aftermath. In the biological-sociological continuum, this requirement operates at both ends. To be able to anticipate the result of our actions requires a neurological structure capable of retaining

¹⁵ I have so far covered two of four possible positions: Ayala's (and assumedly Haidt and Prinz's) that biology forms the scaffold within which morality forms, and my position that biology provides the essential moral objects. The remaining two positions are that morality exists purely as a product of culture, completely free of biological constraints; and that morality is purely a product of biology, immune to social forces. However, the former cannot be held by a naturalist as all mental processes must be accounted for by the brain, and is ultimately bounded by biology. The latter position cannot account for moral plurality in genetically similar populations – a liberal child to conservative parents, for example – nor can it account for changes in moral beliefs such as a meat-eater being convinced to become vegan.

and utilizing the relevant information, as well as past exposure to the functioning of firearms.

My disagreement with Ayala's first condition arises where he states that this ability is exclusively found in the human domain, as anticipating future consequences is thinking in terms of means and ends. This condition, according to Ayala, was selected for through the advantages granted by tool working: "The ability to anticipate the future, essentially for ethical behaviour, is therefore closely associated with the development of the ability to construct tools" (1987: 238). However, tool use has been found in the animal kingdom, such as with chimpanzees (Tomasello et al., 1987: 175, 179-180). Furthermore, considering only immediate concrete tools as embodying means-and-ends thinking is far too limiting in terms of evolutionary pressures. For example, when hunting as a pack, wolves communicate with one another in order to maximize their chances of catching their prey (Feddersen-Petersen, 2000: 395). The wolves use each other as means to a common end. In essence, individuals within a society may willingly, or unwillingly, be used in the exact same manner as a concrete tool. Thus, means-to-end thinking is not necessarily the sole jurisdiction of human beings. There is no reason why any consciousness capable of communication with others is not also capable of means-to-an-end thinking – or in other words, anticipating the results of its own actions.

The second necessary condition for ethical behaviour according to Ayala is the ability to make value judgments, to "perceive certain objects or deeds as more desirable than others" (1987: 238-239). And this, according to Ayala, is only possible through the ability for abstraction, of seeing actions or objects as members of general classes or categories. Again, Ayala states this ability exclusively belongs to the human domain.

Despite Ayala's insistence that this ability and behaviour is not found in animals, there is an abundance of evidence in animal behavioural studies that show that at least certain species and groups of animals can engage in what appears to be conceptual

thinking, and perceive “certain objects or deeds as more desirable”. Chittka & Jensen (2011: R116) refer to studies that demonstrate pigeons can differentiate between photographs with people in them and unpopulated photographs. This suggests that a pigeon has the capacity to have an abstract category HUMAN, and recognize members of that class, regardless of the variability among individual members. Chimpanzees have also been shown able to sort a mix of tools and food into separate piles, even when individual items within each category had no obvious physical similarities to one another (*ibid*), again suggesting the ability to form abstract categories. Whether or not any animal is capable of abstract or categorical thinking is nonetheless treated as a contentious topic, the reasons for which de Waal might consider a form of prejudice and terms “anthropodenialism” (1997: 50). Anthropodenialism will form an important aspect to this thesis, and will be more thoroughly discussed in Chapter 3.

What is far less contentious is that animals are capable of valuing certain objects and deeds over others. The valuing of objects is most simply demonstrated by watching dogs at dinner time as they beg for food from the table even if there is food in their bowls. Further, Darwin's work on the role of sexual selection in the evolution of animals is based upon the idea that certain individuals can have desirable or valuable features such as plumage or food-gathering (1872: 46-47). While this doesn't translate to moral or ethical capacities in itself, it does satisfy the second requirement Ayala demands for ethical behaviour.

The third and final requirement for ethical behaviour, according to Ayala, is free will (1987: 239). However, on my reading of Ayala's argument I found there is a degree of conceptual confusion between free will and freedom in his work. Ayala states that “pulling the trigger can be a moral action only if I have the option not to pull it” (*ibid.*) The availability of options is a question of freedom – a person chained to a wall has no freedom but may still have the metaphysical spark of free will that allows that person to transcend

from the web of causality that a determinist would insist we are entrapped. My interpretation is supported by Ayala's statement that "free will is dependent on the existence of a well developed intelligence" as this enables us to "explore alternative courses of action" (*ibid*), i.e. intelligence expands our list of options and thus freedom, but does not necessarily grant free will.

For this third point Ayala does not insist that the capacity for free will is the exclusive domain of the human intellect, but it does offer a delineating factor: certain forms of life, and certain forms of intelligence do not have the capacity to explore alternative courses of action, and thus lack the freedom of mind required for ethical behaviour. Plants do not choose to grow in such shapes, but are bent and drawn towards the light as a result of biochemical interactions and there are no other options on the table. Unicellular life such as amoeba respond to surrounding stimuli in the same manner, with pseudopodia reaching towards food, cell membranes hardening in arid conditions and so forth (Van Haastert, 2011: 1-2).

The point of this, then, is that even assuming that Ayala is correct in his requirements for ethical behaviour, there is no reason why these requirements are fulfilled exclusively by human intelligence alone. Despite this, in English for example, we can see how the idea of morality and humanness go hand in hand, as we describe the "inhumane" acts of a murderer or the "inhumanity" of the Nazis. This, I think, reflects our attempts at a linguistic level to divorce those people from the moral community. A similar pattern can be found in Japanese where "ningen" (人間) means "human", and by adding the prefix "hi" (denoting a negative) and the suffix "teki" (denoting adjective) results in "hi-ningen-teki" (非人間的) or "inhumane" (or directly translated, not-human-ness). The same pattern can be found in the Japanese word "jindou" (人道) or "the-way-of-human". By adding the same prefix and suffix we have "hi-jindou-teki" (非人道的) which means "brutal" (or not-the-way-

of-human).¹⁶ In Hlubi the phrase “awu ngomntu” is used to describe someone who shows a lack of sympathy, and means that that person is not human (literally: awu = not, ngomntu = person).¹⁷ In African philosophy, the naming of the ethical system “Ubuntu” is an Nguni Bantu term meaning literally “humanness”.

One of the keystones to this thesis is not that we ought to “extend” the concept of morality to non-human animals, promoting savage animals to “civilized” status. Rather it is to recognize that non-human animals have moral capacities in their own right, independent of our own moral capacities. To achieve a proper realization of this, we need to be aware of the tacit assumptions we make of both morality and humanness, even in our everyday speech. This tacit assumption of morality being uniquely human will be more closely examined in Chapter 3 as well.

In this section I argued for an understanding of moral capacities that is heavily dependent on the biology of the individual. However, one should not get the impression that this is identical to a “genes-only” understanding of morality; that there is gene for honesty, a gene for violence. As I stated earlier, there is a complex interplay between the biological structures and the sociological factors. This section will briefly examine this interplay in more detail.

There are five fundamental dynamics that inform our moral-intuitive responses, and these dynamics are common to all humans, as is shown by other studies (Brown, 1991; de Waal, 1996; Fiske, 1992; Haidt & Joseph, 2004; Schwartz & Bilsky, 1990; Shweder, et al.: 1997). The first dynamic is that of the avoidance of harm, and the promotion of care: the “harm/care dynamic”. The second dynamic is promotion of fairly sharing resources and reciprocal treatment: the “fairness/reciprocity dynamic”. Haidt labels these two dynamics as “liberal” dynamics. The third dynamic involves the obedience to authority, and the

¹⁶ Dr. Masako Osada of the Centre for Japanese Studies, personal communication.

¹⁷ Sifanele Xwazi, personal communication.

respect of social hierarchies: the “authority/respect dynamic”. The fourth dynamic is concerned with group identity and loyalty: “in-group/out-group loyalty dynamic”. The fifth and final dynamic is that of spiritual purity and the sanctity of certain spaces and institutions: the “purity/sanctity dynamic”. These last three are identified as “conservative” dynamics (Haidt, 2001: 826; Haidt, 2004: 288-289; Haidt & Bjorklund, 2008: 202-203).

A strict “genes-only” reading of the above dynamics might imply that we should see a great degree of moral homogeneity within population groups but not necessarily between cultural groups. This is to say that if these dynamics arose from and are exclusively sensitive to our genetic data, then children would have near-identical moral values to their parents, regardless of upbringing. The most powerful counterexample to this would be the hippy movement of the 1960s, where long standing social traditions were overturned in a single generation. The cultural schism of the sixties was a moral schism where conservative dynamics gave way to liberal dynamics. Thus the variety of moral codes found among cultures seem to arise through learning, socialization with peers and adults, and the media.

A gene-first reading simply does not hold up under scrutiny. This is not to say that genes play no role in forming our moral selves, but rather, as De Waal (2006: 38-39) argues, our genes are the clay of our morals, and the specific expressions formed depends on the influences of our environment and society. Haidt speculates that these influences may increase or diminish sensitivity in some of these five dynamics, in a similar fashion in which babies lose sensitivity to phonemes that are not used in their home language (Haidt, 2001: 826-828; Haidt & Bjorklund, 2008: 210-211). If this analogy is sufficiently accurate, then we ought to see a direct correlation between levels of exposure to cosmopolitan lifestyles with a belief that morality is relative. In any case, the formulation of a coherent moral system, like a coherent language, requires a large amount of input from social sources.

The most important of these sources, according to Churchland, appears to be the early maternal bond (2011: 52-53). As mentioned above, arginine vasopressin and oxytocin are two hormones which are intimately tied with both moral behaviour and maternal actions. Experiments on voles showed that pups that were hand-raised or were from otherwise neglectful mothers showed a marked decrease in basal levels of oxytocin and arginine vasopressin, and behaved in a manner that was disinterested and uncaring towards other voles in the experimental colony compared to the control group (*ibid.*)

In summary, moral behaviour is totally dependent on biological systems for their existence, but are not necessarily totally dependent on them for their exact nature and expression. Sociological factors, primarily that of the early child/guardian relationship, can still have powerful effects on one's moral values and intuitions.

2.4 Conclusion

We are now armed with a thorough understanding of what constitutes the moral domain: it is emotional, and yet also rational. It is a product of biology, and yet also is subject to culture and learning. It encompasses concerns including the care of others, identity and loyalty to the group, and the evanescent "purity" of the self.

The understanding of the moral domain as being so broad in scope is supported by numerous fMRI investigations into brain activity during moral deliberation (Feldmanhall et al., 2013: 1), especially compared to non-moral dilemmas (*ibid.*: 5). Yet there exists a single thread that weaves through this entire definition. It is that moral concerns are inextricably concerns for others. Moral thinking is social thinking, or perhaps more accurately stated as a special breed of social thinking. It is social thinking that is framed by issues of harm and care to other individuals, as well as social cohesion within the group.

Contra Gert, who I referred to for a definition of morality in the introduction to this chapter, morality is not about the conduct of the individual. It is the conduct of the

individual *within a social milieu*. A statement echoed by Churchland and De Waal (2011: 8, 12; 2006: 161, 163; 1996: 30).

The details of what morality is – a fusion of emotion and rationality, inborn but shaped by the environment– are all important details that will come in handy when constructing the moral community and defining who might belong in Chapter 4.

In the next chapter we now expand the discussion to include other species. We will deal with the problem of anthropomorphism in animal studies – of attributing human-like qualities where they don't belong. We will discuss various forms of altruism as it is found in several species. We will also discuss the issues of inferring the existence of moral minds from apparent moral behaviour. These discussions will be crucial if the moral community is to have any hope of including minds that are not human.

3. Morality beyond Humanity

3.1 Introduction

In Chapter 2 we saw how moral thinking is social thinking; it is ultimately the concern for others as individuals and as a cohesive group. We also saw research discussed by Churchland and Greene et al. which suggested an overlap in the biological bases for moral capacity among species, mammals in particular (Churchland, 2011: 63). The question of this chapter is as to whether or not this understanding of morality can be found in intelligences other than human beings. The answer to this question will have important implications on the shape of the moral community. Additionally, the answer serves as a reply to accusations that any discussions of moral communities will place *human* morality at the top of the communal hierarchy. While it is true that in order to recognize morality in others there must be some conceptual overlap – that is, some features of “human morality” must exist in “wombat morality” for either species to recognize the other’s morality – and thus it is from the overlaps of moral capacities that the moral community is created, but this does not necessitate a hierarchy of moralities.

In recent years it has been increasingly accepted that animals, especially the higher-order primates, are capable of altruism, moral emotions, and genuine concern for others. This has not always been the case; “anthropodenialism” had previously been the reigning paradigm in the scientific community (de Waal, 1997: 50). I will discuss this issue first as it is the most immediate challenge to recognizing morality beyond humanity. Next, I will look at the issue of altruism in the animal kingdom, at its causes – proximal and ultimate – and argue that not all altruism is equal and that this needs to be taken into account when forming our ideas of morality beyond humanity. Finally I will consider some objections that may be made against using behavioural studies to infer the presence of morality.

3.2 Anthropomorphism and anthropodenialism

If we are to ever apply the findings of this thesis to society, we must perform a task which appears simple at first, yet has been shown to be fraught with difficulties in both popular and academic spheres. We need to be able to recognize human-like qualities¹⁸ in non-humans, while restraining ourselves from projecting human-like qualities where there are none. The latter is a case of anthropomorphism – specifically what I will later define as “emotional anthropomorphism”, while to deny the possibility of the former would be called anthropodenialism. I will deal with these two issues before moving on to case studies of morality beyond humanity.

The primatologist Frans de Waal recounts the reprimands he received from his supervisors when he was studying chimpanzee behaviour:

More than once I was asked whether the term "reconciliation" was not overly anthropomorphic. Whereas terms related to aggression, violence, and competition never posed the slightest problem, I was supposed to switch to dehumanized language as soon as the affectionate aftermath of a fight was the issue. A reconciliation sealed with a kiss became a “postconflict interaction involving mouth-to-mouth contact.”

(1996: 18-19)

This is the crux of the anthropodenialism. It is the belief that non-humans cannot have traits that are commonly believed to be uniquely human, specifically goods traits as in the above extract. In Chapter 2 I discussed how the ideas of morality and humanness are linguistically and conceptually linked. What de Waal recounts can be seen as another manifestation of that link.

Anthropodenialism is used in defence against charges of anthropomorphism.

Anthropomorphism is the projection of what is generally regarded to be human features

¹⁸ Or at least, qualities nominally claimed to be unique to human beings, especially positive qualities. See the discussion in section 2.3 regarding a linguistic and conceptual overlap between “moral” and “humanness”.

onto non-human subjects, such as in the example of the PETA advertisement discussed in the introduction to this thesis, or even onto inanimate objects. There are certainly cases of this, and sometimes, anthropomorphism can derail or impede intended efforts or even lead to false claims.

Let us take a look at the case of the rhino maggot. In 1995 an endangered white rhino was brought to the United Kingdom from South Africa, and it was soon found to host a juvenile rhino maggot, an endangered parasite which lives exclusively on the equally endangered white rhino (Kaplan, 2004: 40). The maggot was removed and left to starve to death, leading the entomologist Martin Hall to dryly remark: “‘Save the rhino maggot’ doesn’t have much appeal as a conservation slogan” (*ibid.*) It seems that conservation efforts are motivated so much through anthropomorphism that we cannot recognize legitimate conservation efforts without the emotional effect provided by anthropomorphism.

One can of course counter that animal rights organisations do not, in principle, favour anthropomorphised and neotenized species.¹⁹ Rather, the prevalence of rabbits, puppies and the like on their various media platforms is merely a political decision, not a philosophical one, a “tug at the heart strings” to engender broader support. There are two replies to this claim. First, it does not explain the apathy towards the death of a being which, in terms of conservation of biodiversity, is of equal value to the white rhino. Had the rhino been left to starve to death instead of the maggot, the outcry from all corners would have been enormous. Second, wellsprings of political influence are often derived from philosophical beliefs, whether implicit or explicit, examined or unexamined.

I argue that anthropomorphism and, as a related issue, neoteny, can be grounded in speciesist assumptions. It is to not merely identify human features in non-human

¹⁹ Neoteny is the retention of juvenile features, in laymen terms it is “cuteness” (Jones *et al.*, 1995: 728, 736-737). There is substantial literature showing that neoteny is sexually selected trait not only in humans, but other animal species as well cf. (Crow, 1993; Grammer *et al.*, 2003; Jones *et al.*, 1995). It follows that, if neoteny is related to success as it is selected for in acts of procreation, then it ought to be limited to within the same species. However, if it is extended to species beyond one’s own, then anthropomorphism of the other species is implied.

subjects, but also by virtue of this identification to hold those subjects in higher regard than those subjects who do not appear to have human features.

It is important to note that there is a distinct difference between recognizing similar traits in others, and projecting those traits where there are none. Recall the quotation from de Waal above; the kissing chimps were not kissing, their conciliatory, intimate behaviour had to be mechanized, “dehumanized” for lack of a better word. In this case, the fear of misapplying anthropic traits outweighs the use of accurate descriptors.

So caught between these two positions, describing creatures with empathetic words or being clinically detached, how ought we to proceed? I support de Waal's recommendation: that anthropomorphism be used as an explanatory tool (2006: 65). This is to say that anthropomorphism should be applied when it increases the predictability of observed behaviour. This is in contrast to the form of anthropomorphism which the anthropologist Marshall Thomas uses, cited by De Waal (64), where “wolves set out for the hunt without pitying themselves” and “virgin bitches save their virginity for their husbands”. These statements add an emotional dimension which does not provide us with explanatory powers over the animal's behaviours. Further, it may skew our understanding of the subject's psychology by projecting our own constructs onto their minds.

The use of anthropomorphism as an explanatory tool also plays into the principle of evolutionary parsimony, where two species act the same, the underlying mental processes are probably the same too – this is especially true where the species are related (62). We must keep in mind, as discussed in Chapter 2, that evolution is a tinkerer, not a designer; more complex cognitive systems came about through the expansion of simpler cognitive systems which were successful in previous generations. There is no reason why this principle should not hold true for moral systems too. If we accept Darwin's argument

(2009: 50, 68, 71-72)²⁰ that morality is an evolved trait in humans, there is no reason we cannot find it in other species, either in a “simpler” state, or as a product of convergent evolution.

By now we can identify a distinction between what might be called “emotional anthropomorphism” and “rational anthropomorphism”. Emotional anthropomorphism is extremely useful as it is one of the underpinnings of empathy. It is to see yourself in another, not in a cold, analytical manner but as an immediate emotional connection that helps establish a sense of shared identity, of community. De Waal uses the term “uncritical anthropomorphism” (2006: 62), however in addition to the negative connotation which the term “uncritical” carries, it also seems to gloss over a fundamental component to this breed of anthropomorphism, specifically the emotional character of the act.

Rational anthropomorphism by contrast, is the kind which is only called upon when it adds to our explanatory powers. In the following section, in which we look at the apparent altruistic and moral lives of various animals, we will see what it takes to be “truly” altruistic over “functionally” altruistic, what the different causes of altruism are from an evolutionary paradigm, and which of these types of altruism we might value. As we have seen earlier with statements from de Waal, observations of apparent altruistic actions in animals are often explained through much more complex terminology than that of admitting moral capacities in the animal. This is done to avoid supposed unprofessional or unscientific appeals to emotional anthropomorphism. However, if we accept rational anthropomorphism as I have discussed in this section, we will see that there are animals who live rich social and moral lives, lives which are traditionally conceived of as being exclusively human.

²⁰ Contemporary studies in evolution corroborate Darwin’s position as well. See Churchland (2011), and de Waal (2006).

3.3 The computing ape and the kamikaze honeybee

Scores of bees defend their nest against an interloper, sacrificing their lives for the good of the hive. A flock of flamingos take off as one because one of their members signals an approaching threat, endangering itself in the process. An ape carries a bucket of drinking water over to an unrelated, sickly ape without training or apparent reward. These are three cases of altruism, the act of assisting another at the expense of the self. In this section, I will examine these cases and argue that, while broadly they all belong to the category of “altruism”, not all acts of altruism are equal, and only the third case of the helpful ape is of the kind that can be said to be moral in nature.

From an evolutionary perspective, altruism can be puzzling. It is defined as occurring when “an organism reduces the number of offspring it is likely to produce itself, but boosts the number that other organisms are likely to produce” (Okasha, 2013).²¹ Given the process of natural selection, we would expect an eventual extinction of the altruistic traits in a given population in favour of more selfish traits.

One explanation for the favouring of altruistic traits is that of “kin selection”, where an individual will promote the fitness of its relatives (Wade, 1980: 665). The starkest example is drawn from the honeybee (*Apis mellifera*), where sterile worker bees work to serve the hive, and die to defend it while in a condition of “eusociality”: the highest order of social cohesion (Wilson & Holldobler, 2005: 13367). In fact, their cohesion among individuals of beehives and ant nests is so high that Richard Dawkins has suggested that it may be easier to conceive of the hives and nests as macro-organisms themselves, and the bees or ants as organs of those organisms (2006: 171-172). Kin selection theory posits that in serving the hive, the worker bees effectively serve themselves. By assisting the queen bee, with whom they share half of their genes, they ensure the continuation of their own genes for another generation. This theory produces a testable prediction: a hive which

²¹ “Fitness” or “reproductive fitness” is the term used to describe a subject’s probability of producing offspring (Okasha, 2013).

loses its queen will show a reduction in altruistic behaviour. This is the prediction which Naeger et al. set out to test.

When a hive becomes queenless, it is possible for some worker bees to develop ovaries and begin laying haploid eggs (Naeger et al., 2013: 1574).²² This means that even in a queenless colony, it is possible for the hive to perpetuate some of its genetic material with ovary-bearing workers functioning as ersatz queens. A queen and its worker have a genetic commonality of 75%, while the genetic commonality between the haploid eggs and workers is half that (Queller & Strassmann, 1998: 167). As predicted by kin-selection theory, Naeger et al. found that worker investment in colony maintenance was reduced once the colony entered a queenless state and reproduction was left to haploid-laying workers (2013: 1576).²³

The question here is: Can we call this kin-altruism *true* altruism? The answer to that question is dependent on another: What exactly constitutes the self? The everyday understanding of the term relies on an individual-oriented conception of the self. The *individual* bee reduces its chances of survival to assist another *individual* bee in surviving and reproducing. An evolutionary concept of the self are gene-oriented so, biologically, a set of genes is acting in such a way that the majority of itself will survive to the next generation. In queenless colonies, less of the gene-self will be replicated in the next generation and so altruistic interests are likewise reduced. As an explanatory model, the biological paradigm is able to account for the behaviour of the bees in queenless colonies to a greater degree than other accounts which rely on rationality or other motivations.

If one accepts the biological paradigm, the I-am-my-genes paradigm, then we can say that kin-altruism is not true-altruism as it is ultimately a self-serving act to assist in the

²² "Haploid" means to only contain one set of chromosomes as opposed to the usual two sets in diploid eggs (Naeger *et al.*, 2013: 1574). The haploid eggs of bees are the eggs of male drones (*ibid.*)

²³ There have been other studies which have shown that interactions between bees are dictated by their genetic similarity, such as Greenberg's 1979 study (cited in Rushton et al., 1984: 183) which demonstrated a very strong correlation ($r = 0.93$) between a bee's genetic similarity to a queen and its likelihood of being allowed entry to the hive by the guard bees.

continuation of your own genes. It is, of course, absurd to say that every bee in a queenless colony has consciously calculated that there has been a 50% reduction in genetic similarity with the haploid eggs, and thus a corresponding drop in altruistic actions is in order. Just as it is also absurd to say that a human parent only cares for their baby due to some calculation of genetic overlap, and they selfishly want their genetic data – which they may not even be aware of if they lacked formal education – to be continued for another generation. Herein lies the distinction between ultimate and proximate causes.

In evolutionary and sociobiological discourse, an ultimate cause is described as the agent of natural selection which explains the given behaviour. A proximate cause is the immediate physiological or environmental factor which would explain the behaviour (Francis, 1990: 401). For example, the ultimate cause of a parent caring for their child would be the preservation of genes, while the proximate cause is described in terms of the strong emotional bond a parent has with the child – explained biochemically through heightened levels of oxytocin and arginine vasopressin (Churchland, 2011: 50-53).

One might protest at this point and say then that this implies that there are no truly altruistic acts. If morality is the result of evolutionary pressures, and natural selection operates through whatever might improve gene survival, then all supposedly “moral” acts are self-serving acts. This argument engages in the reductionism which the cognitive neuroscientist Steven Pinker cautions against:

When a person's public stance is selfless but his private motives serve his interests, we can call it hypocrisy. However, when a person's public stance and private motives are both selfless but those motives came about because they once served the interests of his ancestors' genes, we have not uncovered hypocrisy; we have invoked a scientific explanation couched at a different level of analysis...the evolutionary causes of our motives can't be judged as if they are our motives.

(cited in de Waal, 1996: 238)

Does this counter my earlier claim that kin-altruism is not true altruism? It does temper the strength of the claim, but we cannot ignore the fact that most parents care exponentially more for their own children than they do about a stranger half a world away. It is also highly unintuitive to say that to serve your children dinner is equally morally laudable as to feed the homeless. One may of course point out that we are duty-bound to care for our children, while we are not duty-bound to care for the homeless. But this only answers half the dilemma: failure to perform a duty ought to at least bring about a sense of disapprobation, but then why does fulfilling an oft ignored non-duty bring about a sense of approbation?

I believe the answer to be that morality is the overgeneralization of kin-care. This is not a recent idea, and can be traced back Lecky's 1869 text *History of European Morals* where he describes what came to be known as "the expanding circle of morality" (Singer, 2006: 226; de Waal, 1996: 213). This circle entails the beginning of care with the self, and as available resources increase the scope of our care grows to incorporate family, then friends, class, and so forth (*ibid.*) Churchland also cites compelling studies in neuroendocrinology which suggest that, in mammals at least, oxytocin "has been recruited in organizing the brain to extend self-care to infants, and thence to a wider circle of caring relationships" (2011: 63).

While extremely rare, there are also some reports of altruistic acts between members of different species, such as the hippopotamus which saved a drowning zebra and wildebeest in Tanzania (van Rooyen, 2010). These reports are especially confounding for the earlier reductionist argument. If all behaviours are necessarily and perfectly aligned with gene-propagation, then altruism has no business with cross-species non-reciprocal aid (Darwin, 2009: 108; Wyatt et al., 2013: 1854, 1861). Care for others is essential for child rearing, and very useful in reciprocal alliances, but it is only the overspill of that care which is *true* altruism.

So what about the bees? Naeger et al.'s study showed how intimately tied the behaviours of the bees are to the first level of care for others. The moment there was a loss in genetic similarity of progeny, there was a corresponding loss of altruistic behaviour in the hive. Bees, and a number of similarly organized animals, cannot be said to be *truly* altruistic, to hold a *true* moral capacity.

It should be clear by now that for the purposes of the normative claim of this thesis the biological definition of altruism is insufficient. There is another term from biology which might serve: "succorance" or "succorant behaviour" which is defined as helping, caregiving, or providing relief to distressed or endangered individuals other than progeny (de Waal, 1996: 40-41). This term is more in line with the above statement that true altruism is the overflow of care. But is succorance alone a necessary and sufficient reason to welcome an individual or species into a pragmatic moral community? To answer this question, let us now consider the case of the crying flamingo.

In the introduction to this thesis I argued that a defining characteristic of a community is the interactions which occur among the concerned individuals. This implies that membership to a community is dependent on two criteria: the capacity to treat another in a manner which defines the community as such, in this case a moral community requires moral consideration, and also the capacity to incorporate the rest of the moral community within the scope of this treatment. Bees failed at the first criterion, their altruistic behaviour being driven by kin-care. But what of the flamingo, feeding at a lake's edge with its flock, who upon seeing a leopard stalking through the cattails, gives out a warning cry that launches the flock into flight? The cry would single itself out to the predator, but serves to protect the rest of the flock. Does this altruism achieve the two criteria needed to be a member of the moral community?

One might say that the flamingo is acting no different to the bees. The individual endangers itself to protect its family, and as argued earlier from an I-am-my-genes

perspective, this is not true altruism. This argument is dependent on the common idea that a flock of birds is analogous to a hive or family. However, a series of genetic studies have shown that while genetic similarity among birds can be roughly drawn along geographic lines (Foerster et al., 2006: 4555), flocks themselves do not follow family structures (Westneat & Webster, 1994: 92). Adults and emancipated juveniles intermingle among flocks without any apparent familial organisation.

The flamingo then exhibits succorant behaviour. This fulfils the first criterion listed above, but does it fulfil the second? The flamingo has the capacity to engage in actions to help others in distress. Frans de Waal asked a similar question (2006: 25). He argued that the bird was acting reflexively, a reflex that had come about by powerful evolutionary pressures for protecting its immediate community alone. While de Waal's stance is that we should not consider this reflexive behaviour as altruistic, I would argue that it is indeed altruism—a relatively simplistic form of altruism that may have formed the primitive basis of more evolved forms of morality, but altruism nonetheless. That it is a reflexive act is, however, still an issue when attempting to bring this sort of individual into the moral community. The fact that it is a reflexive act, evolutionarily selected for use only on a small, exclusive group means that the second criterion cannot be fulfilled. In other words the flamingo may hold others in a moral regard, but these others are restricted to such an extremely limited group of individuals that the flamingo cannot pragmatically belong to the entire moral community.²⁴

What is necessary to belong to a moral community then, is the ability to “get under the skin” of others, to see from their perspective, understand their needs, and enact a plan to assist them. This is not merely an empathetic emotional projection as I discussed earlier with emotional anthropomorphism. It is also a rational act of identifying concerns and developing strategies which are adaptable to varying contexts, and are not merely

²⁴ The issue of moral agents not extending moral considerations to deserving others will be discussed in Chapter 4.4.

reflexive cues. The emotional aspect is implied by succorance itself. The flamingo flew off out of fear, and gave its cry out of fear for its flock-mates. The two rational components required to “get under the skin” are those of perspective taking, and targeted helping.

Perspective taking is a rational estimation of what another being is capable of perceiving (Kaminski et al., 2008: 980). For example, Kaminski et al. found that when trained dogs are given a “fetch” command, and two balls are in the dog’s field of view, but only one is in the field of view of the owner, the dog will always fetch the ball which the person can see (985).

How does this relate to moral communities? When acting within the context of a community, members must take the perspective of many disparate others into account while engaging with them. The crying flamingo, by contrast, will only give a warning cry when there is a threat to their immediate surroundings. Had the leopard been spotted stalking another flock of birds on the far side of the lake, the flamingo would not have given a warning cry. The perspective of the flamingo is limited to itself. If we forgo the requirement of perspective taking, and incorporate the two flocks on the lake in a single community, it could not be recognized as such, since their perspectives are limited to themselves, they could not mutually assist each other.

Frans de Waal defines targeted helping as “altruistic behaviour tailored to the specific needs of the others, even in novel situations” (2006: 31-32). Perspective taking is a prerequisite to targeted helping, since you need to identify those specific needs which may differ from your own. Targeted helping is thus the behavioural expression of a combination of altruism and perspective taking, and is a crucial component needed to contribute to a moral community and thus to belong to one.

De Waal provides an illustrative example of targeted helping at work in the animal kingdom which he observed while researching chimpanzees at a zoo (*Ibid*): Old tyres are

given to the animals to play with in their pens. The pens are also hosed down from time to time to keep them clean. One afternoon after a cleaning, De Waal noticed a young chimpanzee who was attempting to pull a tyre over to the niche where she usually played. However, the tyre had caught a lot of water during the cleaning process which made the tyre unusually heavy. After a few minutes of struggling, the chimp gave up and went over to her niche. A larger adult chimpanzee had been watching this, and once the juvenile had quit, lumbered over to the tyre, and dragged it to the youth, and left her to play with the tyre. The adult had received no training to this effect, and did not appear to receive any reward.²⁵

Why did the adult chimpanzee act in such a manner? De Waal is satisfied that this was an example of targeted helping. The chimpanzee observed the juvenile's struggle, understood her intentions and goals, and assisted her in achieving these things. We might say that this particular example is not, strictly speaking, a moral action, but that is beside the point. The claim here is not that targeted helping is a moral property in itself, but is a property necessary to belong to a larger community of moral agents.

Another response, and one which is fairly common in explaining apparent altruism in animals, is that selfless actions such as these are part of a larger strategic plan (de Waal, 1996: 151; de Waal 2006: 33). We might argue that the adult chimpanzee was making early moves in a long term strategic plan to win over the juvenile for later support as the group alpha. Again, this response does not take into account that targeted helping is not a moral property in itself. It may explain the ultimate cause of targeted helping, but that does not diminish its pragmatic value in the moral community. Secondly, when used to explain purportedly altruistic acts in animals, this strategic explanation endows the animal with an incredible mental repertoire of memory, long-term planning, and manipulative guile. It turns a relatively simple act of aiding a conspecific in achieving a task for no real

²⁵ There are numerous other examples of targeted helping, mostly in chimpanzees. Cf. (Pruetz, 2011; Melis et al., 2011; Yamamoto et al., 2012)

benefit, to massively complex political manoeuvring. The net effect being that we can explain a behaviour without resorting to moral concepts in animals. We create a computing ape to avoid discussing a moral ape. While we cannot say for certain that the adult chimpanzee is not a sly politician, it is the simpler and more likely explanation to say the adult was helping the child.

It should be noted with de Waal's anecdote that the adult chimpanzee's *ability* to move the tyre alone does not constitute targeted helping. Had it been another juvenile that tried and failed to move the tyre, we can still surmise that it was in an effort assist the first juvenile. We can say then that the assisting juvenile understood the goals and desires of the first, and at least attempted to fulfil them. We should also note that the conceptual difference between perspective taking and targeted helping is twofold. Firstly, perspective taking is purely a mental exercise which is occasionally evident in behaviours as we saw with the dogs fetching the balls, while targeted helping incorporates the physical act of helping another once we have understood what their needs or goals are. Furthermore, perspective taking does not imply concern or altruism, it is merely the estimation of another's point-of-view. Targeted helping does at least suggest altruism, unless one accepts the argument of the computing ape.

In summary, I have covered three primary points in this section. First, there is an important distinction between kin-altruism and true-altruism. Kin-altruism is ultimately self-serving as we see in the kamikaze honeybee and cannot be extended beyond family members. True-altruism is the overgeneralization of kin-care to non-kin. I would also like to point out that in this discussion I have focused on the term "altruism" as opposed to "morality". The reason being that altruism is a publicly observable behaviour and subject to much research, while morality also includes private emotional states, deliberations, and the avoidance of "bad" behaviours. Given further research or the capacity to see into the

minds of the studied subjects, observations of morality would be a preferred topic of discussion.

Second, I have argued that the capacity for altruism is in itself insufficient to belong to the moral community. One needs to be able to take the perspective of others by estimating their point-of-view in order to understand their needs.

Thirdly, I argued that targeted helping is a functional expression of perspective taking which strongly implies altruistic concerns.

One may reject these three points, arguing that it is all inferred from behaviours. We do not know for certain if the bees care for the welfare of other hives as much as their own. We do not know if the crying flamingo would help another flock if it could. We do not know if the ape is really helping, or playing a political game. The animals in question cannot tell us their intentions and motivations, so can we make any estimation of them based on behaviour? This question will be addressed in the next section.

3.4 Behaviourism, Batman, and Solipsism

In the above discussion of morality in animals, we have been largely limited to behavioural studies. The obvious language barrier between species forces us to look at the actions of animals rather than their stated intentions, since they have no way of telling us what they might be!

Studying an animal through its behaviours alone comes with a major drawback, namely that we can only clearly judge its actions through the consequences of its behaviour. For example, if I come upon a nomadic family where the mother is trying to reach up into a tree to grab some low hanging fruit, but is too short to do so, she can still tell me that she wants the fruit to feed her child. On the other hand, if I came upon a troupe of chimps in a similar situation, I have no idea if the chimpanzee reaching for the fruit wants it for its own dinner, or feed another chimp, or whether it simply wants to play with a

colourful, ball-shaped object. It is only once the chimp has succeeded in its attempts to seize the fruit that I can come to a reasonable conclusion on the nature of the chimp's actions, and hence whether they can be said to be a moral act or not.

Consequentialists such as utilitarians²⁶ may scoff and ask why that is an issue, since it is only the net result that matters and not the intent. However, to read morality into actions and their consequences alone is not completely satisfactory, as can be illustrated in the following "Batman" thought experiment:

- 1) As a vigilante, Batman has saved B people in Gotham.
- 2) Batman became a vigilante because of the murder of his parents.
- 3) Thus, the murderer of Batman's parents (Mr Frost) saved B – 2 people in Gotham.

By a purely consequentialist reasoning then, Mr Frost is always almost as much of a hero as Batman. To argue "but he killed two people!" or that "Frost could not foresee such an outcome!" is to abandon the consequentialist paradigm, as so long as Batman saves more than two people, the net consequence of Mr Frost's actions is a positive.

A steadfast consequentialist may accept this argument, and state that Mr Frost is, indeed, almost as much a hero as Batman. But here we come upon a paradox: Mr Frost is in a moral superposition of villain, civilian, and hero. His quantum moral character only collapses into a definite state once we solve for B. Immediately after the murders, Mr Frost is a villain. However, sometime later he becomes morally neutral as, through no further actions or decisions on Mr Frost's part, Bruce Wayne returns to Gotham as Batman and saves two lives. Again, a short time later Mr Frost's moral state increases to hero as Batman saves more lives. To accept a consequentialist line of reasoning then, is to also accept that the moral status of a person may vary radically as time passes even if that

²⁶ We may wish to also recall in Chapter 2, where I argued that pure utilitarianism may not even be feasible. Consider the standard trolley problem I mentioned where the choice is between five strangers, and your own child. It may be more useful to think of utilitarianism as the default position we take, even in non-moral issues, and it carries a certain "inertia" that must be overcome by moral emotions.

person performs absolutely no further actions. While we may be able to say “this action was ultimately good” or “that action was ultimately bad”, this approach gives us no insight into the moral character which is the target of our investigation. From this approach it makes as much sense to say that Mr Frost has a morally good character based on the consequences of his actions, as it is to say a tree has a morally bad character for falling over and crushing a home.

This paradox is a result of what I will call “consequential absolutism”. I use this term to differentiate between it and a softer form of consequentialism. This softer form allows the second objection to the Frost problem above: Mr Frost is only morally responsible for the consequences that he could reasonably foresee.²⁷ Of course now we are incorporating some measure of intent, so why do we call this a form of consequentialism? Because, at least when applied to animals and other consciousnesses that have great hurdles in communication, our moral estimations are intimately tied to the success of the concerned actions – recall our nomads and chimps earlier.

Of course, one might argue in the other direction, that consequences have no bearing on morality, but that intent does, which can only be truly revealed through language. This would mean that behavioural studies have lost a lot of their impact, since we are only viewing behaviours and from them inferring intent. This may fall under the banner of anthropodenialism which was discussed at the beginning of this chapter. I will not re-tread the same ground here, but I will make this rebuttal: to accept the above argument is to imply a form of solipsism.

My rebuttal is based on the question: what is the nature of speech? Why do we hold human language in such high regard over the barks of a dog, or the information-rich scent of a gazelle’s glands? We may quite confidently state that human languages are far more

²⁷ This line of thinking is endorsed by the judiciary, referred to as “the test of negligence” (Green, 1928: 1029).

complex, and capable of more creativity and adaptability than the instinctive languages of animals – although one could go further and say we simply don't know how complex, creative, or adaptive a dog's bark is, since we are not dogs. However I do not believe I need to go so far to defend my rebuttal. At the root of it, human speech is simply another observed behaviour.

One's understanding of a state of affairs is never completely divorced from inference made from behaviour. Consider how often a speaker makes some error in their speech, and we almost unconsciously "patch it up" in our own minds for them so as to understand what is being said (Cummings, 2005: 98-99). This is especially true if one considers the uniqueness of each communicating mind. Nearly every word carries its own connotations, its own memories, and its own impact, such that an utterance made by one and heard by another will never result in identical hermeneutical horizons. For there to be meaningful communication, there must always be some measure of inference, from tone, posture, facial expressions, or even eye contact. Inference in language is largely possible thanks to observation of behaviours.

The core of this rebuttal, however, is that speech is merely another observed behaviour. Faster, more detail-rich than reaching for a fruit, more adaptable than a hug, but a behaviour nonetheless, from which we read the intentions of others. Sometimes, as in other "bodily" behaviours, we project our own meaning over the other's. But to claim that there is an unbridgeable divide between human and non-human minds on the basis that the latter lack human-style speech is to ultimately admit that you cannot know that there are other intention-capable minds beyond your own. While this may be an intellectually interesting position to debate, for all pragmatic purposes it is useless navel-gazing.

3.5 Conclusion

In this chapter I had three aims. First, to provide a compelling case that, contra the dominant view in animal studies, non-human animals are capable of moral capacities. This

was argued through the discussion with the concept of anthropodenialism, the reluctance to see attributes generally considered to be “owned” by humans as being present in animals. The reluctance to do so, and the strategies employed to avoid anthropomorphism, reduce our own explanatory and predictive powers.

My second aim was to broadly delineate the three properties necessary for a being to belong to a moral community. These properties are true-altruism, perspective taking, and targeted helping, with the second being a prerequisite for the third. Altruism demonstrates that the subject has the capacity to engage in a moral community – it holds “currency” for trade if we return to my earlier analogy of an economic community. Perspective taking gives the subject potential to engage with other members of the moral community, and targeted helping demonstrates the fulfilment of that potential. Without these two properties a subject cannot truly be a member of the moral community as its moral actions will be too limited.

“Altruism” was the preferred concept of discussion here as it is by definition a publicly observable act, open to study and scrutiny. “Morality”, by contrast, includes private deliberations, emotions, and the avoidance of immoral actions which are harder to study with non-communicating subjects. This discussion was limited to altruism for practical reasons, but ideally ought to be about the moral capacities of others, and not simply the behaviour of altruism.

My third aim in this chapter was to argue that being reliant on behavioural studies is, while not ideal, both necessary and rational. To deny this would force one to concede a form of solipsism, as not only is most communication between humans based on body language, but spoken languages are themselves nuanced behaviours reliant on interpretation.

In the next chapter, I will be presenting an exposition of my metric for moral consideration. However, before moving on I would like a moment to restate that, while the entire chapter was reserved for the discussion of animal morality, this metric is not limited to the animal rights debate. It is a metric which ought to work with any conscious being, and the concepts and principles laid down in this chapter are general. A machine intelligence, for example, would still need to exhibit altruism, perspective taking, and targeted helping if it is to be considered a member of the moral community. A brief treatment of the metric beyond animals will be given in Chapter 5.

4. The Metric and Moral Communities

4.1 Introduction

At the heart of this thesis is a pragmatic claim, which is to say that it is a prescription for selecting the most useful course of action, moral or otherwise. This is not the same as a utilitarian claim, which advocates the maximum benefit, happiness, or minimum suffering depending on which school of utilitarian thinking one subscribes to (Driver, 2009). A pragmatic claim is one which ideally aims to be maximally useful, to maximize our powers over our own lives. To a certain extent then, a pragmatic claim is an economic claim. Benefits may be derived from this usefulness, and happiness from benefits. But this is first and foremost a claim about usefulness. In this chapter I will, however, also offer a secondary, utilitarian claim which is dependent on the first pragmatic claim.

I consider the utilitarian claim secondary as it is dependent on certain facts of history and economics which are in flux and so only hold true at a particular point in time. The first claim, the pragmatic claim, will always hold true. That which might change with time is the necessity of having a high degree of usefulness. In times of extreme excess, pragmatic claims which stress the importance of practical consequences— or of power to speak very broadly – might seem barbaric or antiquated. Should we ever achieve a world free of want, I would not be surprised if all discussions such as this are relegated to the history books; not because they are no longer true, but because they are no longer relevant. However, so long as we need to make choices about where and how finite resources are allocated, pragmatic arguments will continue to be relevant and they will continue to be important.

This chapter will begin with a discussion of two pragmatic claims, followed by the secondary utilitarian claim. The discussion will also suggest a tertiary claim on the metaphysics of morals, which I will afford a short discussion. I will then close the chapter with a discussion on some possible objections to my thesis.

4.2 A claim, three arguments, and an observation walk into a bar...

Before I begin constructing my argument, I would like to reiterate the two most pertinent points argued for in this thesis thus far. First, morality is not entirely either emotional or rational. If you exhibit an inability for complex thinking, or have a deficient IQ or some other cognitive disorder, then that does not necessarily mean that you are immoral, non-moral, or less-moral than another who does not have those traits.²⁸ An “evil genius” may be a cliché, but it is not an oxymoron. Second, to be a moral being is not in itself sufficient to belong to a moral community, you must be able to incorporate and assist the rest of the moral community in your moral considerations. In other words, in order for the moral community to exist there must be “trade”, to once again return to my early analogy of an economic community.

At last we come to the primary pragmatic argument:

- 1) Beings which are capable of morally considering others ought to be included in a moral community.
- 2) Membership to this moral community carries with it certain rights and privileges, principally the second maxim of the Kantian imperative that they should never be treated merely as means, but only as ends in themselves.²⁹
- 3) Additionally, these rights and privileges of members take precedence over those of non-members.

Let us pause for a moment and consider the intentional usage of the term “beings”, although it should not come as a surprise after reading the first and third chapters of this thesis. Any person, animal, alien, or intelligent machine who has demonstrated the capacity for moral concern may be brought into the moral community; as per the

²⁸ Having a lack of emotion is another matter, and will be discussed in section 4.4.

²⁹ I use the Kantian imperative somewhat reluctantly, and I acknowledge that “certain rights and privileges” is terribly vague. The discussion of these issues is in section 4.3.

discussion in Chapter 3, we should be able to acknowledge that morality is not an exclusively human trait.

I have also refrained from using the term “species” (we might speak of alien species, species of machine intelligence etc.) for two reasons. First, the aggregate capacities of a species does not exist as an absolute rule for every individual. Humans are generally speaking morally capable beings. Yet there are individual exceptions, such as psychopaths or the arbitrarily cruel, who would be denied access to the moral community. One should not be given a free pass simply because the people to your left and right have the necessary prerequisites. That being said, it is eminently practical to operate on a species-by-species basis until individuals distinguish themselves otherwise.

The second reason to not use the term “species” stems from the initial intent of this thesis to present a non-speciest thesis. This has proven to be a herculean task as all research in biology, ethology, and primatology is steeped in the language and thinking of the speciest distinction. I must, however, concede the usefulness of this distinction. We cannot expect some lone honeybee to overnight sprout an anterior dorsolateral prefrontal cortex, anterior cingulate cortex, and all the other neurological and morphological panoply needed for it become recognizably moral. So, while this argument is not strictly anti-speciest, it is not wholly speciest either. Lines are not drawn arbitrarily by the taxonomy of species, but rather by the moral capacities of the given individuals and species.

But of course we can ask: *why* ought we to do this? Why construct this moral community at all? I will stress again that this is not an argument for moral desert. You are not a member of the moral community because you are good and good ought to be rewarded. We should not endorse the moral community out of the expectation of a sense of approbation or moral pride. Rather we ought to do this on the basis of the economically-minded principle of reciprocity. To put it clinically:

- 4) We should do this to maximize our return on investment.

The “we” and “our” here does not point, as it usually would, to humanity. Rather it points to the moral community itself³⁰, which may exist even if there are no human members.

For a simple illustrative example, let’s assume that all dogs are beings capable of moral concern, as well as targeted helping, while all cows are not. (This need not reflect reality for the purpose of the following example.) We are then faced with three options: treat the dogs and cows under the same category of “animal”, as outsiders to the community, which is more or less the current norm, although of course there are differences across cultures; dogs are eaten in traditional communities in China and Korea while cows are revered in India and so forth. Or we could treat everyone as having equal rights to ourselves, as some moral theorists argue we should; think of, for example, Masolo’s “boundless society”³¹ (2002: 566-567), Goodpaster’s “to be alive is to be equal” (2013: 384-385), and Singer’s “to suffer is to be equally considerable” (2002: 17). Or, we could distinguish between dogs and cows on the basis of their aforementioned moral differences, treating dogs as a part of our moral community but not cows.

In the first case, dogs would have the capacity to contribute to the moral community but might not be able to realize that capacity. They could, like de Waal’s helpful chimpanzee, identify needs and concerns of others, and help to answer them. They could, that is, if dogs are given the rights and freedoms necessary to be fully capable of exercising their abilities. The dog cannot come to the aid of a mugging victim if, as a result of its non-member status, it is kept in a cage. This option wastes the positive social potentiality which dogs possess in our example.

³⁰ This phrasing suggests that I assume that both I, the writer, and you, the reader, are qualified members. Neither case is necessarily true, but I will keep this phrasing for the ease of linguistic convention.

³¹ Although it is worth noting that Masolo was more interested in a *human* boundless society, there is no reason why his arguments could not be applied to non-humans.

In the second option, the rights of both dogs and cows are respected. The dog is now free to assist the mugging victim, and the cow is free to go about minding its own business. The moral community profits, some might argue, while non-members are unaffected. However, granting equal consideration of rights to others is not a cost-free act. The bodies of cows are extremely useful, and are used in much more than the food industry. They are used in the creation of glass, adhesives, gelatines, plastics, fertilizer, nitro-glycerine, phosphorus, a wide variety of medicines including antibiotics, insulin, multivitamins and many other products (Centre for Disease Control, 2013; European Medical Agency, 2013; Klinkenborg & Modica, 2001; Robertson, 2004). To no longer use their bodies as a means to an end would mean the surrender of a plethora of useful resources, all for no return on investment.

Furthermore, once they are ends in themselves, the millions of cattle on Earth today would need their own habitats of suitable size and nature. After all, we could not simply dump them in the Sahara and let them perish. It would mean that enormous tracts of prime agricultural land would need to be relinquished as cattle reservations. Regardless of whether this option is instated piecemeal or all at once, there will be substantial short and long-term costs attached, without any pragmatic return.

Were the situation reversed, and cows were capable of moral consideration and targeted helping, there would be an increasing, mutual reward as they are incorporated into the moral community which would offset the costs attributed to incorporating a species into the moral community. For our example at least, this is not the case. It stands then that until resources become sufficiently plentiful, equal rights and consideration for all beings will not be the most pragmatic course of action.

The third case is the option which I advocate for in this thesis: dogs are brought into the moral community while cows remain outsiders. Under these conditions, dogs would be able to maximally contribute towards the moral community, enriching it for others as well

as themselves. Additionally, the usefulness of cows is not lost without recompense. This is not to say that incorporating the dogs in the moral community is without cost. Various legal, social, and economic restructuring would need to take place, similar to the situation with the cows discussed above. However, the costs of restructuring would be limited, once-off costs, while the reward of mutual assistance – of moral trade – would be ongoing into perpetuity.

A reading of this argument may give the impression that members of the moral community are “doers”, since membership is predicated on one’s ability to act in certain ways, and outsiders are “resources” for members to exploit as they see fit. However, being an outsider to the community does not relegate one to the status of object. While it is true that this argument advocates that outsiders *may* be used as a means to an end, their usage must still be justified. After all, a member causing unjustified suffering would indicate a potential lack of moral capacity which would flag them for possible expulsion from the moral community.

There is also a second pragmatic argument in support of my conception of the moral community. It is that beings capable of moral concern for others are capable of self-policing their actions. Beings not capable of this pose an unmediated threat to others, and would require the expenditure of resources to monitor or otherwise restrain them to protect others. This is not to say that members of the moral community are incapable of harming others – fits of rage or greed or desperation can push anyone to committing a crime. However, a well-adjusted individual does not require the same wary eye as a psychopath warrants. The moral community would provide the grounds on which we make such a decision as to who ought to be watched and who ought to be left to self-police until proven otherwise.

In short, equality under law assumes the law is being applied to those who are equal in nature. The right to freedom exists as it is at first assumed that all have the

rational, emotional, and moral capacity to responsibly exercise this right without harming others. Again, this does not relegate outsiders to the moral community to living in gibbets – recall the above discussion regarding unjustified harm – but it does mean that freedom is not an inalienable right for them. This is not necessarily true for all rights of course. A proponent of the right to life as an ultimate right may argue that there is no responsibility tied to that right, and it may never be violated. A closer examination of rights in the moral community is provided in the next section.

The secondary utilitarian argument that I want to make is that indiscriminate altruism with finite resources is sub-optimal, and favouring altruism towards members of the moral communities optimizes the utility of available resources.

For a stark example of what I mean, imagine that in the near future the ambitions of politicians have finally pushed the world over the edge, and left it a charred, lifeless wasteland. Fortunately, NASA had constructed a massive, completely self-sufficient space station in orbit that is crewed by twelve astronauts. Unfortunately, construction was not yet completed when the bombs fell, and even after all the ingenuity the crew could muster, the space station can barely produce enough food to sustain eight people. To equally share their food would only grant the astronauts a very slow death. Four must die and the astronauts must discriminate amongst themselves to decide who will live. They might make the decision based on sex, arguing that women are far more valuable in reproductive terms, or on their respective skill sets, or some other practical criteria. The point is, when resources are lacking neither indiscriminate altruism nor random selection should appeal to the utilitarian.

But the above scenario does not fully capture the argument for the moral community. Members are members because of their capacity for concern and their ability to in some way contribute to the community. Resources spent sustaining and aiding a member are an investment towards that member returning in kind with their own skills and

resources. An outsider will have no compunction in consuming the resources given, and then ignoring the needs of others. In this case, unless the members have some way to compel an outsider to assist, the outsider becomes a veritable black hole of valuable resources. In times of scarcity members of the moral community have justification to not afford the needs and rights of outsiders equal consideration.

Finally we come to the third claim. The claim is that when we judge someone to be evil, we do so not only on the single dimension of “goodness”, but also on their ability to include deserving others in their moral concern. This is not so much an argument towards the creation of a moral community so much as it is an observation that we already have an intuitive sense of its essential structure. This observation is based on our answer to the question, “When do we call someone evil?”

Like the discussion on the moral domain in the second chapter, this is not so much a question of specific acts, as we will find little in common across cultures. A Christian extremist may say that the pornography industry is evil, while a libertarian may say that book-burning is evil, and both will think the other is wrong. Rather, this is a question of the meta-rules which inform our deliberations before coming to the conclusion of a person’s goodness or evilness.

Let us take the case of the happy slave. In this I do not mean some absurd Uncle Tom fantasy. The widespread brutality and outright injustice of slavery is not something to be made light of. But we can imagine a slave owner, the son of a Roman senator perhaps, who is soft-hearted and has never taken a liking to beating or otherwise mistreating the slaves of his household. The day eventually comes when his father passes away and the slaves, seen as possessions, are bequeathed to the kindly son in the senator’s will. The son, very much a product of his family and culture, does not free his slaves, and continues to see them as objects which he owns.

His slaves have no freedom, they must obey his every order, subject to Roman law they would be crucified should they try to run away, they may still be traded at the forum, indeed their very bodies are not their own. And yet under the care of the son there is no unjustified punishment, and are generally well kept so long as they play the role of the good slave.

From our modern, liberal perspective we might call the son evil. We say so not because the son has no moral capacities, or lacks empathy, but because he has not extended his community of equals to those who are deserving of being included. The judgement of “evilness” then, is not on the single dimension of moral capacities but also on the dimension of moral inclusion. Since the days of Imperial Rome the rules which dictate whether one is morally included have changed, leading us to judge the son to be evil.

In the case of the senator’s son, the grounds for inclusion are arbitrary: You are not my equal because you are a slave. This is no more arbitrary than to say you are not my equal because you are not human. We can see an echo of this throughout history. You are not my equal because you are female. You are not my equal because you are homosexual. Of course, we must be wary of not falling into this cycle of thinking: any of these distinctions would be non-arbitrary if they necessarily implied a consequence which would justify exclusion from equal treatment: recall Fjellstrom’s definition of speciesism in the introduction to this thesis.

We have an intuitive grasp that not all can be treated equally, that not all *ought* to be treated equally. The question is as to where the line is drawn. To base our grounds of inclusion upon moral capacities maximizes our return on investment of finite resources, it would utilize self-policing to make efficient use of available resources, and it would support a system where resources are spent on those who can be of future value to others. To do otherwise would either be an inefficient use of available resources, or even an outright waste.

4.3 Means, ends, and rights

I have argued that the capacity for moral concern coupled with the ability of targeted helping are the necessary and sufficient grounds to be considered a member of the moral community. The moral community then endows its members with “certain rights and privileges, principally the second maxim of the Kantian imperative that they should never be treated as means only, but also as ends in themselves” and these members take precedence over non-members. I argued we ought to do so on pragmatic and utilitarian grounds. I will now focus on a more nuanced exploration of these rights and privileges.

In the original outline of this thesis, the analogy made between a moral community and an economic community was much stronger; members held the “currency” of moral behaviour, traded through the process of targeted helping, in exchange for access to moral resources. Unfortunately, due to time and space constraints, a discussion that would define “moral resources” to a sufficient degree is beyond the scope of this thesis. Thus, the second maxim of the Kantian imperative of ends and means and “certain other rights” have been used as a provisional placeholders. To this end, I do not consider this discussion to be either complete or totally satisfactory, and it is certainly an area for future work. Instead of a clear and concise argument, the objective of this section is to act as a marker, a guide for later expansion and improvement.

The first question one might ask is, why make use of the Kantian imperative at all? As I have repeatedly stated, I am making a pragmatic claim while Kant was making a deontological claim (Johnson, 2008). There is, however, an important pragmatic consequence to being treated as an end rather than a means. Being defined as a means implies an immediate curtailment of one’s freedom. In order to achieve some end, another has imposed their will upon your freedom for you to be utilized as needed. This may prohibit you, a member of the moral community, to act as you are morally compelled to do, and thus threatens the integrity of the moral community itself. In order to avoid a self-

destructive contradiction, members of the moral community must be treated as ends in themselves, and never simply as unwilling means. Furthermore, the moral community is construed in such a way that the end beneficiaries of the system are the constituent members. That is to say, the objective of the entire project is for the community members to be ends themselves; to include the Kantian imperative is both an act of surety against abuse and a natural incorporation given nature of the community.

It should be pointed out that the usage of the Kantian imperative is not without difficulties. Kant's exact formulation of the second maxim is as follows:

Act so that you use humanity, *as much in your own person* as in the person of every other, always at the same time as end and never merely as means.

(Kant, 2002: 46-47, emphasis added)

The problem arises in that Kant specifically states that one cannot use one's self as a means to an end. However, the very concept of the economically based moral community is predicated on the fact that members would be willing to allow themselves to be used as a means to others' ends. Indeed this detail of the imperative seems to contradict intuitive notions of moral actions, and certainly the emotivist understanding of morality discussed in Chapter 2. If I use myself to shove another person out of the way of a speeding car, have I performed an immoral act in using myself as a means to an end? To use oneself to the benefit of another is very much the template of a morally commendable act.

An advocate of the Kantian imperative might argue that significant moral weight is assigned to performing a duty such as saving or assisting another (Kant, 2002: 13) and that this overrides the clause in the second maxim. However, if that is the case then why include the clause in the maxim at all? Of course, Kant does specify that one cannot be used as a means to an end *only*, since a student uses a teacher as a means to an end, for example? While this may cover most cases, the speeding car example above may still

prove problematic. If we assume the act kills me, have I now reduced myself to a means only, since all further possibilities of my life have been eradicated in service to this act? To sidestep this issue, as the maxim applies to the moral community, I will simply remove the clause. Members of the moral community may always *voluntarily* use themselves as a means to an end³².

Before entering into a discussion of the specific rights which members of the moral community are endowed with, it is important to note that not all members will necessarily have the same rights. In the previous section I argued that rights are only extended to those who can responsibly exercise them. As Singer points out, it makes no sense to extend the right to vote to dogs as they could never understand the act of voting or its significance (2002: 1-2). These rights fall under one of two categories, the first of which I will refer to as “incidental rights”, not because they are inferior or are dependent on others, but because they do not relate to the functioning of the moral community. Conceivably, there is no reason why in principle the moral community cannot function in an autocracy just as well as in a democracy.

The second category of rights are “necessary rights”, that is they are rights which are necessary for the creation, maintenance, and optimum functioning of the moral community. For example, the right to life and its preservation would be classed as a necessary right. The moral community exists by virtue of the individuals who populate it. To exclude rights of this nature would not only be self-destructive to the moral community, but would limit the moral community’s efficacy, as the members are both benefiter and a class of resource from which the moral community derives benefits. It is these necessary rights which this discussion will be centred upon; however, one should keep in mind that this discussion does not entail exhaustive list of all rights.

³² This may be very similar to Kant’s formulation, especially where duty obligates us to use ourselves as means. However the explicit statement here discourages any possible misinterpretation of the maxim as it applies to the moral community.

While presenting my argument in the previous section, I alluded several times to the fact that members of the moral community have a right to freedom, specifically the freedom of movement. This is a necessary right for members to have in order for them to perform the acts required of them to belong to the moral community in the first place. One cannot engage in targeted helping if you are restrained from doing so.

For the same reason, freedom of expression may be included as well. Targeted helping should not be narrowly defined as entailing an act of physical strength. Teaching, soothing, debating, encouraging and many other acts of targeted helping are performed with utterances, gestures, and writing. Furthermore, the moral community itself will be in need of constant revision. Should marmosets be allowed membership? Should we expand the rights awarded to non-members, given some new abundance of resources? Questions such as these are not answered by a central authority. Rather, it is both the duty and privilege of members to maintain the moral community, and this is best performed when dissenting voices have the freedom to be expressed and heard.

For the optimal functioning of the moral community, members should also be granted a right to training. I use the term “training” to differentiate it from the right to education. Education is usually meant as formal education in literacy or whatever curriculum the state has defined as a minimum standard. The right to training is a right intended to ensuring that the potential skill of a member is fulfilled. Since potential will vary between species as well as individuals – it makes no sense to teach a dog algebra – and it cannot include a minimum standard for all members as a right to education does. Furthermore, the right to training is specifically a teleological right; it has the stated goal of improving the economy and efficiency of the moral community. A right education, by contrast, does not necessarily include such a teleological dimension. Education may be an end unto itself, or be done to enrich life in a way that cannot be quantified or be of immediate use to others. These two rights may seem quite similar, and indeed probably

have considerable overlap, nonetheless I believe it is important to define rights in this way so as to better delineate the scope of the moral community. Being a member of the moral community does not necessarily grant you the wider right to education, and being a community outsider does not necessarily preclude you from that right.

The right to training is not unproblematic. Human beings can understand the goals and objectives of a training programme or a formal education system, and we may voluntarily choose to partake or abstain. Dogs or chimpanzees on the other hand raise an issue, as we cannot receive a meaningful sign of consent. Their training would be imposed upon them because it would be useful to the moral community – they would in a way be treated as means and not ends.

There are two possible responses to this issue, although neither of them is perfect. We can say that even human children are required to partake in some measure of training and education in any culture. Even the most obstinate libertarian must acknowledge that there are certain times where another must make a decision on your behalf and enforce it. Not all first graders are eager for school on a Monday morning, after all. To train members who cannot fully appreciate the goals and rewards of their training may be said to be akin to this. Alternatively, we can say that we do indeed receive some sign of consent from an animal in training. Labradors that gleefully chase decoys while being trained to be gundogs are at least in some way communicating their pleasure in being involved in the training, even if they do not actually understand all facets of the activity.

The problem with the first response is that sending children to school, even against their sullen will, is ultimately for their own benefit. They are still ends in themselves, as opposed to the right to training, where members are trained in order to benefit others. However, we might say that due to the reciprocal nature of the moral community, to help others is to help oneself; and therefore, in the long run the right to training is still to one's own benefit. The problem with this response is that we risk conflating pleasure with

consent. It is pleasurable to eat Turkish delight, but that does not mean that I have given consent to be force-fed it. Although, again, we can say that the Labradors are voluntarily participating in their training and are not being catapulted after the decoys. We may say that we have implied consent of the trainee, and if we predicate this on the argument of interpreting intent from behaviour in Chapter 3, I believe we have a workable, if imperfect, solution.

To summarize then: members of the moral community must always be treated as ends in themselves, even though they may treat themselves as a means. Whereas to use oneself merely as a means to an end is prohibited by Kant, it is a prohibition that I have disregarded as 1) it contradicts the understanding of morality presented in Chapter 2, and 2) the moral duty attached to assisting others might be argued to supersede the prohibition. Members necessarily have the rights to life, to freedom of movement, to freedom of expression, and to the right to training. Furthermore, to ensure the integrity of the community and free moral action of its members, these rights supersede non-members' rights in the event of a conflict. These are rights are not necessarily exhaustive, but are intended as a guide for future work and expansion.

4.4 Clearing the air: responding to possible objections

In this section I will respond to five possible objections that might be raised against the moral community as I have described it.

The first objection is that since membership to the moral community is effectively predicated on the ability to contribute, those who cannot contribute, such as babies, the autistic and so forth, would be excluded. This exclusion might seem unfair or even morally reprehensible, and thus to endorse the moral community is to exclude yourself since you are endorsing a morally reprehensible act, and thus may not have moral capacities yourself. The second objection is that simply having the capacity for moral concern and targeted helping does not guarantee that the subject will contribute to the community. To

include them in the community would be not much different from including community outsiders who do not fulfil the prerequisites, and to exclude them would be to violate the conditionals upon which the moral community is built. The third objection is that in this thesis I have argued for the existence of “true altruism” in Chapter 3, but the moral community is established quite firmly on reciprocal thinking. This is problematic as at the microscopic, individual, level reciprocal treatment of others is insufficient for belonging to the moral community. While the moral community itself entails a system of reciprocity which may ask of its members, who are required to be capable of true altruism, to put that altruism aside in the interests of the group. The fourth objection is that the moral community contradicts itself on premise 2. The community, as a system, uses members as a means to an end. The fifth and final objection is that it is unclear as to what happens to psychopaths and others who cannot engage in the emotional dimension of morality such as machine intelligences, but subscribe to a strict, rationally-based social contract theory. I will now respond to each of these objections in turn:

1. We will exclude those who ought not to be excluded, such as invalids, babies, the autistic, and the elderly.

This objection is based on the moral grounds that the moral community bears an unacceptable cost in exchange for its pragmatism. It would thus follow that anyone who advocates for the moral community does not necessarily belong to the community, and all who would belong to the moral community would not necessarily wish to. However, this objection is born of two errors. First, it assumes that to be excluded from the moral community would entail the end to that being’s rights or some other egregious suffering. This is, however, not necessarily the case. Non-members would still have rights as suits their particular nature or political character³³. These rights would still exist, but would be secondary to the rights of members in the event of conflicting rights. For example, an

³³ Depending on whether these are natural or legal rights, respectively.

insulin-dependent diabetic's right to life, if a member of the moral community, would supersede the relevant rights of the community outsider from whom the insulin is harvested.

Second, the criteria for membership to the moral community include the *capacity* for moral concern, and the *capacity* for targeted helping. It is not a question of how extensive these capacities necessarily are, as long as members are capable of considering the rest of the moral community. An elderly person or an invalid, for example, might not have the same capacity to help as a neurosurgeon or an Olympic power lifter, but as they are not completely incapable of the prerequisites, they would remain full members of the moral community. In addition, should it ever be discovered that people lose these capacities beyond a certain age, it would still be in the interests of the moral community to maintain a "retired" membership. To do otherwise would be for every active member to effectively bet against themselves for future care and consideration. Many of the motivations behind actions are investments in the future; squirrels store nuts for the winter, humans invest in retirement plans and so forth. Having a retired membership would capture this category of motivations, and members may support the moral community without fear of being left out after a lifetime of contributions. Should the above case become reality, it would be necessary to create this exception in the interest of maintaining the moral community.

Infants and people in vegetative states are a different matter, but can still not be considered to be sound counterexamples. The moral community is chiefly founded upon the concept of mutual assistance mediated through long-term investment strategies. With that in mind, the care and investment in infants should in fact be one of the primary concerns of the moral community. People in vegetative states might be considered in a similar manner, but this would vary depending on the specifics of the condition. Should their potential recovery be nothing short of a medical miracle, then expulsion from the moral community would be a rational course of action. Distress and emotionally-based

compulsion aside, this is reasoning similar to that used by people who maintain their loved ones on life support; they are banking on the possibility of recovery.

2. Morally capable but apathetic beings – sullen teenagers, depressives, nihilists etc. – would be excluded.

Again, this objection is based on the same error as above. So long as a being has the capacities, s/he has the membership. That being said, an apathetic member is not only a latent resource to the moral community, but a potential liability as well. In such a case it is very much within the interests of the moral community to motivate the apathetic and turn them into a contributing member.

3. Chapter 3 acknowledges the existence of true altruism and is a requirement for membership, but the moral community itself is a reciprocal system.

On the surface this may appear to be a contradiction; however, there are two errors here. Firstly, to say that true altruism exists in the natural world – as I argued it does in Chapter 3 – and to then say that we must therefore structure our ethics around this would be to subscribe to the naturalistic fallacy. I am arguing for us to take control of this tool with which nature has furnished us and to make the most effective, pragmatic use of it.

Secondly, while the principle of reciprocity is central to the justification and functioning of the moral community, strict reciprocal altruism is insufficient for allowing one to be a member of the moral community. By “strict” I mean that altruistic behaviour ends the moment one party has the opportunity to act with impunity. If I work as a cashier, and I am only motivated by strict reciprocal principles, and I know that due to poor bookkeeping that I can steal from the register without being caught, then I would do so. This violates both the second pragmatic argument and utilitarian arguments for the establishment of the moral community as laid out in section 4.3.

We can, however, say that the community morally operates at two distinct levels. At the microscopic, members-level, true-altruism is in operation, while at the macroscopic, community-level, reciprocal altruism is in effect.

4. The moral community contradicts itself on premise 2. The moral community, as a system, uses members as a means to an end.

This is a legitimate objection to the moral community as I have described it in this chapter. Nonetheless there are two possible defences, of which I prefer the latter. Firstly, while it is true members are used as a means by the moral community as a whole; they are simultaneously the ends to which the community is directed. This holds true even where in a particular time period member A is used³⁴ to the benefit of member B, as the moral community is structured around the principle of reciprocity A may later benefit from member B or C. It is only in a temporally short-sighted view that it might be said that members are used as a means to an end.

Moreover, the phrasing of the second maxim states that one must be treated as an end and “never *merely* as means” (Kant, 2002: 47, emphasis added). We might debate as to how far we can push this qualifier of “merely” a means, since members are means to their own ends; however, I would rather defer to the second defence:

The usage of the Kantian imperative is an imperfect compromise; as such, we should expect some inconsistencies. This particular inconsistency only holds at the system-level perspective. At the interpersonal level, members voluntarily use themselves as means to assist others, which as discussed in the previous section is not a violation of the edited Kantian maxim.

³⁴ That is, member A voluntarily uses himself or herself as a means to an end. For a member to be an unwilling means to an end would be a violation of the rights and privileges afforded to members, rendering membership meaningless.

5. What happens to psychopaths and others who cannot engage in the emotional dimension of morality such as machine intelligences, but subscribe to a strict, rationally-based social contract theory?

While not a direct objection, this question does raise a concern regarding the specific parameters of the moral community. I argued in Chapter 3 that emotions are fundamental to morality, and thus would be just as fundamental to a moral community. So why refer to a “moral” community at all? Why should this not be an endorsement of a rational community or a reciprocal community? After all, being clinically diagnosed as a psychopath does not guarantee that one will commit a murder or some other harmful act, and if we assume commitment to social contract theory, those in question could still form a reciprocal relationship with the moral community.

It should be first pointed out that social contract theory is grounded on the principle of reciprocity (Rousseau, 1895: 56, 73, 112; Binmore, 2004: 19; Sacconi, 2004: 5), and is thus subject to the same criticisms already given above for those who only adhere to reciprocal altruism. But there is an additional critique which can be added against emotionally incapable beings joining the moral community. The nature of a long-term investment strategy is such that it is not always foreseeable that one will eventually be rewarded for one’s efforts. This is especially so in the moral community, unlike in an economic community; the “return” on your investment will not necessarily be from the original target of your investments. It may be impossible to see how someone in a particular predicament could ever recompense you for your efforts; and so, from an absolutely rational perspective, there is no impetus to help. From a moral-emotional perspective, one might still help to either gain a sense of approbation or to avoid a sense of guilt or shame.

This reply is not completely satisfactory. While it is true that psychopaths, or beings who can only engage in reciprocal altruism, fail to answer the three claims delineated in

section 4.2, they are nonetheless of potential advantage to the moral community.

However, as they subscribe to reciprocal tenets this advantage will be lost if they have nothing to gain in assisting the moral community. This would be particularly damaging to the moral community if we assume their reciprocal altruism is a result of a complex cognitive capacity which distinguishes them from other outsiders, as their cognitive abilities could be useful if properly harnessed.

There is an answer with the potential to solve this conundrum. If we were to replace the simpler member/outsider dichotomy with a more complex tiered moral community, emotionally incapable beings might still partake in at least some of the rights and privileges of full members in exchange for their assistance. There are, however several reasons why I have decided to advocate for the simpler of the two models in this thesis: A tiered moral community would only be satisfactorily described with the use of moral resources.

Reciprocal altruism is not simply a case of “do this for me, and I do that for you” if “this” and “that” have vastly different costs in their performance, or if “this” and “that” can be found elsewhere at a lower cost. Moral resources would allow for a more nuanced trade than the current model, allowing for a tiered moral community that could fairly incorporate partial members. Also, the tiered structure itself is not without its own problems, more serious than that of inefficiency. It would necessitate a way of quantifying the degree of assistance the non-emotional being might offer, the degree and complexity of its social contract³⁵, and these two factors would need to be translated in to a value equitable to the moral-emotional value for trade in exchange for moral resources.

I believe that none of these problems are insurmountable, but solving them is far beyond the scope of this thesis.

³⁵ I.e. can the being only understand an immediate quid-pro-quo from a single beneficiary, or could it engage in long-term strategies with numerous and interchanging partners as with de Waal's computing ape previously discussed in Chapter 3?

4.5 Conclusion

In this chapter I have stated that we ought to create a “moral community”, a grouping of beings who differentiate themselves from others as having the capacity for moral concern for others, and who are able to engage in targeted helping. I argued that we ought to do so for three reasons: we ought to do this as it increases our powers over our environment and our own lives (the first pragmatic argument), that we save resources via the utilization of self-policing tactics (the second pragmatic argument), and that as long as need is greater than supply there ought to be a prioritization of investment in accordance with utility (the utilitarian argument). I also discussed an observation on the manner in which humans pass moral judgment on others, that our judgment is based on a subject’s moral capacity and its inclusion of deserving others. This suggests that we already have an intuitive grasp of the essential features of the moral community.

I then discussed in more detail the benefits of membership to the moral community. In no sense are these benefits a matter of moral desert, but rather rights and privileges which either maximize the power of the moral community, or protect it from destruction. They are purely *functional* benefits, which is to say rights necessary to the existence of the moral community and do not include any rights that are incidental to the moral community. These rewards were described in terms of the second maxim of the Kantian imperative – that members are always an end in themselves and never merely a means to an end – as well as the right to life, freedom of movement, freedom of expression, and a right to training. I also stipulated that as a matter of necessity these rights take precedence over any rights an outsider might have.

This discussion was not unproblematic, some issues arising from the fact that the second maxim and the collection of rights and privileges were used in this model as a provisional replacement for the concept of “moral resources”, as a discussion of such resources would require more time and space available to be included in this thesis.

Acknowledging this shortcoming, I also included a short response to five possible objections one might have against the moral community.

In this chapter I alluded to several other moral theorists who have argued for alternate metrics for rights and consideration. In the following chapter I will be looking at their arguments and will consider whether their arguments are valid or whether their models are preferable to my own.

5. Denouement

5.1 Areas for future research

Due to the limited scope and available time allowed for this thesis, a number of issues had to be left either unaddressed or only partially discussed. It should also be evident that much of the discussion in Chapter 2 regarding the nature of morality is grounded on very recent neurobiological research, and has perhaps has not withstood enough genuine attempts at falsification for the findings to be acceptable to the majority of academics. For example, Fumagalli et al.'s experimental results were puzzling and at some points contradictory, with only cautious speculation offered in explanation. Furthermore, the exact role mirror neurons play in human empathy, if any, is still a point of contention (Cook et al., 2014: 177).

In Chapter 2, I discussed the CAD triad hypothesis in support of a biological root of morality. However if it can be argued that not all altruism is equal, as I did in Chapter 3, can we not also ask if all morality is equal? The moral emotions of contempt and anger are theorized to function as a motivation to defend communal and individual integrity respectively, and thus they may have pragmatic benefits, but does moral disgust have the same usefulness? The answer to this question will shape the characteristics potential members will need to hold in order to belong to the moral community.

In my estimation, the most important area of research that would benefit this model would be research on the concept of moral resources. How might these be defined? Very loosely I would say moral resources are the generally unquantified costs attached to providing unreciprocated aid and upholding rights. But even accepting such a slipshod definition, the question arises, are all these resources traded in a finite, common currency? Is it ever possible to say that the capacity to share food is enough to earn the right to unlimited freedom of movement? The relatively simple dichotomy of community members and outsiders presented in this thesis sidestepped such problems, but at the cost of

nuance to the model. As was acknowledged in section 4.4, this loss of nuance led to a loss of efficiency, as those who can contribute in a limited way, and can also effectively strike or rebel in protest and thus become active liabilities to the community, are completely removed from the moral community.

Work that clarifies the concept of moral resources will open up another avenue for exploration, that of the possibility of a tiered moral community, which I briefly discussed in Chapter 4. The tiered moral community is a restructuring of the moral community where individuals such as psychopaths would have limited access to moral resources based on their contributions to the moral community.

The full implications of adopting my model of the moral community have also not been explored completely. In the introductory chapter, I alluded to the fact that the moral community could include machine or alien intelligences, just as much as animal intelligences. With regards to alien intelligences, Robert Freitas raises many relevant questions as to how we might relate to them, and tentatively suggests that aliens' moral capacities should determine our relations (Freitas, 1977: 15-17). I believe my conception of the moral community would be sufficient to allow for this issue, but this question is deserving of more thorough analysis than can be provided here.

How a machine intelligence might be treated by the moral community is also a question deserving greater investigation. I have already argued in this thesis that moral-emotional capacities are essential to the moral community. It then follows that how the machine is treated is entirely dependent on the nature of emotions and the machine in question. Under current computing architecture, can the emotions of rage, love, pathos, shame, or any other be emulated, if not outright produced? Is there an architecture that could ever do so?

Finally, should it come to pass that this model is put into practice, written into law and enforced, the largest research project of all would be the cataloguing of those who are members, and those who are outsiders. In the illustrative example in Chapter 4, I assumed that cows to be outsiders, but are they really? Are dogs, horses, chipmunks, trout, iguanas or tens of thousands of other species outsiders? How might we need to restructure our societies to accommodate an influx – or expulsion – of so many species from our current circle of concern?

5.2 Conclusions

The original intent of this thesis was to create a non-arbitrary, non-speciest model for determining whether or not a being is worthy of moral consideration. Evolution and pragmatism were accepted as the operative paradigmatic points of departure.

Speciesist models, which are implicit in many of our laws and social customs, are arbitrary models. That said, it is apparent that the fields of biology, ethnology, and primatology are still steeped in the speciesist perspective. It follows that any thesis that wishes to base itself on such empirical work will need to adopt at least a quasi-speciesist formulation. While still making use of the species classification system, I have nonetheless endeavoured to avoid any egregious speciesist stances such as “humans first, always.” While I have used such pronouns as “I”, “you”, and “we” to refer to the members of the moral community, I wish to stress this was only out of consideration for easier reading. The pragmatic theory put forward in this thesis would still remain true even if no humans populated the moral community.

A legitimate objection to this point might be to say that of course humans will populate the moral community, as other beings petitioning for entry to the community would be judged by the parameters of human morality. Thus, human principles and understandings would still be held at the pinnacle of the moral community’s hierarchy. To counter this objection, a two-pronged response was prepared in the second and third

chapters: (i) to define the nature of the moral domain, and (ii) to untangle the concept of humanness and morality, respectively.

In the defining the moral domain my objective was not to provide moral prescripts: “This is morally acceptable, that is not”. Rather, it was to analyse the nature of morality itself. This was done on two fundamental and often debated continuums, emotive theories versus rationalist theories, and core versus veneer theories. On the former continuum, a review of a large number of fMRI studies conducted over the last decade has provided compelling evidence for a dual model of moral deliberation which Greene et al. (2010) advocated for. I argued for a stronger role of emotional forces in the deliberation process than Greene et al. provided. The pertinent implication being that a moral community is not synonymous with a rational or intellectual community. Indeed, the findings of Fumagalli et al. (2010) suggest that emphasizing rationality over emotionality may lead to cruel or criminal behaviour.

With regards to the core versus veneer continuum, I argued strongly in favour of a core understanding of morality. That is to say, the meta-rules of moral behaviours are a product of evolutionary selective pressures and are largely hardwired into the brain. It would of course be as absurd to say that there is a “don’t steal gene” as it is to say there is a “nose-picking gene”. How those meta-rules are expressed, and which are dominant, are shaped by environmental factors. I discussed the CAD Triad Hypothesis and how it exemplified this.

To summarize, in Chapter 2 I argued for a dual process of moral deliberation, favouring emotive pressures. I also argued that morality is necessarily derived from certain inherent morphological and neurobiological structures, while sociological factors played a non-essential formative role. After all, as the primatologist Frans de Waal once observed, if we are not good at our core, how could a society ever spring forth to teach us morality in the first place? (Ross, 2009)

With a clear conception of morality in hand, I proceeded in Chapter 3 to untangle the concepts of humanness and morality. Some contemporary theorists, such as Ayala, still insist that morality is strictly a human characteristic, and argue that to see any of these characteristics in animals is to commit the veritable sin of anthropomorphism – projecting human structures of consciousness onto non-human animals. Thus to clarify my subsequent arguments, I presented a case for a critical or rational understanding of anthropomorphism as formulated by de Waal, where it is permissible to project human traits onto animals if doing so increases our ability to predict their behaviour, and that this projection is always subject to change in the event of a superior projection or method. To do otherwise would be a disservice to our understandings of other animals, an act of anthropodenialism. Rational anthropomorphism was presented in opposition with emotional anthropomorphism. Emotional anthropomorphism which, *contra* de Waal who presented it in a negative light, I argued may play an important role in group identification and empathy.

Given rational anthropomorphism I investigated three case studies of altruism in the animal kingdom, with the aim of demonstrating that not all altruism is equal. The first was that of kin altruism in bees. Here I argued that while these animals performed altruistically in the biologically defined sense – acting in such a way that increases another's chance of propagating their own genes while reducing your own – this is not true altruism – due to the close genetic relationship between the bees involved, their acts were in fact selfish acts.

The second case of altruism investigated was that found in birds, where a bird would give a reflexive warning cry of an approaching danger to its own flock. While this is indeed in line with true altruism – given empirical studies which show that birds flock as a mixture of families – I argued that in itself, the capability of true altruism alone is not enough to qualify one as a member of the moral community. There are two further

requirements: 1) a potential member also needs to have the ability to extend moral concerns to the rest of the moral community, which birds have not been found to do, and 2) one must also be capable of targeted helping.

Targeted helping is the ability to identify the cause of distress in another, formulate a plan to alleviate that distress, and then attempt to execute that plan. This is exemplified in the third case study of a helpful chimpanzee witnessed by Frans de Waal, where an adult chimpanzee dragged a heavy tyre over to a juvenile's niche after the adult had seen the juvenile fail to do so with her own strength.

The aim of this section was threefold: to put an end to any notion the morality is an exclusively human trait, to argue that there are different breeds of morality, and to outline the required capacities one would need to be a member of the moral community. These include a morality which supports an encompassing true-altruism – that is, a moral capacity which can have an unselfish concern for others who may be different or distant from the self – and that of targeted helping.

With these findings in hand, Chapter 4 dealt with the heart of the thesis: the moral community itself. The argument may be summarized as follows:

- 1) Beings that are capable of morally considering others ought to be included in a moral community.
- 2) Membership to this moral community carries with it certain rights and privileges, principally the second maxim of the Kantian imperative that members should never be treated as means, but only as ends in themselves.
- 3) Additionally, the rights and privileges of members take precedence over those of non-members.

And we ought to do these things because:

- 4) This will maximize the moral community's return on investment.

This return comes about through reciprocal assistance members of the community lend one another; through members self-policing their own actions; and through making maximum use of available resources for individuals who in the future may be in a position to contribute to the moral community.

I then provided a preliminary discussion as to what “certain rights and privileges” might encompass. Rights to life, freedom, and training are preferentially given to members as, not only would these serve as motivators for members to support the moral community, but crucially these rights in themselves serve the function of maintaining the moral community and of improving its efficiency.

This thesis has advocated a strong pragmatic position: by structuring our social order according to moral criteria rather than speciesist criteria, we will make more efficient use of available resources. Reading this from a relatively privileged, most likely western, perspective it might seem odd or even downright greedy to speak of society’s limited resources. But one does not need look far to find stark poverty in the world. This is not only for human beings but for a multitude of other species as resources such as habitat space and food sources are placed under greater and greater stress. This thesis is the proposal of a method to ease this increasingly fierce and extirpative competition among certain species, and to make better use of what we have, through moral criteria. Morality would never have evolved if cutthroat competition was always a superior method of survival. To limit the benefits provided through moral capacities by arbitrary barriers such as speciesism is irrational, disadvantageous, and impractical.

References

Ayala, F. 1987. The Biological Roots of Morality. *Biology and Philosophy*: 235-252.

Binmore, K. 2004. Reciprocity and the Social Contract. *Politics, Philosophy & Economics*, **3**: 5-35.

Botterweck, G. & Ringgren, H. eds. 1997. *Theological Dictionary of the Old Testament*. 3rd ed. Stuttgart: Verlag W. Kohlhammer.

Brown, D. 1991. *Human Universals*. Philadelphia: Temple University Press.

Cavlieri, P. & Singer, P. 1993. *The Great Ape Project: Equality Beyond Humanity*. London: Fourth Estate.

Center for Disease Control 2013. *Vaccine Excipient & Media Summary*. Retrieved 1st May 2014, from the World Wide Web:

<http://www.cdc.gov/vaccines/pubs/pinkbook/downloads/appendices/B/excipient-table-2.pdf>

Chittka, L. & Jensen, K. 2011. Animal Cognition: concepts from apes to bees. *Current Biology*, 21(3): R116-R119.

Churchland, P. 2011. *Braintrust: what neuroscience tells us about morality*. New Jersey: Princeton University Press.

Cook, R. et al.(2014. Mirror Neurons: From origin to function. *Behavioral and Brain Sciences*, 37: 177-241.

Crow, T. 1993. Sexual selection, Machiavellian intelligence, and the origins of psychosis. *The Lancet*, 342: 594-598.

Cummings, L. 2005. *Pragmatics: a multidisciplinary perspective*. Edinburgh: Edinburgh University Press.

Darwin, C. 1872. *On The Origin of Species*. 6th edition. London: John Murray.

Darwin, C. 2009. *The Descent of Man and Selection in Relation to Sex, volume 1*. Cambridge: Cambridge University Press.

Dawkins, R. 1996. *The Blind Watchmaker*. New York: Norton & Company.

Dawkins, R. 2006. *The Selfish Gene*. 3rd ed. Oxford: Oxford University Press.

de Waal, F. 1996. *Good Natured: The origins of right and wrong in humans, and other animals*. Cambridge: Harvard University Press.

de Waal, F. 1997. Are we in anthropodenial?. *Discover*, 18(7): 50-53.

de Waal, F. 2006. *Primates and Philosophers*. New Jersey: Princeton University Press.

Dennett, D. 1991. *Consciousness Explained*. Little, Brown and Co.: Boston.

Dennett, D. 1995. Animal Consciousness: what matters and why. *Social Research* 62(3): 691-711.

Driver, J. 2009. *The History of Utilitarianism: Stanford Encyclopedia of Philosophy*. Retrieved 2nd May 2014, from the World Wide Web:

<http://plato.stanford.edu/entries/utilitarianism-history/>

European Medical Agency 2013. *Guideline on the use of bovine serum in the manufacture of human biological medicinal products*. Retrieved 1st May 2014, from the World Wide Web http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/06/WC500143930.pdf

Feddersen-Petersen, D. U. 2000. Vocalization of European wolves (*Canis lupus lupus* L.) and various dog breeds (*Canis lupus f. fam.*). *Archiv Tierzucht*, **43**(4): 387-397.

Fieldmanhall, O., Mobbs, D. & Dalgleish, T., 2013. *Deconstructing the brains moral network: dissociable functionality between the temporoparietal junction and ventro-medial prefrontal cortex*. Retrieved 1st December 2013, from the World Wide Web: <http://scan.oxfordjournals.org/content/early/2013/01/23/scan.nss139.full.pdf>

Fiske, A. 1992. Four elementary forms of sociality: Framework for a unified theory of social relations. *Psychological Review*, **99**: 689-723.

Fjellstrom, R. 2002. Specifying Speciesism. *Environmental Values*, **11**(1): 63-74.

Foerster, K., Valcu, M., Johnsen, A. & Kempenaers, B. (2006). A spatial genetic structure and effects of relatedness on mate choice in a wild bird population. *Molecular Ecology*, **15**: 4555-4567.

Foot, P. 1967. The problem of abortion and the doctrine of the double effect. *Oxford Review*, **5**:5-15.

Francis, R. 1990. Causes, Proximate and Ultimate. *Biology and Philosophy*, **5**: 401-415.

Freitas, R. 1977. Metalaw and Interstellar Relations. *Mercury*, **6**(3): 15-17.

Fumagalli, M.; Vergari, M.; Pasqualetti, P.; Marcegalia, S.; Mameli, F.; Ferrucci, R.; Mrakic-Spota, S.; Zago, S.; Sartori, G.; Pravettonia, G.; Barbieri, S.; Cappa, S.; Priori, A. (2010). Brain switches utilitarian behavior: does gender make a difference?. *Plosone*, **5**(1): 1-9.

Gazzaniga, M.S., Ivry, R.B., Mangun G.R. 2014. *Cognitive Neuroscience: The Biology of the Mind*. 4th ed. New York: Norton.

Gert, B. 2011. *The Definition of Morality: Stanford Encyclopedia of Philosophy*. Retrieved 26th September 2013, from the World Wide Web: <http://plato.stanford.edu/entries/morality-definition/>

Goodpaster, K. 2013. On Being Morally Considerable. In: R. Shafer-Landau, ed. *Ethical Theory: an anthology*. 2nd ed. Oxford: Blackwell Publishers Ltd. pp. 381-389.

Grammer, K., Fink, B., Moller, A. & Thornhill, R. 2003. Darwinian aesthetics: sexual selection and the biology of beauty. *Biological Review*, **78**: 385-407.

Greene, J. 2007. The secret joke of Kant's soul. In: *Moral Psychology*. Boston: MIT Press, pp. 35-79.

- Greene, J.D.; Cushman, F.A.; Stewart, L.E.; Lowenberg, K.; Nystrom, L.E.; Cohen, J.D. 2009. Pushing moral buttons: the interaction between personal force and intention in moral judgment. *Cognition*, **111**(3): 364-371.
- Greene, J.D.; Darley, J.M.; Sommerville, R.B.; J.D., Cohen; Nystrom, L.E. 2001. An fMRI investigation of emotional engagement in moral judgment. *Science*, **293**(5537): 2105-2108.
- Greene, J.D.; Morelli, S.A.; Lowenberg, K.; Nystrom, L.E.; Cohen, J.D. 2008. Cognitive load selectively intereferes with utilitarian moral judgment. *Cognition*, **107**(3): 1144-1154.
- Greene, J.D.; Nystrom, L.E.; Engell, A.D.; Darley, J.M.; Cohen, J.D. 2004. The neural bases of cognitive conflict and control in moral judgment. *Neuron*, **44**(2): 389-400.
- Green, L. 1928. The Negligence Issue. *The Yale Law Journal*, **37**(8): 1029-1047.
- Haidt, J. 2001. The emotional dog and its rational tail: a social intuitionist approach to moral judgment. *Psychological Review*, **108**(4): 814-834.
- Haidt, J. 2004. The emotional dog gets mistaken for a possum. *Review of General Psychology*, **8**(4): 283-290.
- Haidt, J. 2007. The new synthesis in moral psychology. *Science*, **316**: 998-1002.
- Haidt, J. & Bjorklund, F. 2008. Social intuitionists answer six questions about moral psychology. In: *Moral Psychology, Vol 2*. Boston: MIT Press, pp. 181-218.
- Haidt, J. & Joseph, C. 2004. Intuitive ethics: How innately prepared intuitions generate culturally variable virtues. *Daedalus*, **2**: 55-66.
- Haidt, J., Rozin, P., Lowery, L. & Imada, S. 1999. The CAD Triad Hypothesis. *Journal of Personality and Social Psychology*, **76**(4): 574-586.
- Hume, D. 2007. *An Enquiry Concerning Human Understanding and Other Writings*. Cambridge: Cambridge University Press.
- Huxley, T. 1894. *Evolution & Ethics and other essays*. London: Macmillan and Co.
- Jaworska, A. & Tannenbaum, J. 2013. *The Grounds of Moral Status*. Retrieved July 30th 2014, from the World Wide Web: <http://plato.stanford.edu/entries/grounds-moral-status/>
- Johnson, R. 2008. *Kant's Moral Philosophy: Stanford Encyclopedia of Philosophy*. Retrieved 13th May 2014, from the World Wide Web: <http://plato.stanford.edu/entries/kant-moral/>
- Jones, J.C.; B., Loring; Janowiak, W.; Laland, K.N.; Musselman, L.E.; Langlois, J.H.; Roggman, L.A.; Perusse, D.; Schweder, B.; Symons, D. 1995. Sexual Selection, Physical Attractiveness, and Facial Neoteny: Cross-cultural Evidence and Implications. *Current Anthropology*, **36**(5): 723-748.
- Kaminski, J., Brauer, J., Call, J. & Tomasello, M. 2008. Domestic dogs are sensitive to a human's perspective. *Behaviour*, **146**: 979-998.

- Kant, I. 2002. *Groundwork for the Metaphysics of Morals*. New York: Vault-Ballou Press.
- Kaplan, M. 2004. Save the Rhino Maggot!. *New Scientist*, **181**(2440): 40-43.
- Klinkenborg, V. & Modica, A. 2001. *Cow Parts: Discover Magazine*. Retrieved 1st May 2014, from the World Wide Web: <http://discovermagazine.com/2001/aug/featcow>
- Kohlberg, L. 1969. Stage and sequence: the cognitive-developmental approach to socialization. In: *Handbook of socialization theory and research*. Chicago: Rand McNally, pp. 347-480.
- Masolo, D. 2002. Rethinking Communities in a Global Context. In: P. Coetzee & A. Roux, eds. *The African Philosophy Reader*. 2nd ed. New York: Routledge, pp. 558-572.
- Melis, A.P.; Warneken, F.; Jensen, K.; Schneider, A.; Call, J.; Tomasello, M. 2011. Chimpanzees help conspecifics obtain food and non-food items. *Proceedings of the Royal Society B*, **278**: 1405-1413.
- Naeger, N.L.; Peso, M.; Even, N.; Barron, A.B.; Robinson, G.E. 2013. Altruistic Behavior by Egg-laying Worker Honeybees. *Current Biology*, **23**: 1574-1578.
- Nakamura, K. 2012. *The Footbridge Dilemma: Effect of number of victims in moral dilemmas*. Retrieved 27th June 2013 from the World Wide Web: <http://mindmodeling.org/cogsci2012/papers/0148/paper0148.pdf>
- Newman, B. 2014. *Dennett's Consciousness Explained: It's critics and the controversy over the "true nature" of consciousness*. Retrieved 5th August 2014 from the World Wide Web: https://philosophy.stanford.edu/apps/stanfordphilosophy/files/wysiwyg_images/newman.pdf
- Okasha, S. 2013. *Biological Altruism: Stanford Encyclopedia of Philosophy*. Retrieved 27th March 2014 from the World Wide Web: <http://plato.stanford.edu/archives/fall2013/entries/altruism-biological/>
- Paxton, J. & Greene, J. 2010. Moral reasoning: hints and allegations. *Topics in Cognitive Science*, **2**:511-527.
- Preston, S. & de Waal, B. 2002. Empathy: Its ultimate and proximate bases. *Behavioural and Brain Sciences*, **25**: 1-72.
- Prinz, J. 2007. *The Emotional Construct of Morals*. Oxford: Oxford University Press.
- Prinz, J. 2011. *Morality is a culturally conditioned response*. Retrieved 15th May 2012, from the World Wide Web: http://philosophynow.org/issues/82/Morality_is_a_Culturally_Conditioned_Response
- Poulshock, J.W. 2006. *Language and Morality: Evolution, Altruism, and Linguistic Moral Mechanisms*. PhD Thesis. University of Edinburgh: Edinburgh.
- Pruetz, J. 2011. Targeted helping by a wild adolescent male chimpanzee (*Pan troglodytes verus*): evidence for empathy?. *Journal of Ethnology*, **29**: 365-368.

Queller, D. & Strassmann, J. 1998. Kin Selection and Social Insects. *Bioscience*, **48**(3): 166-175.

Robertson, J. 2004. *One cow, hundreds of uses: AR News*. Retrieved 10th May 2014, from the World Wide Web: <http://lists.envirolink.org/pipermail/ar-news/Week-of-Mon-20040105/015167.html>

Rousseau, J. 1895. *The Social Contract or Principles of Political Right*. London: Swan Sonnenschein & Co.

Rushton, J., Russell, R. & Wells, P. 1984. Genetic Similarity Theory: beyond kind selection. *Behavior Genetics*, **14**(3): 179-193.

Sacconi, L. 2004. Corporate Social Responsibility (CSR) as a Model of "Extended" Corporate Governance. *Liuc Papers*, **142**: 1-49.

Schwartz, S. & Bilsky, W. 1990. Toward a theory of the universal content and structure of values. *Journal of Personality and Social Psychology*, **58**: 878-891.

Shweder, R., Much, N., Mahapatra, M. & Park, L. 1997. The "big three" of morality (autonomy, community, divinity), and the "big three" explanations of suffering. In: *Morality and Health*. New York: Routledge.

Singer, P. 1993. *Practical Ethics*. Cambridge: Cambridge University Press.

Singer, P. 1997. *How Are We To Live?*. Milsons Point: Random House Australia.

Singer, P. 2002. *Animal Liberation*. New York: HarperCollins.

Singer, P. 2006. *In Defense of Animals: the second wave*. Oxford: Blackwell Publishing.

Sneddon, A. 2007. A social model of moral dumbfounding. *Philosophical Psychology*, **20**(6): 731-748.

Teffo, L. & Roux, A. 2003. Themes in African Metaphysics. In: *The African Philosophy Reader*. New York: Routledge, pp. 161-174.

The Holy Bible 1984. *New International Version*. Grand Rapids: Zondervan.

Tomasello, M., Davis-Dasilava, M., & Bard, K. 1987. Observational Learning of Tool-Use by Young Chimpanzees. *Human Evolution*, **2**(2): 175-183.

Turiel, E. 1983. *The Development of Social Knowledge: morality and convention*. Cambridge: Cambridge University Press.

Van Haastert, J. 2011. Amoeboid Cells Use Protrusions for Walking, Gliding and Swimming. *PLoS One*, **6**(11): 1-5.

van Rooyen, F. 2010. *Hippos Saves Wildebeest, Zebra Babies: News24*. Retrieved 10th April 2014, from the World Wide Web: <http://www.news24.com/SciTech/News/Hippo-saves-wildebeest-zebra-babies-20101203>

Vining, J., Merrick, M. & Price, E., 2008. The distinction between humans and nature. *Human Ecology Review*, **15**(1): 1-11.

Wade, M., 1980. Kin Selection: Its Components. *Science*, **210**(4470): 665-667.

Westneat, D. & Webster, D. 1994. Molecular analysis of kinship in birds. *Molecular Ecology and Evolution: Approaches and Applications*, **69**: 91-126.

Wilson, E. & Holldobler, B. 2005. Eusociality: Origin and consequences. *Proceedings of the National Academy of Sciences of the United States of America*, **102**(38): 13367-13371.

Wyatt, G., West, S. & Gardner, A. 2013. Can natural selection favour altruism between species. *Journal of Evolutionary Biology*, **26**: 1854-1865.

Yamamoto, S., Humle, T. & Tanaka, M. 2012. *Chimpanzees' flexible targeted helping based on an understanding of conspecifics' goals*. Retrieved 17th April 2014, from the World Wide Web: <http://www.pnas.org/content/early/2012/01/30/1108517109.full.pdf+html>