

THESIS

OF MICE AND KANT:

RE-EXAMINING MORAL CONSIDERABILITY TO NON-HUMAN ANIMALS ON

KANT'S COGNITIVE GROUNDS

Submitted by

William Eric Easley

Department of Philosophy

In partial fulfillment of the requirements

For the Degree of Master of Arts

Colorado State University

Fort Collins, Colorado

Summer 2014

Master's Committee:

Advisor: Bernard E. Rollin

Jane Kneller

Vicki J. Volbrecht

Copyright by William Eric Easley 2014

All Rights Reserved

ABSTRACT

OF MICE AND KANT: RE-EXAMINING MORAL CONSIDERABILITY TO NON-HUMAN ANIMALS ON KANT'S COGNITIVE GROUNDS

In this thesis, I examine Kant's criterion for moral considerability in light of the intersection between the moral, critical, and epistemic principles Kant commits himself to and evidence of advanced cognitive capacities in non-human animals. As I argue, Kant's denial of crucial cognitive capacities in non-human animals represents a flawed attempt at applying a principle of parsimony which threatens to undermine the transcendental base of his theories. Further, expelling the anachronisms and human exceptionalism Kant fell victim to in his theories reveals a robust sense of ethical duties directly to non-human animals, beyond non-cruelty. In Chapter One, I argue that the basis of moral considerability in Kant's ethics ought to extend directly to non-human animals if they possess sufficient degrees of the three cognitive capacities that comprise dignity and the ability to meaningfully set ends: reason, autonomy, and self-consciousness. In Chapter Two, I examine Kant's cognitive theory and argue that it lacks a developmental account in terms of degrees of these capacities that is crucial to completing Kant's ethical project. In Chapters Three and Four, I develop a model for such an account based upon evidence and theories in the philosophy of mind and the sciences, concluding that many non-human animals do possess advanced cognitive capacities and the we, thus, have moral duties directly to most non-human animals.

ACKNOWLEDGEMENTS

Foremost, I would like to thank my thesis advisor, Dr. Bernard E. Rollin, for his guidance and support throughout the stages of this project, most especially for his candidness, thoughtfulness, and for permitting me the freedom to fail and all the lessons that come with it. I would also like to thank Dr. Jane Kneller for her expert direction in navigating Kant's labyrinthine theories, and Dr. Vicki J. Volbrecht for her thoughtful comments and the courage to immerse herself in this often ponderous project. Further, I would like to express my indebtedness to the faculty and staff of the Philosophy Department for their support, discussion, and comments; my fellow graduate students for their attentive ears and seemingly endless talent and insight; and, of course, my family and friends for their unwavering encouragement and love, despite never wanting to hear another word about philosophy. Finally, I must give special thanks to Dr. Steve Naragon, whose scholarship and patient correspondence has been invaluable in uncovering and deciphering Kant's views regarding non-human animals and orienting my project among them.

DEDICATION

To my late father, Wayne E. Easley, who showed me what it is to be human
...and to “the brutes” who continually demonstrate that there is more to being human than
simply being a human.

TABLE OF CONTENTS

ABSTRACT.....	ii
ACKNOWLEDGEMENTS.....	iii
DEDICATIONS.....	iv
CHAPTER ONE: THE MEASURE OF MORAL WORTH.....	1
§1.1 Introduction.....	1
§1.2 Ethical Duties in Kant’s Ethics.....	3
§1.3 Non-Cruelty to Animals, as Perfect and Imperfect Duties to Ourselves	12
§1.4 Respect, Dignity, and Duties to versus Duties with Regard to.....	16
§1.5 The Dilemma of Moral Considerability.....	20
§1.6 Setting the Bar for Cognitive Capacities	26
§1.7 Anachronism, Religiosity, and Speciesism.....	33
CHAPTER TWO: THE GROUNDWORK OF COGNITION.....	41
§2.1 Introduction.....	41
§2.2 Kant’s Theory of Cognition.....	42
§2.3 Why There Must be Degrees of Self-Consciousness, Rationality, and Autonomy ...	55
CHAPTER THREE: DEGREES OF SELF-CONSCIOUSNESS AND COGNITION: Part I: From Transcendental Conditions to Self-Conscious Reflection.....	65
§3.1 Introduction.....	65
§3.2 A Partner in Crime: Gennaro’s Model of Degrees of Self-Concept.....	67
§3.3 Level Zero: The Non-Conceptual Kantian Base of Self-Conscious Experiences	69
§3.4 Level One: Bodily Self-Consciousness	75
§3.5 Level Two: Concepts, Facts, and “I” as Experiencer of Mental States	78
§3.6 Level Three: Reflection, Episodic Memory, and “I” as Enduring Thinking Thing ..	80
§3.7 Interlude: Recapitulation and Clarification.....	86
CHAPTER FOUR: DEGREES OF SELF-CONSCIOUSNESS AND COGNITION: Part II: From Theory of Mind to Moral and Parsimonious Theorizing	93
§4.1 Introduction.....	93
§4.2 Level Four: Implicit Autobiographical Memory and a Theory of Mind	94
§4.3 Level Five: Autobiographical Narrative Enhancement and Moral Agency	104
§4.4 Full Circle: Moral Considerability for Kantian Ethics Based Upon Cognitive Lives.....	112
§4.5 Subjectivity, Scientific Observation, and Parsimony	113
References.....	134

CHAPTER ONE
THE MEASURE OF MORAL WORTH

“All animals are equal, but some animals are more equal than others.”

--George Orwell, *Animal Farm*

“Trouble with mice is you always kill ‘em.”

--John Steinbeck, *Of Mice and Men*

§1.1 Introduction

Grounding an ethical theory that includes non-human animals within the realm of moral concern presents many issues. Some of these issues are practical, such as deciding when human interests are permitted to override concern for animal welfare. But, perhaps the deepest issues are theoretical: how do we account for the differences between humans and non-human animals in such a way that we do not also drive too wide of a gulf between the similarities? Moreover, the application of whatever criteria constitutes moral considerability must not unjustifiably exclude beings which would otherwise meet the criteria. If non-moral considerations—such as race, nationality, gender, handicap, or, in this case, species—are used to exclude beings from moral considerability despite their meeting the criteria, then those considerations are not morally relevant, and, thus, unjustified. Such inconsistent application of moral criteria would violate the stabilizing principles of rational consistency and universality which characterize moral claims in most ethical theories and prevent our moral theories from degrading into arbitrary and pernicious forms of relativism.

Put another way, unless membership in a species can be shown to be morally relevant, we do not want a special ethics for humans and a special ethics for non-human animals. Thus, we require a theory that fits humans and non-human animals on the same scale of moral worth with differences of degree, but degrees of the same, or similar, kind. Constructing such a scale

requires identifying and isolating what criterion constitute moral worth. Though a capacity to suffer often serves as the *prima facie* criteria, determining the presence of such a capacity involves determining the presence of a conscious cognitive mental life, or the capacity to have experiences which meaningfully organize and guide the life and ends of that being.¹ As such, I believe that a conscious cognitive mental life ought to be the criteria for moral worth and, in this chapter, I will ground this scale of moral worth in Kant's ethical theory.

This project presents its own difficulties, as Kant's ethics are often criticized as establishing little to no concern for non-human animals. In fact, Kant explicitly denies moral considerableness to non-human animals, and explicitly for the reason that they do not have the same kind of cognitive lives as humans. So, one might immediately wonder why I would choose to ground my project in Kant's theory. First, recent scholarship has shown that non-cruelty to animals actually holds a significant place in Kant's ethical theory. Second, though I will utilize this scholarship as a starting point, I believe we can extract more out of Kant. Third, I believe that Kant's denial of the right sort of cognitive lives in non-human animals is based upon anachronistic concerns that can be ejected from his theory while still preserving the essence of the theory.

As I will argue, Kant's theory places stress on cognitive capacities as the ground of moral worth in two ways: first, on the cognitive life of the being who has moral obligations and, second, on the cognitive life of the beings one is obligated to. This stress on cognitive lives, I believe, means that Kant's ethical framework ought to apply to non-human animals as well, if they

¹ Suffering, as a criterion, distinguishes between inanimate objects and the life of beings, but the presence or absence of a conscious cognitive life is how we intuitively distinguish between beings which can suffer and beings which cannot in any meaningful way. For instance, some might be tempted to point to a wilting plant as an instance of suffering, but since we do not usually attribute consciousness to plant life, its suffering would not matter in the same way as a human or a non-human animal. Thus, the presence of a cognitive life, here, distinguishes between rocks and plants on the one side, and humans and non-human animals on the other. Whether plants ought to be classified this way is beyond the scope of this thesis.

exhibit similar capacities. Of course, establishing that non-human animals do, in fact, possess these capacities is crucial and I will take this on in Chapters Three and Four. But, first, it must be established that in Kant's view if a being has such capacities, then it has moral worth. I will begin this effort by discussing the non-cruelty aspects of Kant's ethics. Then, I will proceed to place stress on the inconsistencies and anachronisms in Kant's view by utilizing important aspects of Kant's ethics and common sense moral intuitions regarding moral considerableness. Finally, I will conclude that this process of carefully ejecting anachronisms yields a reading of Kantian ethics that stays true to the essence of Kant's theory while placing moral worth exactly where we need it—on the cognitive lives of beings. This, in turn, will yield an ethical framework that includes non-human animals as objects of direct moral concern based on respect for their cognitive lives as importantly similar to our own.

§1.2 Ethical Duties in Kant's Ethics

Generally speaking, a deontological ethical theory is one which concerns what we ought to do, as guided by what is forbidden, required, or permitted rather than by the foreseeable consequences of an action. Kant's ethics is a deontology, but it is somewhat unique in that the law that we are duty-bound to follow does not come from an external source, such as religious dogma or societal laws, but comes from us, internally, as a matter of how we are constituted as cognitive beings. Possible actions are "represented" in us by the same faculties of cognition that represent our thoughts to us in our mind, with the distinction that some possible actions are represented as obligatory—as unconditionally binding on the exercise of our will.² Moreover,

² Cf. Immanuel Kant, *Metaphysics of Morals*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996) 6:222. Hereafter, *MS*. Also, the Cambridge editions of Kant's works follow a numbered format in the margins of the text of volume: page number that references the standard German, or the

our faculty of reason represents the obligatory action as universally binding on all rational agents. As Kant puts it, “Any [moral principle] is really an obscurely thought *metaphysics* that is inherent in every human being because of his rational predisposition.”³ By metaphysics, Kant means that the obligatory nature of the action appears to us as an objective feature of the world accompanying the cognition of the action. These obligatory actions epitomize the moral law and our duty to adhere to it.

Yet Kant’s ethics is also a kind of virtue theory—one which calls us to cultivate our dispositions, and the sort of persons we are, toward acting not only in accord with the moral law, but for the sake of it. Doing so, for Kant, affirms our own humanity—our nature as self-conscious, rational beings with autonomous wills.⁴ Though these features are part of our constitution, it is by regulating our actions through reason that we cultivate virtuous dispositions and live up to our moral obligations. Of course, since we also see ourselves as autonomous—as free in the exercise of our will—we can always refrain from performing an action that is morally obligatory. Nevertheless, doing so goes against reason, or, more importantly, against our rational nature. In this sense, a moral law categorically applies to us in that we ought to follow it, even when we choose not to follow it.⁵

Academy, edition, *Kant’s Gesammelte Schriften*, edited by the Royal Prussian Academy of Sciences (Berlin: Georg Reimer, later Walter deGruyter & Co., 1900-). I have chosen to follow that pagination for each of Kant’s works.

³ Kant, *MS*, 6:376, emphasis in text. Hereafter, since Kant’s original writings heavily use emphasis where he thought it was important, and since the Cambridge editions have sought to preserve Kant’s emphasis for clarity and interpretation, all emphasis from Kant’s writings should be assumed as his unless otherwise noted.

⁴ Some may object to my not including here the capacity to set ends for oneself, and the capacity to legislate the moral law by the use of one’s will. However, both of these criteria are parasitic on the three criteria I do name—which I take as necessary, as one cannot set ends if one is not free to do so, one cannot legislate the moral law unless it is represented to oneself by reason first, and one cannot do either without being self-consciously aware of oneself as a being with a will, capable of taking action in the world. I will return to these points below in my main argument for including non-human animals under moral considerableness.

⁵ In her famous article, “Morality as a System of Hypothetical Imperatives,” *The Philosophical Review* 81 (1972): 305-316, Philippa Foot apparently misses this aspect of Kant’s ethics in her criticism that one has no reason to

Indeed, our autonomous nature is part of the problem for Kant—a problem that can be resolved by constraining our freedom through the regulative use of reason. As Kant sees it, it is entirely natural to act out of a desire for happiness—not only are we free to do so, in fact we can assume that this is an end of all rational beings.⁶ But, exclusively doing so out of pure desire and inclination tends to lead to a degradation of virtue, to a focus on the outcomes we desire, and to making exceptions for oneself at the cost of the happiness of others “unless a good will is present which corrects the influence of these on the mind and, in so doing, also corrects the whole principle of action and brings it into conformity with universal ends.”⁷ For Kant, worthiness to be happy is a matter of morally pursuing happiness as an end within a community of beings, rather than selfishly pursuing it.⁸ The moral pursuit of happiness entails recognizing the ends of these other beings and restraining one’s pursuit so as not to interfere, or harm, others, but, instead, to promote happiness for oneself and others. In other words, our freedom can be abused and needs to be restrained and guided by maxims which could be adopted by any rational being with a good will, not because of what it accomplishes immediately, but because it is the proper use of our will as regulated by reason.

Moreover, there are teleological aspects (i.e. a doctrine of design) to Kant’s theory—which, as I shall criticize below, often carry religious connotations. As Kant says, humans are “unholy enough that pleasure can induce them to break the moral law, even though they

follow a moral law without an incentive beyond the law applying to oneself. For Kant, though, the agent’s own reason, itself, represents the moral law to the agent, suggesting not only the law, but a reason to follow it as well. In other words, for Kant, there is always an agent-centered reason to follow the moral law, specifically because it is the agent, not something external, which represents the law to the agent.

⁶ Immanuel Kant, *Groundwork of the Metaphysics of Morals*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996) 4:415. Hereafter, *GMS*.

⁷ *Ibid.*, 4:393.

⁸ *Ibid.*, “a good will seems to constitute the indispensable condition even of worthiness to be happy.”

recognize its authority; and even when they do obey the law, they do it *reluctantly* (in the face of opposition from their inclinations).”⁹ Kant held that human nature has a dual-aspect: part “animal,” in terms of having sensibility and a will that is stimulated by the senses, and part “rational,” in terms of not merely being stimulated by the senses, but possessing understanding and reason which we can deploy in a regulative fashion over our wills and desires. Exercising our wills in this way allows a human to “raise himself from the crude state of his nature, from his animality...more and more toward humanity...so that he may be worthy of the humanity that dwells within him.”¹⁰ To continue in Kant’s words, “In other words, the human being has a duty to cultivate the crude predispositions of his nature, by which the animal is first raised into the human being.”¹¹ Thus, the morality that arises from within us, as self-conscious, autonomous, rational beings, comes in the form of ways in which we are obligated to constrain our wills by reason, both in the actions we take and in the dispositions which guide why we take those actions. These obligations come in the form of duties, the taxonomy of which will be vital to the arguments below.¹²

For purposes here, we can divide these duties into perfect and imperfect duties, each of which can be duties to ourselves or to others. Perfect duties have to do with rights in that they prohibit specific overt actions which would infringe upon the free exercise of a being’s will. As

⁹ Kant, *MS*, 6:380. See also 6:386.

¹⁰ *Ibid.*, 6:387. See also Immanuel Kant, “Universal Natural History and Theory of the Heavens,” in *Natural Science*, ed. Eric Watkins, trans. Olaf Reinhardt (Cambridge: Cambridge University Press, 2012) 1:356.

¹¹ Kant, *MS*, 6:392.

¹² It is worthy of note in regards to the preceding paragraph that this somewhat religious, historical, and perhaps outdated teleological view of Kant’s will become critical to deflating Kant’s divide between humans and non-human animals in the discussion below. Essentially, this view, by itself, only establishes aspects of human nature, it does not establish that non-human animals cannot have similar aspects of rationality and refrain from sensible desires, even to lesser degrees. Indeed, as a terminological note, Kant has a frustrating tendency to often use “animal” and “merely animal” to refer to aspects of human sensible perception, and rarely establishes anything directly regarding his views about non-human animals.

Kant puts it, a right is “the sum of the conditions under which the choice of one [being] can be united with the choice of another in accordance with a universal law of freedom.”¹³ This includes societal rights, such as freedom from harm, the essence of which is a respect for the dignity and humanity of oneself and the community of rational beings one belongs to.

Importantly, dignity and humanity are ambiguous concepts on Kant’s view, despite also being critically important to his moral theory, as well as many moral discussions in general (especially ones which exclude non-human animals from moral consideration). On the one hand, there is a religious undertone to these terms, insofar as their usage seems to invoke a reference to an inherently noble, even sacred, character that human life possesses by default. But, Kant seems well-aware that such usage is at least somewhat empty, as not all humans exhibit this character and, really, dignity is a feature of a moral person—i.e. one with a good will, as discussed above. A moral man or exemplar, for instance, may exhibit a dignified character and be said to possess much virtue or dignity, making their character seem noble and an inherent good in the world. In this sense, dignity is not a feature all humans have by default, but a capacity Kant believes all humans could, and should, develop as far as they can. On the other hand, though dignity and human dignity are often used by Kant as synonyms for humanity, as we shall see below, humanity, for Kant, is not a property exclusive to humans, but is, instead, a descriptive term for rational, autonomous beings who can self-consciously direct their wills. And so, as intuitive definitions for my purposes here, dignity comes down to rational autonomy, or freedom to direct one’s will toward one’s ends as guided by reason. Hereafter, all usages of these terms should be understood in this way, as a baseline of cognitive capacities that putatively constitute moral considerability. Though one ought, on Kant’s view, develop these capacities virtuously, in terms

¹³ Kant, *MS*, 6:230.

of perfect duties one must first be free to do so—not just in terms of a capacity to do so, but also in terms of not being interfered with by the actions of others.

In this sense, perfect duties represent what actions are prohibited, as to take them would directly undermine the dignity—i.e. the mental capacities and moral pursuits—of oneself or others, and, therefore, can be duties to oneself or to others. Moreover, these duties are necessarily good (logically), not because they are enforceable by law, but because they are logically required given a sense of duty.¹⁴ For instance, if it is immoral to hurt others, then I, logically, ought not stab someone. However, why one upholds a perfect duty can vary—as perfect duties limit actions, not the maxims which guide actions. For an action to be ethical in Kant’s view it must be done for the sake of the moral law, and though it is possible to uphold a perfect duty out of respect for the moral law, it is also possible to uphold perfect duties for other reasons as well. For instance, one may refrain from hurting someone else because one fears the repercussions, in which case one upholds the perfect duty they have to that other person, but not for the right reason and, thus, not morally, but only coincidentally.

Imperfect duties, however, are different and are, perhaps, much more important as they constitute the essence of what is ethical for Kant, as they are about ends. An end, for Kant, “is an object of the choice (of a rational being), through the representation of which choice is determined to an action to bring this action about,” and while the constraints of perfect duties can limit or require certain actions they cannot mandate particular ends or maxims, “only I myself can *make* something my end.”¹⁵ However, there are some ends we have a duty to adopt and attempt to bring about, as they are prescribed by the very concept of duty. We must keep in mind

¹⁴ Cf. Lewis White Beck, *Commentary on Kant’s Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), 151-152.

¹⁵ *Ibid.*, 6:381.

that, for Kant, morality is “the relation of actions to the autonomy of the will,”¹⁶ and “the necessity of my action from *pure* respect for the [moral] law is what constitutes duty, to which every other motive must give way because it is the condition of a will good *in itself*, the worth of which surpasses all else.”¹⁷ In other words, as we saw above, the representation of the morality within us comes in the form of obligations to constrain our wills and, while perfect duties constrain us externally (due to the freedom and dignity of others) and necessarily, imperfect duties require a self-constraint in the effort of developing the virtuous dispositions which constitute a good will and comprise our worthiness to be happy. Put simply, imperfect duties require us to choose virtue and virtuous ends, for virtuous reasons. Though these duties are wide in scope and contingent, insofar as maxims, intentions, and ends cannot be specifically restricted or specified in terms of how one ought to fulfill them, having the right maxims, intentions, and ends is a matter of the virtuous self-governance of one’s will.

To summarize and rephrase, then, perfect duties are prescriptive obligations to respect one’s own rational, autonomous nature, as well as to not interfere with the autonomy of others. Imperfect duties are prescriptive obligations to cultivate one’s own autonomous will and contribute to the happiness of others, as well as to the development of their autonomous wills. As such, there are two ends which we have a duty to adopt: our own virtue, and the virtuous ends of others. In terms of our own virtue, there are two parts: cultivating one’s faculties and cultivating one’s will. As we saw above, Kant holds that humans have a duty to rise above their animal nature. But, they also have a duty to respect their animal nature in terms of “self-preservation...the preservation of the species...and the preservation of his capacity to enjoy life,”

¹⁶ Kant, *GMS*, 4:439.

¹⁷ *Ibid.*, 4:403.

or, put negatively, to not commit suicide or defile oneself through gluttony, inebriation, or unnatural or lustful sex acts.¹⁸ Without much of a stretch, this can also be thought of as taking one's health and well-being as an end, including constraining one's desires. The cultivation of one's will, however, involves making the "maxims of his will [consistent] with the *dignity* of humanity in his person."¹⁹ This project means that one has a duty to "cultivate his natural powers...as means to all sorts of possible ends," "to be a useful member of the world," and to strive for perfection though this can only be met through continual progress.²⁰ In this sense, one has a duty to oneself to be a good person, or to affirm his worthiness to be happy by cultivating a good will.

Adopting the happiness of others as an end stems from love (or beneficence) and respect. Love, understood as "satisfaction in the happiness (well-being) of others; [with] beneficence [being] the maxim of making others' happiness one's end,"²¹ leads us to helping others (or taking their ends as our own), thus promoting their happiness. Though there are logical reasons that we have a duty to help others—namely, that we are limited beings and require assistance to achieve our own ends, thus we ought to help others if we expect them to help us²²—there are also deeper reasons to help others, stemming from a recognition of belonging to a community of beings with natures similar to our own, and the resulting obligation to constrain our wills in how we interact

¹⁸ Kant, *MS*, 6:420.

¹⁹ *Ibid.*

²⁰ *Ibid.*, 6:445-446.

²¹ *Ibid.*, 6:452.

²² See Kant, *MS*, 6:451. Also, 6:387 and the *GMS*, 4:423.

with others.²³ While helping others fulfills a duty to them, it also provides means to fulfill the duties to oneself, insofar as one makes the morally permissible ends of others one's own end (that is, one cultivates one's own virtue by helping others virtuously pursue their happiness). However, respect involves recognizing the basic dignity inherent in all rational, autonomous beings—or an “absolute inner worth” which constitutes their status as moral persons as ends in themselves rather than merely as means to one's own ends.²⁴ This respect is similar to the respect one must have of one's own dignity, but, here, this means that one cannot use others in ways which violate that person's dignity.

Thus, in Kant's ethics we are required to recognize in others what is worthy of respect in ourselves—namely, humanity, or our nature as self-conscious, rational beings with autonomous wills. Due to this nature, we set ends conducive to our happiness and through respect and love of dignity we have duties directly to ourselves and to other rational, autonomous beings as constraints on our freedom and ends both individually and collectively. However, these duties can also be understood as promoting freedom through the regulative use of reason, which leads to the development of the virtues constitutive of a good will and our worthiness to be happy. Thus far, this discussion has centered on obligations we have to other beings who have natures similar to our own, that is, to humans, with only broad strokes toward any non-human beings.

²³ Here, and in what follows, I am attempting to combine Kant's formulation of the Categorical Imperative in the “Formula of Humanity” and the “Formula of the Kingdom of Ends” in terms of love and respect for humanity. The purpose for this attempt is to circumvent a long discussion that would stray too far afield of my purposes in this chapter. For instance, not only would a full discussion require examining each of Kant's several formulations of the Categorical Imperative, there is also the issue that each formulation may not be equivalent, despite Kant's insistence that they are—an issue at least partially resolved by Bernard E. Rollin in “There is Only One Categorical Imperative,” *Kant-Studien* 67, no. 1-4 (Jan 1976): 60–72. For my purposes, what is important is Kant's emphasis on humanity, or autonomy, which is a consistent emphasis even if his formulations are not consistent. For the specific formulations, see Kant, *GMS*, 4:429 and 4:433, respectively.

²⁴ Kant, *MS*, 6:435. See also the “Formula of Humanity” referenced in the previous footnote.

With these resources in mind, we are now in a position to see how modern Kant scholarship applies this to non-human animals.

§1.3 Non-Cruelty to Animals, as Perfect and Imperfect Duties to Ourselves

As noted in the Introduction (§1.1), Kant's ethics are often criticized as exhibiting scant regard for non-human animals. A reading centering on this sort of criticism might point to passages in Kant such as, "*Respect* is always directed only to persons, never to things. The latter can awaken in us *inclination* and even *love* if they are animals...or also *fear*...but never *respect*,"²⁵ and "As far as reason alone can judge, a human being has duties only to human beings."²⁶ Since the basis of perfect duties is respect, and the basis of imperfect duties is respect coupled with beneficence, holding that respect can only be properly directed to persons and that animals are not persons suggests that we have no basis of regard for animals in Kant's ethics and makes it seem "as if Kant thinks an animal is no more worthy of our concern than is a turnspit on which we might choose to roast it."²⁷ But, this reading is not quite accurate. As some commentators point out, another reading is available which focuses on other passages such as, "We have duties, though, only towards other people; inanimate things are totally subject to our will, and the duties to animals are duties only insofar as they have reference to ourselves."²⁸ In this case, respect is present, but only in an indirect fashion—as respect for humans, not for

²⁵ Immanuel Kant, *Critique of Practical Reason*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996) 5:77. Hereafter, *CPrR*.

²⁶ Kant, *MS*, 6:442.

²⁷ Patrick Kain, "Duties Regarding Animals," in *Kant's "Metaphysics of Morals": A Critical Guide*, ed. Lara Denis (Cambridge: Cambridge University Press, 2010), 210.

²⁸ Immanuel Kant, "Moral Philosophy: Collins's Lecture Notes," in *Lectures on Ethics*, ed. Peter Heath and J.B. Schneewind, trans. Peter Heath (Cambridge: Cambridge University Press, 1997) 27:413. See also 27:347. Hereafter, *LE (Collins)*.

animals. Some criticize that this indirect respect is insufficient to capture our moral intuitions, and, in fact, my argument is based on this same point. However, it will be useful to first explore this reading of indirect respect as it is far more robust than it appears.

As Lara Denis has argued, paying special attention to specific passages in Kant yields a much deeper and important place for non-human animals in Kant's ethics—one which bars cruelty to animals due to its violation of our imperfect duties to ourselves because of an undermining effect on cultivating virtuous dispositions, and violating our perfect duties to ourselves by failing to uphold our own human dignity. One passage in Kant captures this so well that it is worth quoting in full:

With regard to the animate but nonrational part of creation, violent and cruel treatment of animals is...intimately opposed to a human being's duty to himself, and he has a duty to refrain from this; for it dulls his shared feeling of their suffering and so weakens and gradually uproots a natural predisposition that is very serviceable to morality in one's relations with other people. The human being is authorized to kill animals quickly (without pain) and to put them to work that does not strain them beyond their capacities...But agonizing physical experiments for the sake of mere speculation, when the end could also be achieved without these, are to be abhorred.—Even gratitude for the long service of an old horse or dog (just as if they were members of the household) belongs *indirectly* to a human being's duty *with regard to* these animals; considered as a *direct* duty, however, it is always only a duty of the human being *to* himself.²⁹

We see in this passage references to scientific experiments, to agricultural practices and animal husbandry, and, most importantly, to our relations with other people. As has been well-documented, cruelty to non-human animals can, potentially, have a detrimental desensitization effect on a person and his treatment of other people,³⁰ and, in Kantian terms, this effect runs contrary to our duties to respect human dignity and to cultivate our ethical virtues (as discussed

²⁹ Kant, *MS*, 4:443.

³⁰ For instance, see, as Denis points to, Stephen R. Kellert and Alan R. Felthous, "Childhood Cruelty toward Animals among Criminals and Noncriminals," *Human Relations* 38, no. 12 (December 1985): 1113-1129 or Temple Grandin, "Behavior of Slaughter Plant and Auction Employees Toward the Animals," *Anthrozoos* 1, no. 4 (1988): 205-213. Also, see Bernard E. Rollin, *Animal Rights and Human Morality*, 3rd ed. (Amherst, NY: Prometheus Books, 2006), 48-49. Hereafter, *ARHM*.

in the previous section). On the other hand, treating non-human animals properly can help foster our cultivation of virtues by developing the right dispositions.

As we saw in the previous section, Kant claims that we have duties to respect our partially animal nature. According to Denis our recognition of the value of our animal nature includes recognizing our “susceptibility to love and sympathy” and not acting in ways which impair that susceptibility, as to do so would violate our respect for our humanity, thus violating our perfect duties to ourselves.³¹ Moreover, since imperfect duties partially rest on love and beneficence “fostering dispositions that help us practically express...appreciation and concern for others is itself a duty to others” and by fostering these dispositions we also cultivate our virtues; failing to foster these dispositions by treating animals cruelly, therefore, violates our imperfect duties to ourselves and to others.³² Patrick Kain agrees, adding that “sympathetic feeling is a necessary precondition for moral obligation” and “we have a perfect duty to ourselves to preserve and neither denigrate nor demean [these feelings], in addition to...[having] an imperfect duty to ourselves to cultivate [these feelings].”³³ But, Denis extracts even more from this reading, adding that these considerations mean that “a good person cannot *adopt a policy* [or maxim] of ignoring animal suffering,” and that in addition to the issues Kant discusses in the passage above, we also may have duties to refrain from jobs which involve the harming or killing of animals, to refrain from meat-eating due to the harsh conditions involved in raising the animal, and, moreover, a duty to educate ourselves on where cruel treatment of animals persists

³¹ Lara Denis, “Kant’s Conception of Duties Regarding Animals: Reconstruction and Reconsideration.” *History of Philosophy Quarterly* 17, no. 4 (October 2000): 409. Hereafter, “Kant’s Conception of Duties Regarding Animals.”

³² *Ibid.*, 410.

³³ Kain, “Duties Regarding Animals,” 223.

and refrain from any knowing “support or benefit from that way of acting.”³⁴ Thus, on Denis’s reading the place for non-human animals in Kant’s ethics is both deep and important in the implications it has for how we ought to treat non-human animals.

For some, this might be enough, as in these considerations Kant not only grants that non-human animals can suffer, but that we must have regard for that suffering insofar as we ought not to cause more suffering than absolutely necessary. Thus, it appears this Kantian reading yields concern similar to the regard that Peter Singer’s ethics does—concern based upon mitigating suffering. However, though Denis argues that these duties are “well-grounded, genuine duties,” many will argue that this reading has concern for animals too easily overridden by human concerns, as the source of obligation has nothing to do with the animals directly, but only the effect such suffering has on humans.³⁵ For instance, as Bernard E. Rollin criticizes, “If we had good psychological evidence that certain sadistic individuals could expunge their brutality... say by torturing kittens, and thereby become more moral toward people, this view would not only permit the torturing, but would also seem to make it morally obligatory!”³⁶

It is also worthy of note that violating our perfect duties to others is not among the duties Denis, or Kant, discuss in regard to animals—as, presumably, not only do non-human animals not have rights, they do not possess the basis for such rights: the dignity inherent to self-conscious, rational, autonomous beings. Thus, animals are never considered persons and there is, then, concern that the duties yielded by Denis’s sort of reading may evaporate or collapse under pressure from practical human concerns into treating animals no better than things.

³⁴ Denis, “Kant’s Conception of Duties Regarding Animals,” 411-417, emphasis in text.

³⁵ *Ibid.*, 417.

³⁶ Rollin, *ARHM*, 49.

§1.4 Respect, Dignity, and Duties to versus Duties with Regard to

This worry about Denis's reading can be best understood by reference to Robert N. Johnson's discussion of Kant's distinction between duties to and duties with regard to. According to Johnson, there are three directions in which ethical obligations can extend: directly *to* myself, directly *to* others, and indirectly *with regard to* others. The former two are captured in our perfect and imperfect duties, and the directionality of the duty references the fact that we can only be obligated to beings with rational free wills (or, as I have used the terms above, humanity, or dignity).³⁷ In other words, the respect aspects of our duties can only map onto beings which are worthy of respect (persons) and only humanity is worthy of respect; though humanity (as a descriptive term) may apply to more than humans, only humans are known to possess humanity; thus, as above, respect "is always directed only to persons, never to things," and we can only have duties directly *to* human beings. For Kant, if one thinks one has duties *to* beings that have no humanity, then one has fallen victim to an amphiboly—a sort of equivocation based, here, upon an ambiguous sense of respect and humanity—and thereby mistakes what is actually a duty to oneself as a duty to that being by "mistaking his duty *with regard to* other beings for a duty *to* those beings."³⁸ Thus, the confusion is in the directionality of the obligation.

The third direction (with regard to) arises only when our direct obligations also indirectly reference another object or being. Johnson uses several examples to illustrate this. For instance, if I borrow my neighbor's lawnmower, I now have an obligation *to* my neighbor, not *to* the lawnmower, and my obligation is *with regard to* the lawnmower (i.e. that I ought to return it in

³⁷ Robert N. Johnson, "Duties to and Regarding Others," in *Kant's "Metaphysics of Morals": A Critical Guide*, ed. Lara Denis (Cambridge: Cambridge University Press, 2010), 195.

³⁸ Kant, *MS*, 6:442. Also, Cf. Kain, "Duties Regarding Animals," 221.

good condition).³⁹ Similarly, though more provocatively, “If I promise you that I will help to take care of your child this afternoon, I have acquired a duty *regarding* your child, but no duty *to* your child. My duty is *to* you, the one to whom I have made the promise.”⁴⁰ For Johnson, obligations *to* another being are a matter of being constrained by the will of another, specifically because of the interdependence of humans that we saw above (§1.2—that we are imperfect beings and require the assistance of others in achieving our ends). This interdependence creates a relationship between wills which implies reciprocal constraints.⁴¹ Thus, though our duties *to* another can only arise by reciprocal constraints and out of respect for the humanity of the person to whom you are obligated, duties *with regard to* can reference just about anything, but only as an extension of one’s direct obligation *to* another person or *to* oneself, and not as a direct obligation *to* that thing. As in Denis’s reading, when one thinks one has a duty to a non-human animal, in truth one’s duty is to oneself, but with regard to the animal (i.e. to not treat it cruelly). So it is no surprise that some might worry that whatever duties we have regarding non-human animals might collapse under pressure from practical human concerns—as they are, in fact, centered on humans and not at all on the non-human animal.

Thus, there are two notions at work here in qualifying a being as capable of being obligated *to*: reciprocal constraints, in terms of interdependence of wills, and reference to the cognitive capacities of the being—i.e. that it must possess humanity, or human dignity. However, though Johnson’s account is correct regarding the directionality of any single duty, it may not exhaust the complexities of our obligations in light of Kant’s theory and our moral intuitions.

³⁹ Johnson, “Duties to and Regarding Others,” 204.

⁴⁰ *Ibid.*, 203, emphasis in text.

⁴¹ *Ibid.*, 195-202.

First, it must be noted that on this account there are actually two duties incurred: one *to* another being and one *with regard to* something else—which if I fall for the amphiboly I mistakenly take as a duty *to* that something else, and if I do not fall for the amphiboly I correctly take as a duty *to* myself. Moreover, these two duties are incurred in addition to all the other perfect and imperfect duties which are already present, creating a complex web of duties.⁴² Second, Johnson’s account does not explicitly consider that one might have additional constraints on being obligated beyond interdependence and reciprocal constraint. For instance, what we are obligated to do cannot itself be unethical. On one hand, this simply means that any obligation I have to someone else cannot violate the perfect and imperfect duties I have to myself or to any other beings. Thus, when coupled with the first point a deeper and more complex interplay between duties and reciprocal constraints becomes apparent—one in which further duties and constraints may emerge and restrict our fulfillment of duties. On the other hand, this interdependence itself makes reference to the cognitive capacities of the being we are obligated to—i.e. they must possess humanity and they must have ethical ends. But therein lies the problem.

As we will see, the ability to set ends requires the same cognitive capacities which Kant denies to non-human animals, such that if he is wrong to deny those capacities (and I believe he is) then we may have duties *to* non-human animals. In this case, ejecting Kant’s anachronistic views regarding the nature of non-human animal mental life provides an easy fix to the project of providing a model of moral considerability which is based upon the cognitive life of a being and places humans and non-human animals on the same evaluative scale. On the other hand, if one must not only be able to set ends, but to set *ethical* ends, then Kant’s view of moral

⁴² In fact, to be accurate the duty which I incur *to* another is a result of perfect and imperfect duties I already have—namely, in addition to my perfect duties to respect myself and others, I have imperfect duties to cultivate my virtues and, in so doing, to take on the ends of others. These imperfect duties are the basis of the duty *to* which is incurred. So, to complicate matters, my duty *to* is actually a duty *to* myself *with regard to* another.

considerability based on cognitive capacities may be too restrictive and eliminate obligations to most people—as setting ethical ends requires cognitive capacities developed to advanced degrees which many humans either do not possess or do not utilize, including children.

Such a restrictive view would be entirely alien to our moral intuitions. Consider the differences between housesitting, babysitting, and dog-sitting. If I agree to housesit for my neighbor, I now have an obligation *to* my neighbor *with regard to* her house, as well as a duty to myself to fulfill that duty. As the house is an inanimate object, I have no duty directly to the house, though I do have a duty *to* my neighbor to keep her house in good condition. But, does it make sense to say that in agreeing to babysit my neighbor's child I incur no duty *to* the child, but only *with regard to* the child? Certainly my duty *to* my neighbor can be understood this way, but in light of the fact that keeping a house in good condition and keeping a child in good condition involve vastly different things, it would appear possible that I might incur additional obligations when beings with certain, even underdeveloped, capacities are involved. Similar obligations may arise out of dog-sitting, as one is not just obligated to insure the dog or child does not die, but also that it is happy and flourishes.⁴³ But, in watching the dog or child, how is keeping that being happy not an additional direct obligation *to* that being, but, rather, only an obligation *with regard to* that being? Moreover, is it really an amphiboly (a mistaken logical equivocation) to think we have, and ought to have, such obligations? I assert that our moral intuitions would suggest that in babysitting or dog-sitting we also incur a direct obligation *to* the child or dog, the fulfillment of which is a necessary means to fulfill our obligations *to* the neighbor. For instance, if one can at

⁴³ This, of course, assumes that your neighbor's request is ethical—e.g. that she does not want you to burn her house down while she is away so that she can get the insurance money, or that she does not secretly want you to harm her child or her dog by, for instance, feeding them poison.

all say that my neighbor has a duty to her own child, then one can see that I might incur that same obligation as a surrogate in the case of babysitting.

Of course, one may object that the child or, especially, the dog have no rational will (i.e. humanity) and thus, I cannot be obligated directly *to* them. Or, perhaps one may object that the child (and maybe the dog) have humanity, but to an insufficient degree for us to have obligations directly *to* them. But, this is exactly what we need clarified. Recall that on Denis's reading, Johnson's reading, and Kant's theory we have no direct duties to other beings unless they meet the two criteria of reciprocal constraints and the right cognitive capacities (i.e. humanity). Throughout the rest of this chapter I will put pressure on these two notions in various ways in order to reveal that we need a better account of the mental capacities of beings than Kant has provided if we are to understand when we have obligations to other beings and to which beings we can have obligations. Further, rejecting certain anachronisms in Kant's view of non-human animals will reveal that his account of obligations *to* a being ought to include non-human animals as beings we can have direct obligations to if they do, in fact, possess the right cognitive capacities.

§1.5 The Dilemma of Moral Considerability

By now it should be clear that cognitive capacities play an important role in Kant's ethics, but that to ground a theory of moral considerability which includes non-human animals will require some work. Further, there is a building sense of mystery regarding which beings possess the cognitive capacities that constitute humanity and dignity, such that we have direct obligations to those beings. To begin resolving these issues, I suggest that a dilemma arises if we base moral considerability on cognitive capacities. However, like with most dilemmas in philosophy, the

idea here is to show that a middle position between two undesirable extremes is where our considerations ought to settle. As a formulation of the extremes of this dilemma I propose:

The Dilemma of Moral Considerability: any being that is worthy of moral consideration is either A) worthy entirely in virtue of the capacities of the being that is obligated to it, or B) worthy entirely by traits it itself has that create the obligation in others.

In the case of A, my own cognitive capacities are the source of my moral obligations.

This makes sense, as the initial source of obligation on Kant's theory is our own faculty of reason coupled with the presence of our will. In other words, as we saw in §1.2, the fact that we see ourselves as beings with free will generates the thought that there are obligations for us in directing our will. Moreover, for Kant, for an action to be ethical it must not only conform with duty, but must be done for the sake of the moral law. This places primacy on my actions, my maxims, and my reasons for acting—thus, on my cognitive character. This is captured in Kant's famous formulations of the Categorical Imperative (the supreme moral law), which not only suggests how we ought to direct our will, but also how we ought to treat other beings: as a universal law, "*act as if the maxim of your action were to become by your will a universal law of nature;*" as how to treat humanity, "*act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*"⁴⁴

Additionally, this means that only moral agents (rational beings with autonomous wills who act in accordance with the Categorical Imperative) can have obligations—or, more accurately, that only rational beings, beings with humanity, have obligations to be moral. This is because it is their faculty of reason which represents the moral law along with representing possible actions.

However, no part of this horn (A) references what beings a moral agent is obligated to. For Kant, it does reference that there is some quality of beings called humanity and that it ought

⁴⁴ Kant, *GMS*, 4:402 and 4:429.

to be respected, but it does not tell me what exactly constitutes humanity such that I can recognize it in others. Part of this issue arises from the epistemic limits (or limits of knowledge) Kant places on our cognitive powers. Specifically, we cannot perfectly cognize ourselves, and, thus, we cannot know that we are free in the exercise of our will; moreover, I cannot know that any other being is free either.⁴⁵ In modern philosophical terms this is much like the problem of other minds—i.e. how can I know that any other being not only has a mind, but has a mind like mine, when I have no access to their minds, only my own? Because I cannot know this, I also cannot know for sure which beings are persons (i.e. possess humanity) and which beings are not. Perhaps the best I can do is assume that persons must be like me—but in what will that consist? Given that I do not perfectly cognize myself, nor do I have direct access to the cognitive life of another being, if my obligations are a matter of my mental constitution, how am I to extrapolate this to others? However, I could, perhaps, choose to treat all beings as if they possess humanity. Thus, on this horn, there is no reason to exclude non-human animals from moral considerableness.

The other horn (B) seems accurate historically, as Kant denied that we have duties to non-human animals, specifically in reference to their lack of the right cognitive capacities. Further, we intuitively seem to think that we ought to be obligated only to beings that are also obligated to us—hence the respect for humanity captured in our perfect duties to others, as well as the qualifications present in Johnson’s reading of duties *to* versus duties *with regard to* (§1.4). Nothing seems irrational about limiting moral duties to moral agents, and having moral duties to

⁴⁵ As Kant says it, “I cannot cognize freedom as a property of any being to which I ascribe effects in the world of sense, because then I would have to cognize such an existence as determined, and yet not as determined... nevertheless, I can think freedom to myself,” Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998) Bxxviii. Hereafter, *CPR*. Also, the pagination of the *CPR* varies slightly in the Academy edition from Kant’s other works, in that its format is edition (A or B) followed by section number. I have chosen to follow that pagination here. See also, Bxxix, and B562. See also, Kant, *GMS*, 4:456 and 4:459.

moral agents seems to always pass the formulas of the Categorical Imperative as a test for the morality of an action.⁴⁶ However, this means that only certain beings with certain capacities count as moral agents, and, therefore, as morally considerable. This also means we need an account of what capacities are required, and to what degree they are required for moral considerableness (i.e. the necessary and sufficient conditions of moral considerableness). So, if we accept B, then we must ask what traits a being must have in order to be worthy of direct moral consideration, and, more importantly, to what degree one must have them. On horn A this was a matter of recognition—how do I recognize that the capacities in me, which make me moral, exist in another being—but, on horn B this is a matter of definition coupled with recognition—i.e. I can define what constitutes a moral being separately from what I am, but still must be able to recognize it in another being.

Thus, there are epistemic issues on horn B as well. Since we cannot perfectly cognize ourselves our own motives are often not clear to us. We may not ever be able to know exactly why we have acted the way we have, nor must we always have acted from a conscious maxim.⁴⁷ In fact, as Kant says, “it is absolutely impossible by means of experience to make out with complete certainty a single case in which the maxim of an action...rested simply on moral grounds...for we like to flatter ourselves by falsely attributing to ourselves a nobler motive.”⁴⁸

What this means is that if we are obligated to another entirely due to its moral agency, including

⁴⁶ What I mean here is that if a being is moral, then I have no reason to not help him in his ends, i.e. it never leads to a contradiction to do so. For more on the Categorical Imperative as a test of the maxims guiding our actions, see Onora O’Neill, “Consistency in Action,” in *Ethical Theory 2: Theories About How We Should Live*, ed. James Rachels (Oxford: Oxford University Press, 1998), 260-261. As O’Neill develops in her reading, the Categorical Imperative, specifically the Formula of Universal Law, presents a test for our maxims—a test that we need not always apply, but that our actions must always pass if they are to be considered morally worthy.

⁴⁷ Cf. O’Neill, “Consistency in Action.”

⁴⁸ Kant, *GMS*, 4:407.

the right cognitive capacities, and we can never be sure about the maxims which guide the actions of ourselves or others, then we can never be sure that an action is ethical and, thus, that a being is moral, despite the fact that we have a definition of what is moral. As Onora O’Neill notes, the Categorical Imperative, as a test of the morality of one’s action, “is a test that agents must seek to apply to their own proposals for action” since only they can adopt maxims for their actions, and, thus, “the test is one that we can propose to but not impose upon” other agents (as we cannot determine their maxims for them).⁴⁹ To put this another way, if a moral being is a rational being *and* an autonomous being, then its freedom, coupled with my epistemic limits, means it can always act unethically.⁵⁰ Thus, if obligation is a matter of the capacities of another being, I can never know what beings actually have and deploy those capacities, and, therefore, I cannot know that I have obligations to that being.

There is a further issue with horn B, though, which stems from the criteria of reciprocal constraints seen in Johnson’s reading of duties to (§1.4) and any attempt to suggest that we have moral duties only to moral agents. We can see this through a thought experiment. Suppose that Jones is a perfectly rational moral agent—he is as perfectly morally virtuous as one human can possibly get. Smith is also a moral agent and therefore can have direct moral obligations to Jones. In a freak accident, Jones gets hit with a rock that falls out of the sky and causes a severe mental handicap such that Jones loses his ability to be a moral agent. It would certainly go against most of our moral intuitions if we held that Smith is suddenly absolved of all possible direct moral

⁴⁹ O’Neill, “Consistency in Action,” 264-265.

⁵⁰ Steve Naragon raises an interesting point about this in that, in Kant’s system, only rational beings can be autonomous, but it is possible for a being to be rational and yet not autonomous. This makes sense given the epistemic limit that human autonomy, in the metaphysical sense, cannot be known. See Steve Naragon, “Reason and Animals: Descartes, Kant, and Mead on the Place of Humans in Nature” (PhD diss., University of Notre Dame, 1987) 53, re-formatted 190 p draft, accessed October 20, 2013, <http://users.manchester.edu/facstaff/ssnaragon/Naragon/Research.html>. Hereafter, “Reason and Animals.”

obligation to Jones—i.e. that any duties he may previously have had to Jones are now just imperfect duties to himself because Jones now has less value than what is required for direct moral consideration. Clearly, Smith ought to have some direct moral obligations to Jones still, but if we settle for non-cruelty then we only get obligations that Smith cannot do whatever he wants to Jones, but no obligations that Smith must respect and aid Jones in whatever ends he might still have (such as living a life as dignified as he can). As part of our moral intuitions, we do still have some positive obligations (i.e. not just non-cruelty obligations) to people like Jones—for example, the handicapped, the infirm, or the senile. Perhaps we recognize that they still have some dignity and that we must respect it.⁵¹ Of course, the difficulties of definition and recognition on horn B still apply here. If our definition includes beings like Jones, then though the epistemic constraints might be loosened (i.e. it becomes easier to identify the class of beings we have obligations to, because that class has been enlarged), the line establishing what cognitive capacities are needed for moral obligation becomes blurred, because it now includes not only beings who are, putatively, perfectly rational, but also beings that are only possibly or minimally rational.

Alternatively, we might take a middle ground between A and B, acknowledging that the source of obligation is within the moral agent, but certain features of the being we are obligated to qualify it for moral consideration in ways that other beings are not so qualified. This is the view I am arguing for. Further, we can relax the epistemic limits by letting A and B work together—allowing for a sort of practical catch all in that I have moral obligations because of my

⁵¹ In fact, Kant holds that we must treat criminals, and even the most vile and vicious human beings, with respect despite the fact that their deeds make them unworthy of it (see *MS*, 6:463). However, the difference here is that the guilty criminal who faces the death penalty has the capacity to reason but, for whatever reason, is not doing so—is not living up to his obligation to do so given the type of being that he is—whereas Jones has lost the ability to reason through no fault of his own. This raises an interesting question as to why we would place more value on having reason and failing to use it, than we do on having had reason and then lost it entirely.

capacities, and the capacities I see in others can help act as an intuitive mirror for detecting beings like me, to which I have, or can have, ethical obligations. In terms of the Jones case above, this would allow my duties to Jones to be preserved not because of what sort of being Jones is (though that does matter), but because of what sort of being I am. Further, as we will see in §1.7 (and, more thoroughly, in Chapter Two), Kant suggests that what capacities non-human animals do have are analogues to what humans have. To follow suit, we can say that beings to whom I can be obligated have capacities analogous to my own—for the most part, these will be human beings, and indeed this is how many see Kant as answering the problem of other minds. However, on my view this will include non-human animals based on their cognitive lives. To see why, we need to return to the question of what traits are necessary, and what degrees are sufficient, for moral consideration, which the dilemma of moral consideration highlights. In other words, we must ask what traits a being must have in order to be worthy of moral consideration, that are common to me and to it, and, more importantly, to what degree one must have them.

§1.6 Setting the Bar for Cognitive Capacities

In the discussion so far, I have repeatedly referenced that a being must have humanity in order to qualify as an object of our respect, and, thus, as a being we can be obligated to. Now, I would like to explicitly focus on what constitutes humanity as a set of cognitive capacities which qualify a being for moral consideration in themselves (i.e. directly). As noted above, for Kant, humanity cannot merely be comprised of being a human being, but instead must be a quality which human beings possess. This is because though what we respect in others, and in ourselves, is humanity, if the moral law is universal it “must hold not only for human beings but for all *rational beings as such*” and, thus, must be derived “from the universal concept of a rational

being as such” and not from a “*special property of human nature*.”⁵² Moreover, though the representation of the moral law can “*occur only in a rational being*,” if the moral law is universal it “does not hold only for human beings, as if other rational beings did not have to heed it,” but must hold for all rational beings.⁵³ For Kant, the category of rational beings includes aliens, angels, and spirits—to which we cannot have duties solely because “it would have to be possible for the human understanding to gain knowledge of these spirits, their nature, existence and constitution,”⁵⁴ but who, apparently, are more perfect in nature in direct proportion to how far they are from the Sun.⁵⁵ Of course, we may wonder, given the epistemic issues noted in the previous section, how we can know that non-human animals do or do not have rational capacities, yet commit to the claim that humans have them, as well as possibly angels, aliens, and spirits—which we have never encountered. I will return to this issue in the next section, but, what should be clear now is that the population of rational beings does not just include humans.

So why not non-human animals? As previously noted, apparently, non-human animals do not have the right capacities. Reason and autonomy are the most likely candidates for these capacities, but neither is straight-forward on Kant’s account. In fact, Kant’s account of autonomy is somewhat confusing, for though the animal part of nature has choice, animals are also “necessitated [by stimulus], so that a dog must eat if he is hungry and has something in front of him; but man, in the same situation, can restrain himself.”⁵⁶ Additionally, as we have already

⁵² Kant, *GMS*, 4:408, 4:412, and 4:425, respectively.

⁵³ *Ibid.*, 4:401, and 4:389, respectively.

⁵⁴ Immanuel Kant, “Kant on the Metaphysics of Morals: Vigilantius’s Lecture Notes,” in *Lectures on Ethics*, ed. Peter Heath and J.B. Schneewind, trans. Peter Heath (Cambridge: Cambridge University Press, 1997) 27:711. Hereafter, *LE (Vigilantius)*. See also *LE (Collins)*, 27:459-460, and *MS*, 6:442-444.

⁵⁵ Cf. Kant, “Universal Natural History and Theory of the Heavens,” 1:360.

⁵⁶ Kant, *LE (Collins)*, 27:267 and, on choice, *MS*, 6:442.

seen, the epistemic commitments of Kant's philosophy include that we cannot know that we are autonomous, yet we nevertheless see ourselves as autonomous. Seeing ourselves this way is a product of our reason, in that reason provides the rules by which our cognitions of ourselves are constructed by our cognitive faculties, and our imperfect cognitions of ourselves represent our will as being free. Thus, reason and autonomy are tied together.

In fact, it is reason that allows humans to have motives and to set ends, and it is the “capacity to set oneself an end—any end whatsoever—[which] characterizes humanity (as distinguished from animality).”⁵⁷ Yet, again, the epistemic limits discussed above indicate that we do not have very good access to our motives and intentions, let alone those of others. Further, earlier, in §1.4, I mentioned that there is a difference between setting ends and setting *ethical* ends in terms of what cognitive powers would be required. This is because the development of virtue depends upon the development of our capacity to reason. As Kant puts it, the duty to cultivate one's virtue can “consist only in *cultivating* one's faculties (or natural predispositions), the highest of which is *understanding*, the faculty of concepts and so too of those concepts that have to do with duty.”⁵⁸ As we saw above in §1.3, non-cruelty helps us cultivate these virtues. But, in §1.2 we saw that human nature is part animal and that reason allows us to rise above this nature, which suggests a developmental account of our rational capacities. According to Kant, every human being has the requisite predispositions for morality, including reason (from which we get the concept of duty), but those capacities must be developed for us to be ethical beings.⁵⁹ However, some of us never succeed in this, as “some human beings remain at [a] stage of

⁵⁷ Kant, *MS*, 6:392 and, on motives, *LE (Collins)*, 27:267.

⁵⁸ Kant, *MS*, 6:387.

⁵⁹ *Ibid.*, 6:399.

development” which does not rise above sensible desires and inclinations and into ethical ends.⁶⁰ So, clearly, being human is not equivalent to being ethical.

More importantly, the idea that we can cultivate our capacities implies that reason and autonomy admit of degrees. In fact, freedom does admit of degrees for Kant, as “the less a human being can be constrained by natural means and the more he can be constrained morally (through the mere representation of duty), so much the more free he is.”⁶¹ So, though setting ends may allow us to be ethical, it is by setting *ethical* ends that we become more free. In this sense, though humanity is distinguished by the ability to set ends, this humanity is also increased in us as we become more virtuous. Because of this, I suggest there is a more necessary foundation of setting ends—self-consciousness—which allows us to be aware of our ends and motives and constrain them with reason toward virtuous, ethical ends. Thus, to set ends a being must meet at least three necessary criteria: it must be self-conscious, rational, and autonomous—one cannot set ends if one is not free to do so, one cannot act from the moral law unless it is represented to oneself by reason first, and one cannot do either without being self-consciously aware of oneself as a being with a will, capable of taking action in the world. Thus, it is these three criteria that comprise our ability to set ends, as well as our ability to cultivate the ends we set—thus, these three criteria at least partially comprise humanity, such that a being which has them is a good candidate for a being which possesses humanity and, thus, moral considerability.

However, as I will discuss in Chapter Two, self-consciousness also admits of degrees for Kant.⁶² Since autonomy, reason, and self-consciousness all admit of degrees, comprise our

⁶⁰ Kant, “Universal Natural History and Theory of the Heavens,” 1:356. Also, see *MS*, 6:387.

⁶¹ Kant, *MS*, 6:382.

⁶² See Kant, *CPR*, B415.

humanity and our ability to set ends, and must be developed for us to be ethical beings, we ought to wonder what degrees of these capacities are necessary for direct moral consideration. When does one develop the capacity to be ethical, or to have moral worth? Frustratingly, Kant does not say much about these degrees, nor how much our inborn capacities must be developed before we become morally worthy. This leaves us to speculate on where this threshold must be.

On the one hand, Kant licenses us to assume that happiness is an end of all rational beings, but he also says that “nature would have hit upon a very bad arrangement in selecting the reason of the creature to carry out this purpose,” as “we find that the more a cultivated reason purposely occupies itself with enjoyment of life and with happiness, so much the further does one get away from true satisfaction.”⁶³ In other words, reason is a terrible guide to happiness when compared to instinct, but, nevertheless, it is required for cultivating a good will, which is the measure of our worthiness to be happy.⁶⁴ Moreover, not only do humans not reason perfectly, reason often leads us to mistakes and illusion—a problem that Kant spends incredible amounts of time in his *Critique of Pure Reason* discussing in terms of how we can reign in our reason and avoid error. In this sense, we might think, then, that a lower degree of rationality would be more conducive to happiness and properly cultivating our capacities.

For instance, as Kant admits, it is doubtful “whether any true virtue is to be found in the world” and “there may never have been a sincere friend.”⁶⁵ However, my childhood friend Richard has severe mental handicaps, but is also the most loyal friend I have ever had, specifically because he is too handicapped to ever think of being devious, dishonest,

⁶³ Kant, *GMS*, 4:395. On happiness as an end, see 4:415.

⁶⁴ *Ibid.*, 4:396.

⁶⁵ *Ibid.*, 4:407-408.

backstabbing, or anything other than genuine. Insofar as a good will is one which is virtuous in what it wills, perhaps what Richard wills is more pure and virtuous in its motives than the vast majority of people. A similar notion applies to companion animals such as dogs, who are often referred to as “man’s best friend” and, assuming they are capable of happiness, do not have reason interfering with their happiness and the happiness of others. Perhaps the most loyal friends, and, perhaps, the most moral beings, then, are the ones who are simple enough in their cognitive capacity to not be misguided by reason.

Of course, one may say that neither Richard nor an animal companion (nor Jones from the thought experiment in §1.5) knows any better and, therefore, their friendship has no moral content as they must be friendly or moral from respect for the moral law and not from any other considerations—such as needing help or food and shelter. How much one must think about duty in acting morally is actually a somewhat disputed feature of Kant’s ethics.⁶⁶ But, perhaps we can say that at least one must be capable of thinking this way, even if they don’t actually do so. So, perhaps we can say that to be a moral being one must be capable in this way, but, like in the first horn of the dilemma of moral considerability, this does not tell us anything about whether or not we might have moral obligations *to* someone who does not possess this capacity, just that one must have this capacity in order to have obligations to others. The issue here, though, is in regard to what beings can be counted as directly morally considerable—which is a classic problem for ethical theories, as if one counts children, the handicapped, the infirm, the senile, or beings like Richard or Jones as morally considerable based on the presence of developed cognitive capacities, then one must reconcile why we cannot consider non-human animals, which often exhibit more developed cognitive capacities than some humans, as morally considerable.

⁶⁶ Cf. O’Neill, “Consistency in Action.”

On the other hand, one might reasonably say that in order to be directly morally considerable one must not only be capable of thinking that there are right and wrong ways to act, but also that one must understand the principle of a good will—the Categorical Imperative. However, the vast philosophical literature that attempts to explicate Kant’s ethics and the various formulations of the Categorical Imperative indicates that these elements of Kant’s ethics are extremely difficult to understand—even for the best and brightest philosophers. This, then, restricts the amount of moral beings that have ever existed to such a low number that Kant’s ethics would be virtually useless for “normal” people, which comprise the vast majority of humans that do, have ever, or will ever exist. In other words, Kant’s ethics would be useless if we only have duties to others that actually have a good will (especially since humans are imperfectly rational)—since cultivating a good will requires actually applying rationality, and the vast majority of beings with potentially good wills may not have understood how to apply reason properly to the will—as evidenced by the fact that we even need a moral theory at all.

What these considerations ought to show, then, is another middle ground. We cannot too tightly restrict what beings we can have direct moral obligations to, or else the moral theory becomes useless, as it applies to next to no one. However, if we loosen constraints too much, we have no reason to not include non-human animals as objects of direct moral duties. With the dilemma of moral considerability and the Jones thought experiment, we see that we might have obligations to others regardless of whether or not they can or do fulfill their obligations to us—Kant’s is not a *quid pro quo* theory of reciprocity.

And so, the discussion so far has led to this conclusion: the source of moral obligation is intrinsic to the moral agent, but what beings an agent is obligated to is a matter of meeting a certain threshold of reason (cognitive capacity) and autonomy (freedom to set one’s ends) that

has yet to be established. There must be some minimal conditions—such as being able to reason at all or being able to direct one’s will with some degree of freedom based upon one’s degree of ability to reason. But, we have no account of when this develops in a being. What should be clear here is that on Kant’s account there is a deep divide between whatever cognitive faculties non-human animals possess and those which humans are said to possess, despite the fact that Kant sees humans as part animal. Some of this divide is accounted for by the ability in humans to develop their inborn capacities through autonomy, reason, and self-consciousness—which allow humans to rise above and cultivate their virtues and ends. Self-consciousness, autonomy, and reason are necessary conditions of setting ends—thus, perhaps a degree of these three yield the ability to set ends and the presence of these three can indicate such an ability even in non-human animals. The question, then, becomes whether or not non-human animals do, in fact, possess degrees of these cognitive conditions. That will be the goal of Chapters Three and Four. However, it should be clear that if they do, then they also possess the necessary conditions for direct moral obligations to them, rather than merely with regard to them.

§1.7 Anachronism, Religiosity, and Speciesism

Before moving on to Chapter Two, it is worthwhile to note, as a way of addressing possible counterarguments, why Kant might further resist granting any degree of self-consciousness, reason, and autonomy to non-human animals. The simple, relevant answer is this: that to do so, in Kant’s era, would threaten to undermine the ethical system that Kant is trying to develop. In other words, there must be a difference in kind between humans and non-human animals or else one can easily push back on his ethical theory by saying “Why should I not just be like other beings in nature? Why should I not just let my desires rule me? Why should I not

use my reason and autonomy merely to further my own ends?” Kant’s answer is, of course, that we have dignity and humanity in terms of respecting our rationality and autonomy, and we are limited beings who need assistance in achieving our ends—both of which call us to be consistent and to universalize our maxims in ways which account for recognizing the value of other beings like us. Thus, if non-human animals have rationality and autonomy, they would become part of this universalization and one might wonder why we are called to be any better than them, when it seems relatively clear that they do not bind themselves to the same moral law. On the other hand, with my argument we may be able to find a relevant difference in degree, rather than kind, and still uphold the call to be ethical. To do this though, we must eject two anachronisms which support Kant’s resistance to granting rationality and autonomy to non-human animals: a religious teleological view, and an outdated scientific view based on the principle of parsimony.

We have already seen, in the previous sections, Kant’s commitment to a religious teleological view—that is, that humans have a special and unique purpose. For one, humans are part animal and “the human being has a duty to cultivate the crude predispositions of his nature, by which the animal is first raised into the human being.”⁶⁷ This suggests that what we actually respect is the ability to rise above—a sort of teleological view that was historically popular, due especially to its tie to religious thought. In other words, we often fall into vice when we let our “animal” desires and inclinations rule over our actions, and we become more holy, or virtuous, through developing self-restraint—we rise out of sin and toward the perfection of our human nature by acting less like beasts. Interestingly, though Kant believes non-human animals have souls, he denies non-human animals the development of virtue which allows them to rise above

⁶⁷ To reference the same passages as used in §1.2, see Kant, *MS*, 6:387—“raise himself from the crude state of his nature, from his animality...more and more toward humanity...so that he may be worthy of the humanity that dwells within him” and 6:392—“the human being has a duty to cultivate the crude predispositions of his nature, by which the animal is first raised into the human being.”

aspects of their nature. For instance, Kant says that an animal soul is a “brute soul [and] can develop itself to infinity, grow, but always only sensitively, never up to a rational being”⁶⁸ and they will “never become human souls, because they differ from human beings not merely by degree... but by kind.”⁶⁹ This fundamental difference between humans and non-human animals is also represented in the view that human dignity is “of a worth that has no price,”⁷⁰ or that human life is beyond price while animal lives are not, due to their lack of the reason and autonomy which are unique to humans among known species. Thus, as in the discussion of humanity and dignity above, much like religious teleology, there is a special nature, a special purpose, to human life in its call to “rise above” and achieve the ideal it was meant for.

We also saw other religious notions in Kant’s mention of extraterrestrials, angels, and spirits—beings which, on Kant’s view, are more perfect than humans. Though he admits we have never encountered such beings, his willingness to entertain their existence, along with their degree of perfection above that of humans, is reminiscent of the Great Chain of Being—a religious hierarchy of beings with the most perfect or spiritual being, God, at the top followed by the types of angels and celestial bodies, then humans, then animals, then plants, and finally rocks, minerals, and non-beings at the bottom. Indeed, in the *Lectures on Metaphysics*, Kant speculates that non-human animal souls might “persist for eternity and also stand in service to human beings in the next life as in this one”⁷¹—a view reminiscent of the Biblical view that grants humans dominion over animals. Additionally, God and the afterlife play important roles in

⁶⁸ Immanuel Kant, “*Metaphysik Dohna*,” in *Lectures on Metaphysics*, ed. and trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997) 28:690. Hereafter, *LM (Dohna)*.

⁶⁹ Immanuel Kant, “*Metaphysik Mrongovius*,” in *Lectures on Metaphysics*, ed. and trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997) 29:906. Hereafter, *LM (Mrongovius)*.

⁷⁰ Kant, *MS*, 6:462.

⁷¹ Kant, *LM (Mrongovius)*, 29:906.

Kant's ethics in the forms of a distributor of justice and reward, and a life in which we might actually achieve the perfection we strive for.⁷² So, clearly there are religious aspects to Kant's view, and it is relatively easy to speculate that if Kant had suggested that there is not a difference in kind between non-human animals and humans—no special characteristic allowing mankind to “rise above”—then Kant's call to be ethical may appear undermined, or at least challenged in ways that Kant may not have wished to take on.

To be fair, however, Kant does not entirely rely on these notions. Human perfection, as an end or aim, is also a result of how a human sees himself. Recall that we imperfectly cognize ourselves on Kant's view, and a robust reading of Kant's ethics can still be had by connecting the call to be ethical, or to cultivate virtue, exclusively to the nature of human beings as perceived by the limits of human cognition. We see that we can be better, we see that things ought to have happened that did not happen, and that we are free to guide our will towards bringing about what we think ought to happen. Fully providing this reading is beyond the scope of this thesis. But, the point here is that the religious aspects of Kant's view reveal a vested interest in the special nature and purpose of human dignity, but we need not follow Kant on these anachronistic views.

For one, even if we grant that an ethical theory must rely on God and an afterlife for incentive and motivation, we need not also embrace religious views that place non-human animals as subservient to humans, nor that our animal nature is somehow sinful. Indeed, as we saw in §1.3, we violate our respect for our own humanity if we fail to recognize the value of our animal nature. This raises an important question: if we must respect our own animal nature, why would we not respect animal nature in non-humans, even if that is all there is to them? Further,

⁷² Cf. Kant, *CPR*, especially B841—“Thus without a God and a world that is now not visible to us but is hoped for, the majestic ideas of morality are, to be sure, objects of approbation and admiration but not incentives for resolve and realization...”

we need not cut non-human animals off from progression in cognitive capacities, at least into beings with a certain degree of rationality and freedom. Perhaps it is too much to think that non-human animals can “rise above” their nature and, somehow, become human in some relevant sense, but we can accept a lesser notion—the sort of psychological continuity of traits that Darwin and Romanes thought was likely given the phylogenetic continuity of biological traits. In other words, we need not embrace a sort of human exceptionalism, or speciesism, that humans are unique among known species and that their interests are, therefore, the only ones worthy of moral consideration. Instead, we may embrace an evolutionary account that suggests that some non-human animals do have some degree of reason, autonomy, and self-consciousness in their cognitive lives, and thus, more worthiness of moral consideration.

Second, we need not feel hesitant to abandon anachronistic aspects of Kant’s view, as long as we do so mindfully. Kant himself embraced such methods when he said, “it is a foolish mistake to attribute preeminence in talents and good will to the ancients in preference to the moderns just because of their antiquity.”⁷³ Indeed, if accepting Kant’s view at all means we are stuck with all aspects of Kant’s view, then we would need to accept his blatant anti-homosexuality, his stance against pre-marital sex, and his view that selling one’s tooth is a way “of partially murdering oneself.”⁷⁴ Additionally, we would need to endorse the blatant racism evidenced in his views that “Humanity has its highest degree of perfection in the white race,” that “The inhabitant of the temperate zone, especially in its central part, is more beautiful in body, harder working, more witty, more moderate in his passions, and more sensible than any other

⁷³ Kant, *MS*, 6:456.

⁷⁴ *Ibid.*, 6:424-425 and 423, respectively.

kind of people in the world,” and that the Hottentots sometimes eat worn-out shoe soles.⁷⁵ Most modern Kantians would eject these views without much argument, as they are inessential to Kant’s more substantive views. Though ejecting all religious or teleological notions from Kant’s views would require argument, we have little reason to follow Kant on his religious distinction between humans and non-human animals.

There is, however, a more problematic anachronism. In various passages, Kant denies that non-human animals possess consciousness, understanding, apperception, or judgment.⁷⁶ However, in grounding our duties *with regard to* animals (or our imperfect duties *to* ourselves to not be cruel to animals), Kant says in his *Lectures on Ethics* that “animals are an analogue of humanity... the acts of animals arise out of the same *principium* from which human actions spring, and the animal actions are analogues of this”⁷⁷ and, in the *Lectures on Metaphysics*, “The principle that guides animals, as analogue of reason... is called instinct, the faculty for performing actions without consciousness, for which human beings require consciousness.”⁷⁸ What seems to be at work in this denial is a principle of parsimony, much like Occam’s Razor—we can explain non-human animal behavior by instinct, and, thus, need not attribute the more robust capacities of reason and autonomy to non-human animals. As Steve Naragon reconstructs Kant’s argument, “if we can adequately explain the behavior of brutes without attributing reason to them, then we should refrain from such an attribution on grounds of parsimony.”⁷⁹ As

⁷⁵ Immanuel Kant, *Physical Geography*, in *Natural Science*, ed. Eric Watkins, trans. Olaf Reinhardt, (Cambridge: Cambridge University Press, 2012) 9:316, 9:318, and 9:320, respectively.

⁷⁶ See, for instance, Kant, *LE (Collins)*, 27:267 and 459; *LE (Vigilantius)*, 27:710; *LM (Dohna)*, 28:690; *LM (Mrongovius)*, 29:907. Also see Naragon, “Reason and Animals,” 42.

⁷⁷ Kant, *LE (Collins)*, 27:459.

⁷⁸ Kant, *LM (Dohna)*, 28:690.

⁷⁹ Naragon, “Reason and Animals,” 48. Also see 43.

Naragon goes on to explain, this denial is further supported by observations that non-human animals lack language, “are easy to deceive, they do not improve their skills, and they are unaware of future time”—or, in other words, their overt behavior indicates a lack of the complexities that are clearly present in humans.⁸⁰

We see here a historical position that posits and maintains a gulf between the cognitive lives of humans and non-human animals, a view which is still prevalent today in various forms. What looks like rational behavior, or social behavior, or autonomy is deemed instinct—some rudimentary analogue of what is far more complex and developed in humans. More commonly, brains structures are used to explain human capacities and a lack of those structures is used to deny those capacities in non-human animals—relegating whatever remains unexplained in non-human animals as explicable by unthinking instinct. And, perhaps most commonly, though we must admit that some non-human animals clearly communicate with each other, their apparent lack of language is considered the ultimate barrier against granting them reason—as only with language could they have complex and abstract concepts. As the mind is complex, tricky, and mysterious, the skepticism present in each of these views may be warranted and they could, possibly, support the idea that humans and non-human animals are different in kind.

But, like the religious teleological view, we may have good reason to eject this argument from Kant in place of a better theory. Instinct, as a posited faculty, mostly serves as a negative barrier of empty explanatory value. There is no real description or explanation of what instinct actually is beyond an analogue of reason that often looks like reason, but isn’t. As an argument from parsimony then, Kant’s denial resists granting a more complex capacity that might explain the observable features of non-human animal life, by positing a mysterious and unexplained

⁸⁰ Ibid., 48.

faculty that does not really explain anything. As a historical analogue, this is somewhat like trying to hold onto the idea that other planets move in epicycles, because one does not want to accept that Earth's orbit is an ellipse rather than perfectly circular. As seen above, one may not want to admit that non-human animals can reason to any degree, because it might be perceived as making human reason less special. In other words, parsimony really should dictate that we ought to have the same scale of evaluation for all related things. Just as planetary orbits can be more or less eccentric, what looks like reason can be reason of greater or lesser degrees. Similarly, what looks like consciousness and self-consciousness is consciousness or self-consciousness, but these admit of degrees to. In other words, different degrees of the same kind of cognitive capacity, rather than different kinds, might actually allow us to explain more—about ourselves, about non-human animals, and about the relationship between the two.

However, completing this argument from parsimony against Kant's argument from parsimony requires the resources of the Chapters 3 and 4—the goal of which will be to develop a model for degrees of these cognitive capacities and show that at least some non-human animals can and ought to be considered rational agents on that model, with human rational agency being a difference in degree of development, not in kind. I believe this model yields far more explanatory value than instinct when it comes to explaining the cognitive lives of beings, and, indeed, yields a fecund framework for explaining even what does appear unique to humans, such as language and narrative concepts of the self. Further, if the arguments of this present chapter are correct, and on Kant's moral theory the cognitive life of another being determines whether we can be obligated to that being, then placing non-human animals on the same scale of cognitive capacities as humans may yield moral duties directly *to* non-human animals.

CHAPTER TWO
THE GROUNDWORK OF COGNITION

“Guy don’t need no sense to be a nice fella. Seems to me sometimes it jus’ works the other way around. Take a real smart guy and he ain’t hardly ever a nice fella.”

--John Steinbeck, *Of Mice and Men*

§2.1 Introduction

We saw in the previous chapter that Kant’s ethical theory is fundamentally tied to his cognitive theory through its stress on the cognitive capacities of reason, autonomy, and self-consciousness as the ground of moral considerability. As I argued, if some non-human animals can be said to possess degrees of these capacities, then we ought to have duties directly to those non-human animals, in addition to the indirect duties of non-cruelty that we already have with regard to non-human animals. The difficulties of reliably ascribing cognitive capacities to other beings was noted in Chapter One, and establishing degrees of these capacities, as well as that some non-human animals possess them to morally relevant degrees, is the project of Chapters Three and Four and the model I will establish in those chapters. What is needed for that project is an account of the necessary conditions of these cognitive capacities, such that the grounds for my model and for ascriptions of capacities to other beings is provided. The most natural place to turn for this ground is, of course, Kant’s own theory of cognition.

As we will see in this chapter, Kant’s theory of cognition purports to establish the necessary conditions for the possibility of having the sort of spatiotemporal experiences that characterize human cognition. As such, I will first present a brief overview of Kant’s theory that reveals the roles of memory, conceptual understanding, and self-consciousness (or apperception) in human cognition, as well as the subjective features of phenomenal experience Kant’s account

establishes. As we will see, the elements of Kant's model are necessarily and comprehensively related such that no part works without the other components. These features and elements are the tools with which I will extend Kant's principles to construct both the base of my model and each subsequent level of it in Chapters Three and Four.

However, we will also see that Kant's theory jumps from describing cognition at necessary and minimal levels to describing full-blown human reason without accounting for any degrees or development in between. This jump brings into question whether Kant's theory is sufficient by itself, or if more is needed to flesh out his theory—both for his moral project, and for my own. Thus, in the second half of this chapter, I will place pressure on Kant's theory in terms of the absence of a developmental account of cognition. Further, I will place pressure on Kant's denials of crucial cognitive capacities to non-human animals and how these denials threaten to undermine the structure and argumentation of his theory, in terms of its necessary structure and universalizability. Pressuring Kant's theory in this way will reveal hints as to what Kant may have had in mind for a developmental model, and motivate the model I will construct in Chapters Three and Four based upon Kant's principles and arguments (though without his anachronistic denials). As I will argue, following Kant's clues and extending his principles to account for what is importantly lacking in his theory, will provide good reason to assume that non-human animals cognize much like humans do and, therefore, to ascribe the fundamental features of cognition to non-human animals. This ascription will also entail extending moral considerability to non-human animals based upon their cognitive lives.

§2.2 Kant's Theory of Cognition

Kant's theory of cognition is far too complex and comprehensive to cover all aspects of it in the room I have here, especially since the systematic interrelations of Kant's arguments are at least as important as what they establish. As such, I will focus on two central features of Kant's theory which are important for my purposes here: the Transcendental Synthesis of the Imagination (TSI) and the Transcendental Unity of Apperception (TUA), which is both an implication of the TSI and a vital component of it. Essentially, the TSI is the underlying constructive framework necessary for a conscious experience, and the TUA is the fact that a conscious experience necessarily relies on there being a conscious experiencer, or self (in some sense of the word). However, neither of these ideas are quite that straight-forward, so to assist understanding both I will focus on explaining the TSI in terms of the three cognitive faculties of sensibility, imagination, and understanding which contribute to it. On the way, I will explain the TUA and the important focus on memory, self-consciousness, and reason in Kant's theory.

From an epistemic standpoint, Kant's theory is an attempt to secure the foundations of human knowledge by explaining not only what experiential and cognitive faculties must be present for us to interface with the world, but also what limits, rules, and faculties must be present in order to make that interfacing possible. To establish this, Kant adopts what he calls a transcendental method—or, taking something that is known to be actual and uncontroversial and analyzing it to discover what the conditions of its possibility are *a priori*, or without reference to the particular experience itself.⁸¹ It is not controversial for Kant to assert that humans have experiences, nor that our experiences largely share similar features. Thus, if we can deduce what

⁸¹ Cf. Ralph C.S. Walker, "Kant and Transcendental Arguments," in *The Cambridge Companion to Kant and Modern Philosophy*, ed. Paul Guyer (Cambridge: Cambridge University Press, 2006).

features are necessary for the possibility of human experiences, then we have deduced something transcendental—something that must be the case, given that we have experiences, because it makes those experiences possible to have.

Further, for Kant, all experience is manifold and must be combined and organized by us via a synthesis: a bringing together of the pieces of the manifold into a cohesive experience. What this means is that our cognition works via mental representations (or constructions) of objects of experience, which are comprised of various perceptual inputs combined into one unified image. For instance, seeing a red chair not only involves perceiving the shape and constitution of the chair, but also the color and lighting and shading, each of which must be processed and united into one image. This same process of combination also explains cognitive error and illusion, such as seeing a straw or oar in water as broken—i.e. the object itself is not broken, but our representation of it is, due to how we perceive light differently through varying mediums such as air or water. Though this combination will largely take place unconsciously, it is an act of the mind and is, therefore, brought to experience by us. Thus, “This world is a represented world”⁸² and “we can represent nothing as combined in an object without having previously combined it ourselves.”⁸³

To get technical, the synthesis involved in our mental organization is a culmination of three distinct faculties of intuition separate from but propaedeutic to reason: sensibility,

⁸² Wilfrid Sellars, "...this I or He or It (The Thing) Which Thinks...," *Proceedings and Addresses of the American Philosophical Association* 44, (1970-1971): 7, accessed November 30, 2013, <http://www.jstor.org/stable/3129681>.

⁸³ Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998) B130. Hereafter, *CPR*. Also, the pagination of the *CPR* varies slightly in the Academy edition from Kant's other works, in that its format is edition (A or B) followed by section number. I have chosen to follow that pagination here. Further, all emphasis from Kant's writings should be assumed as his unless otherwise noted.

imagination, and understanding. Each faculty is itself a synthesis which contributes to the TSI, which is itself a three-fold synthesis of apprehension, reproduction, and recognition deriving what it synthesizes from the three faculties. Each part of the TSI works with inputs synthesized from each faculty, no faculty works entirely separate from the other two, and each part of the TSI is necessary for the resulting representation. To make this more clear, I will first overview the faculties in terms of the TSI and then briefly recapitulate the synthesis explicitly by connecting the pieces from the overview of the faculties.

Sensibility is the faculty for passively interfacing with objects of experience, or as Kant puts it, “the *capacity* to acquire representations through the way in which we are affected by objects.”⁸⁴ In some ways we can think of sensibility purely as perceptual apparatus and, thus, as physical, biological, or neuronal processes⁸⁵ much like, for example, those involved in color vision, where a wavelength of light which is not, properly speaking, colored becomes a color sensation “[derived] entirely from processing by the nervous system.”⁸⁶ The job of this apparatus is to passively sort inputs into manageable chunks, or impressions of an input located in space and time that can be worked with in the TSI to yield a conscious mental representation. Since this sortal work is already a kind of synthesis, it should be noted that sensibility does not function

⁸⁴ Kant, *CPR*, B33.

⁸⁵ I use this comparison only to aid in understanding. Though such a physicalist view of Kant might be partially correct, the trouble is that sensibility also grounds our perception of time and space, and for Kant time is part of our experiencing of inner sense, or introspection, as well as thought. Thus, it is not clear that a full biological reduction can be had on Kant’s view of sensibility. In other words, Kant seems to commit to a kind of dualism, though one not nearly as strong as Descartes’s. However, it seems an open question as to how important this dualism was to Kant and whether or not it might collapse under pressure from views such as eliminative materialism or a physicalist reductionism. Patricia Kitcher holds that no such reduction can be had due to the epistemic limits Kant places on his own transcendental psychology against speculative metaphysics—see Patricia Kitcher, *Kant’s Transcendental Psychology* (New York: Oxford University Press, 1990), 203-204. This debate is, however, outside the scope of this chapter.

⁸⁶ James Gordon and Israel Abramov, “Color Vision,” in *The Blackwell Handbook of Sensation and Perception*, ed. E. Bruce Goldstein (Malden, MA: Blackwell, 2001), 93.

entirely on its own (i.e. without the other faculties); instead, sensibility acts as a sort of passive filter of what could possibly be cognized from what cannot be. Since the world around us is filled with possible inputs, it would be overwhelming if they came all at once, unfiltered, as a “great blooming, buzzing confusion.”⁸⁷ Though some inputs, such as the pressure from molecules of air around us or wavelengths of light beyond the visible spectrum, may fail to register as inputs because our faculties are not sensitive enough, others, such as individual voices within a crowd, must be differentiated in order to register. This differentiation comes as impressions encoded with spatial and temporal information for use in the TSI.

The encoded spatial and temporal information is crucial, as we do not receive all of the impressions that will be combined into a mental representation at the same time. Instead, we receive them over time and these separate elements occurring at different times must be unified and represented at a single moment to the mind, by the mind.⁸⁸ For instance, at a brief glance this page contains words made of letters, each letter of which is perceived at a slightly different time and strung together in a sequence to form a word. Without this sequence, the words would be jumbled and nonsensical. Thus, in the TSI, elements that occurred moments prior (even nanoseconds prior) must be reproduced by the mind into a mental image in a single moment in the present, which requires a rudimentary sort of memory. As Kant puts it, “if I were to always

⁸⁷ This point should not be confused with attention, where one can pay more attention to some stimulus (bringing it into the foreground of their experience) over others. That will require the other faculties and the full synthesis of representation. The point here is that we do not experience everything around us at all times. For the quote, see William James, *The Principles of Psychology, Vol. 1* (New York: Cosimo, 2007), originally published in 1890, 462.

⁸⁸ Cf. Kitcher, *Kant's Transcendental Psychology*, 152.

lose the preceding [impressions]...from my thoughts and not reproduce them when I proceed to the following ones, then no whole representation...could ever arise.”⁸⁹

This reproduction is carried out by the imagination, or “the faculty for representing an object even without its presence in intuition.”⁹⁰ Thus, imagination is fundamental to the memory and reproduction involved in synthesis, where synthesis is “the mere effect of the imagination, of a blind though indispensable function of the soul, without which we would have no cognition at all, but of which we are seldom even conscious.”⁹¹ However, the reproduction could not be carried out without rules that govern it—rules that determine how to order the spatially and temporally encoded impressions to construct the cohesive representation in the present moment. In other words, the synthesis of reproduction over a series of moments which is required for a conscious experience requires a schema—a blueprint of sortal concepts for individuation and re-identification—and this schema comes from the categories of the understanding (more on these in a moment). Thus, the imagination is a bridge between sensibility and understanding—a conceptual territory wherein encoded impressions meet with a set of instructions to build a cohesive conscious representation out of the impressions as they continually unfold over time. Indeed, Kant says that part of the imagination belongs to sensibility and part to understanding, insofar as our intuition is sensible (receptive and passive) but our bringing forth of representations is spontaneous (active), and both spontaneity and receptivity are required for the

⁸⁹ Kant, *CPR*, A102.

⁹⁰ *Ibid.*, B151.

⁹¹ *Ibid.*, A78.

synthesis of the manifold into a representation by the TSI.⁹² Thus, imagination, reproduction of sensible impressions, and the memory required to do so play central roles in the TSI.

Finally, the understanding is “the faculty for bringing forth representations itself, or the spontaneity of cognition”; it is “the unity of apperception in relation to the synthesis of the imagination”; and it is “the faculty for thinking of objects of sensible intuition.”⁹³ This faculty gets the most attention in Kant’s arguments and contains most of Kant’s original contributions to philosophy of mind, as it includes the origins of self-consciousness as apperception, the rules of our cognition as deployed in the transcendental synthesis of the imagination, and the capacity for thinking beyond what is given in experience. As Kant says in a famous passage, “Thoughts without content are empty, intuitions without concepts are blind... The understanding is not capable of intuiting anything, and the senses are not capable of thinking anything. Only from their unification can cognition arise.”⁹⁴ Thus, where sensibility provides content for experience, and the imagination synthesizes that content by reproducing it in a representation, the understanding provides the requisite concepts for connecting and cognizing the content, both by the synthesis carried out in each faculty and by the TSI. In other words, the concepts of the understanding provide the active part of the synthesis, and the rules that govern that synthesis, which produces mental representations via the TSI.⁹⁵

⁹² Ibid., B151-152.

⁹³ Ibid., B75 and A119.

⁹⁴ Ibid., B75.

⁹⁵ To be clear, the ordering of the previous discussion ought not to imply a temporal order to the TSI itself. In other words, the Transcendental Synthesis of the Imagination is not like an assembly line where each part of the process feeds the next, but is more like a stew, wherein no one ingredient produces the distinctive flavor (in this case the resulting mental representation) but, instead, the result is the product of all the ingredients cooking together. No one part of the TSI works without all three faculties and all three syntheses.

But, what are these schema? I referred to them as a set of instructions, but we cannot think of this as a set of rules given to us by the objects we are cognizing—e.g. in perceiving a red chair the chair does not provide the instructions for cognizing it. If this were the case, then the sortal concepts deployed in cognition would be empirical concepts—ones learned through experience—rather than transcendental *a priori* concepts which are required for the possibility of having an experience. Instead, it is our own cognitive structures which provide these instructions, specifically, the concepts of the understanding. These concepts serve the function of determining basic judgments regarding quality, quantity, relation, and modality in the impressions that allow them to fit together in the TSI. More specifically, they allow for apprehension (of the manifold of impressions), reproduction (combining the manifold into a representation), and recognition (that the representation corresponds with the object we are intuiting), which together comprise the three-fold synthesis that is the TSI. We can think of this as something like a jigsaw puzzle, in that each impression received from the incoming manifold is a piece, each piece has relations to other pieces (ways in which it fits with other pieces), and the assembled puzzle is supposed to be an accurate representation of the image on the box (and if it isn't accurate, then something is wrong), where the image on the box is the analogue of the perceived object. The difference is that our minds and perceptual apparatus are what do the chopping up into pieces, the molding of the shape of those pieces, and the fitting together of those pieces. They do this through the TSI, as governed by the categories of the understanding.

However, the TSI has important implications for the subject of the experience (or the synthesizer), including the subject's awareness of itself, as for the synthesized representation to be cognized as a unified object (or recognized as corresponding to the object of experience) it must be apperceived under a single, unified consciousness, or subject. For Kant, "no cognitions

can occur in us, no connection and unity among them, without that unity of consciousness that precedes all data of the intuitions, and in relation to which all representations of objects is alone possible,”⁹⁶ and “We are conscious *a priori* of the thoroughgoing identity of ourselves with regard to all representations that can ever belong to our cognition, as a necessary condition of the possibility of all representations.”⁹⁷ Kant calls this principle the Transcendental Unity of Apperception (TUA), and, as Wilfrid Sellars notes, it is “the central concept of the [*Critique of Pure Reason*],” bridging both Kant’s account of cognition and his account of the limits of knowledge.⁹⁸ As seen above, the reproductive synthesis requires a memory of impressions from moments past, and this implies that recognition must be grounded in a numerically identical consciousness that can bridge the past and the present such that the impressions can be synthesized into a representation in the present. To keep with the jigsaw puzzle example, imagine if the assembler of the puzzle was disconnected from herself from moment to moment, forcing each move in the assembly process to start anew as she would lose any cognition of her progress or of any piece other than the one currently in hand. Such an assembly would be either impossible or entirely random. Our production of representations is certainly not impossible, nor does it seem random, suggesting that something about the representer (or assembler or cognizer or experiencer) must be stable over time in order for recognition and synthesis to occur.

In other words, as a necessary condition for the TSI, a reference to a perduring self as subject is inherent in all experience, or “representations not only represent *something*, they

⁹⁶ Kant, *CPR*, A107.

⁹⁷ *Ibid.*, A116.

⁹⁸ Sellars, “...this I or He or It (The Thing) Which Thinks...,” 6.

represent it to *someone*.”⁹⁹ This someone, or self, is the synchronic and diachronic subject of the experience, itself inseparable from the experience. As Kant puts it, “we cannot represent anything as combined in the object without ourselves’ having combined it beforehand.”¹⁰⁰ This means a being’s perceptual experience is filtered and conceptualized by it for it, with the resulting representation having a quality of being “for me,” which, therefore, implicates a “me” as the subject of the experience. In another famous passage, Kant says “The I think must be able to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much to say that the representation would either be impossible or else at least would be nothing for me.”¹⁰¹ Again, think of the assembler of a jigsaw puzzle. With each piece taken into hand, the assembler must recognize not only that she sees a piece, but that she is engaged in a project to connect that piece to other pieces. This sort of self-conscious awareness can be entirely implicit (as will be discussed in §3.4 and §3.5), in that the assembler need not have the conscious thought “I am doing this.” Instead, Kant is after an unconscious necessary condition—that it must be possible that the representation be recognized as belonging to the subject’s own consciousness—or, in this case, it must be possible that the assembler recognizes that it is engaged in some experience.

Typically, this aspect of Kant’s theory (the TUA) is explained as meaning that consciousness entails self-consciousness. As we will see in the next chapter, this explanation is misleading, though not entirely off the mark, because there are varying degrees of self-

⁹⁹ Andrew Brook, “Kant’s View of the Mind and Consciousness of Self,” *Stanford Encyclopedia of Philosophy*, last modified January 22, 2013, accessed March 14, 2013, <http://plato.stanford.edu/entries/kant-mind>, §3.4, my emphasis.

¹⁰⁰ Kant, *CPR*, B130.

¹⁰¹ *Ibid.*, B132.

consciousness. For present purposes, it should be enough to understand the implications of TUA as requiring that any conscious experience requiring synthesis requires at least implicit reference to the experiencer as a stable subject of the manifold of impressions and the resulting representations. For Kant, all of our cognitions require synthesis, and, as we have seen, the categories, the three faculties, and memory are required for that synthesis and are brought by us to our experiences—i.e. they are not given to us by objects of experience, but are required to have experience in the first place. It is in this sense that the TSI, its components, and the TUA are transcendental, as without them, no experience is possible.

Yet, our cognition of objects of perception and our cognition of ourselves is imperfect, insofar as our cognition of either requires a representation, and a representation is, fundamentally, not identical to the object it represents. Think of a mirror, for instance, where the reflection of the person in the mirror looks like the person, but is not identical to that person. It is the same with our representations. If I look at a red chair, but get a mental image of an orange, something is wrong. But, even if I look at a red chair and get a mental image of a red chair, what I cognize is not the chair, but the mental image produced by the TSI. The same holds true of our introspection. As Patricia Kitcher explains, we have no unmediated acquaintance with the thing that thinks, because the “unity of apperception, under which all my cognitive states must stand, is created through synthesis.”¹⁰² In other words, the foundation of what we take ourselves to be is our identity over time as represented in a synthesis either of something else (in which case we show up as the receiver of the experience) or of ourselves through introspection (which is mediated by intuition, and, thus, the TSI), either of which is only an appearance and not identical to the object of perception itself.

¹⁰² Kitcher, *Kant's Transcendental Psychology*, 104. See also 181-204.

As Kant says, “in the original synthetic unity of apperception [i.e. in perceptual experience], I am conscious of myself not as I appear to myself, nor as I am in myself, but only that I am... [while in introspection] I therefore have no cognition of myself as I am, but only as I appear to myself.”¹⁰³ There is no “standing or abiding self in this stream of inner appearances,”¹⁰⁴ and, as a result, the self, as commonly understood, is not given in experience, which is limited to representations and appearances that can never be rid of the self’s own subjectivity. In other words, in attempting to introspect and isolate what we are as thinking beings, we are like dogs chasing our own tails in that “we must always already avail ourselves of the representation of [ourselves] at all times in order to judge anything about [ourselves].”¹⁰⁵ Wilfrid Sellars captures this nicely when he says, “we cannot answer, ‘what is it that represents?’ save, tautologously, by ‘a representer,’”¹⁰⁶ which also applies to the questions: “what is it that experiences? An experiencer” and “what is it that thinks? A thinker.” Asking the question provides us with no new information, since we are limited to what we can know by what we can experience consciously, here of ourselves. In a famous passage, Kant says that he “had to deny knowledge in order to make room for faith.”¹⁰⁷ Here Kant acknowledges that knowledge is limited to what can be cognized in a possible experience, and what can be cognized is limited to how the world appears to the experiencer. Aspects of how the world is, in itself, separate from our interfacing with it, will forever be beyond our knowledge since we cannot possibly experience it without bringing our limited experiential apparatus along with us—in other words,

¹⁰³ Kant, *CPR*, B157-158.

¹⁰⁴ *Ibid.*, A107.

¹⁰⁵ *Ibid.*, B404.

¹⁰⁶ Sellars, “...this I or He or It (The Thing) Which Thinks...,” 10.

¹⁰⁷ Kant, *CPR*, Bxxx.

though we can know *that* something may exist beyond our experience we cannot know anything about it beyond the bounds of what we can possibly experience.¹⁰⁸

Kant argues that despite our limits we are naturally and continually driven to reach beyond these bounds by our faculty of reason: a faculty that is grounded in principles that encapsulate the unified set of rules that governs how our cognition works—i.e. the categories of the understanding.¹⁰⁹ In other words, we take the concepts that make our experiences possible, universalize them into logical principles, apply them empirically to our experiences, and seek to use those principles to go beyond the limits of our cognition. More importantly, we use these principles to try and establish metaphysical claims to knowledge about the world beyond our experiences, which inevitably leads to illusion. We can, of course, think beyond our experiences and representations—say, of unicorns or a future state of affairs—but what we can know must withstand the test of being proven possible or verifiable in experience, or derived logically and *a priori* from experience but still within the realm of possible experience.¹¹⁰ Kant’s famous examples of such a speculative over-reaching of our reason are in regards to God, the immortality of the soul, and, critically important to his ethical theory, the idea that we are free in the exercise of our wills.¹¹¹ Again, we can *think* these things, but we cannot *know* them. For

¹⁰⁸ In his arguments, this move allows Kant to escape the problems of the empiricists (such as Locke), who posited that the world is real but could not adequately explain how we can know the world, but also seems to invite the criticism of a monism like Berkeley’s—that the world exists solely in the mind. Kant escapes Berkeley’s monism by positing that the world is real and exists outside of the mind; however, we can only know the world as it appears to us in our experiences through our mental representations of the world. In this sense, the world is transcendently ideal and the conditions of the possibility of our experience and cognition are matters of our own experiential apparatus and mental structures.

¹⁰⁹ “reason is the faculty of the unity of the rules of understanding under principles,” Kant, *CPR*, B359.

¹¹⁰ See Kant, *CPR*, Bxxvi.

¹¹¹ Unlike unicorns, which we can imagine experiencing perceptually as white horses with horns, there does not seem to be a possible representation that captures God, immortality, or free will as an object of perception—hence, they cannot be known through perceptual experience.

Kant's ethical theory, this means that though we appear to ourselves as free beings, we also imperfectly cognize ourselves and cannot establish the truth, or falsity, of our freedom. This is crucial to the upshot of this chapter, as it means not only that conscious experience entails the possibility of self-conscious awareness, but also that our self-conscious awareness and knowledge of ourselves and the world is limited by our minds and perceptual apparatus. We can be wrong about what we are apart from how we experience ourselves, just as we can be wrong about the world, and the objects and beings in it, apart from how we experience it. Though the categories of the understanding guide the TSI, and the TSI implicates and relies upon the TUA, our use of logical principles derived from the categories (such as negation, causality, unity, and identity) cannot validly reach beyond the bounds of all possible experience and establish metaphysical truths about what and how we are as cognizers apart from our cognitions.

Thus, we have seen the role for rudimentary forms of memory (in the TSI), self-consciousness (in the TUA), and reason (in that the basis of our reason is the categories of the understanding which guide the TSI). We have also seen that these rudimentary forms are the foundations of our knowledge, and that our knowledge is limited to what can be cognized in possible experience. Further, we have seen how we tend to reach beyond our limitations by using our faculty of reason and how this might have implications regarding how we see ourselves constitutionally and ethically. It is now time to see what is missing from Kant's account.

§2.3 Why There Must be Degrees of Self-Consciousness, Rationality, and Autonomy

One issue that ought to stand out from the account above is a jump in levels from what is necessary for experience to how we are mistaken in our mature use of reason when we use it to

engage in speculative metaphysics. The account describes cognition at the extremes, from necessary conditions to full-blown reason, without describing anything in the middle. This jump, however, is a feature of Kant's theory, not just my explanation of it, and it presents two problems of note—one regarding a developmental account, and one regarding universal application. In both cases, I suggest that a model of degrees of cognitive capacities will mitigate these issues by providing greater explanatory range, depth, and grounds.

First, while Kant gives much attention to the categories, he gives almost no attention to empirical concepts or linguistic concepts, which is strange considering that it is by logical principles *and* these concepts that we learn and explore the world and ourselves in an effort to gain and share knowledge. Simply put, Kant gives us a scant account of how reason develops in humans, which is troublesome for his cognitive theory as well as his ethical theory since it requires cultivating one's faculty of reason and deploying it to constrain one's will.¹¹² Presumably, humans are born with the categories of the understanding, but we are certainly not born doing speculative metaphysics, nor with explicit command over logical or linguistic principles, or empirical or moral concepts. We must learn these things over time, and even if these projects rely on the categories and our default cognitive structures we may wonder, for instance, how a child goes from the necessary features in Kant's account to a command over full-blown human reason and cognition.

¹¹² To be fair, Kant does talk about derived *a priori* concepts (i.e. derived from the categories) and empirical concepts, but when he does so it is to differentiate them from the categories of the understanding (i.e. the transcendental *a priori* concepts). Further, he does suggest that reason (in terms of making judgments) must be practiced. But, my point is that he does not give an account of the structure of such notions in relation to his theories of cognition, reason, and morality. The reason for this is probably because Kant's cognitive theory is largely epistemological in nature, as he sought to ground our knowledge as well as its limits.

Answering this question has important implications. As we saw in Chapter One, reason, autonomy, and self-consciousness are the ground of our moral worth. Yet, as we saw in §1.6, the duty to cultivate one's virtue can "consist only in *cultivating* one's faculties (or natural predispositions), the highest of which is *understanding*, the faculty of concepts and so too of those concepts that have to do with duty."¹¹³ Moreover, "the less a human being can be constrained by natural means and the more he can be constrained morally (through the mere representation of duty), so much the more free he is."¹¹⁴ It was in this sense that freedom admits of degrees, but, more importantly here, being ethical involves cultivating one's understanding, which now includes concepts of duty, obligation, and the moral law—none of which are among Kant's twelve categories of the understanding. And so, cultivating understanding requires something beyond the categories, and can be greater and lesser since it can be cultivated.

Further, cultivation would imply a growing self-consciousness in that one must develop greater awareness of oneself if one is to apply reason to oneself, one's motives, and one's actions. For instance, is a child born with the categories, the TSI, and the TUA self-conscious in the same way that she will be when she tries to cultivate herself morally? Is this a different self-consciousness, or a greater, more robust, more developed one? It seems likely it is the same self-consciousness developed through cultivation, and, again, being cultivated implies greater or lesser degrees. Similarly, if reason itself did not need careful developing, we might wonder how it is we come to fall into the mistakes and illusions of speculative metaphysics, and, further, how we may lose the ability to reason due to conditions like tiredness, old age, dementia, or

¹¹³ Immanuel Kant, *Metaphysics of Morals*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996) 6:387. Hereafter, *MS*.

¹¹⁴ *Ibid.*, 6:382.

emotional duress. Thus, the lack of a developmental account means we do not know when we humans gain, or lose, the ability to explicitly deploy reason, nor do we know when we gain concepts crucial to ethics such as the ability to see other beings as having minds and ends, nor do we know when we develop the self-consciousness to regulate our will morally.

A model in which cognitive capacities and reason develop in degrees by building upon a foundation of the necessary features of cognition may resolve these issues. But, building such a model requires close attention to the differences between the categories, empirical concepts, linguistic concepts, and their respective roles—differences which, as noted above, Kant is mostly silent about. As such, I propose the following, admittedly rough, distinctions. As we have seen, the categories are concepts that must be present as a condition of the possibility of experience, as without them no synthesis, and thus no mental representation is possible. These are concepts that allow us to determine basic information—such as determining that an object exists, or that it exists as a unity or plurality—and are the basis of reason. Empirical concepts are less basic than the categories and borrow from experience. Presumably, this means empirical concepts include unconscious heuristics, similar to those found in Gestalt psychology (such as the principle of closure), that build upon the categories through experience and, unlike the categories, are not necessary for every possible experience.¹¹⁵ Instead, empirical concepts are learned (or identified) in one experience and then recognized in a future experience. Finally, linguistic concepts are semantic concepts used in the deployment of symbols, signs, and language. Historically, the lack of this sort of conceptual deployment has been used to deny the presence of reason in a being—from feral children, to non-human animals, to humans who do not speak one's native language. This is because language, presumably, allows us to reason explicitly, abstractly, logically, and

¹¹⁵ Cf. Kitcher, *Kant's Transcendental Psychology*, 226.

universally, as well as, perhaps, morally (since one must, presumably, understand what the moral law means if one is to act out of respect for it). To put this in the terms of the present discussion, linguistic concepts allow us to engage in speculative metaphysics, or understand a being's expressed motives and ends such that we can recognize them, and ourselves, as part of the moral community. However, though linguistic capacities may rely on and build upon the categories and empirical concepts, neither the categories nor the empirical concepts are themselves linguistic in nature.

An example of color vision will help illustrate the relevance of these distinctions. Essentially, when we perceive color we receive an input via sensibility, in this case a visual stimulus consisting of a wavelength of light. That stimulus is channeled through cones to opponent and non-opponent ganglion cells to and through the layers of the lateral geniculate nucleus and, finally, to the layers of the primary visual cortex.¹¹⁶ None of that adds up to the experience of color, such as “red,” but all of it is necessary for the experience. While, the empirical concept red is learned (or identified) in experience and then recognized in a future experience, the necessary biology for color vision must exist prior to the experience, including whatever sortal concepts those biological processes require in order to “label” the input as it is processed—these are the categories and without them, no visual experience occurs, let alone the experience of redness. However, a being may experience the sensation of redness without knowing what to call it linguistically—i.e. that the sensation corresponds with the word “red.” What this means is that the categories are necessary to have the experience of redness, while the empirical concept of red requires the experience of redness (and, thus, the categories). Further, the linguistic report of seeing red requires both the categories and empirical concepts (assuming

¹¹⁶ My description is heavily simplified, for a more full discussion Cf. Gordon and Abramov, “Color Vision,” 104.

one is correctly reporting redness). But, this relation is not bi-directional—i.e. linguistic concepts are not a necessary condition for possessing the categories. In other words, the lack of linguistic concepts does not entail the lack of categories of the understanding or of empirical concepts, but the presence of empirical concepts does entail the presence of categories of the understanding (as without them one could not have experiences). This point will be crucial to understanding the arguments in Chapters Three and Four, but, for now, this brings us to the second issue stemming from Kant’s jump in levels.

Second, Kant’s theory of cognition is supposed to hold for all beings who intuit spatiotemporally like humans do—much like his ethics is supposed to hold for all beings which possess reason, autonomy, and self-consciousness. As we saw in Chapter One, Kant holds that some non-human beings, such as angels and extraterrestrials, could possibly possess reason. We also saw, above, that the categories are the basis of reason and are crucial to our cognizing anything. Though Kant explicitly accounts for the possibility of beings that intuit differently than humans do (such as being able to see the future), if a being intuits spatiotemporally, like humans do, it is not clear on his theory how they would do so without the categories of the understanding—the rules of synthesis and basis of our logical principles.¹¹⁷ Though a being which intuits spatiotemporally may not possess exactly the same set of categories as humans require for cognition, it is not clear that they could possess *no* categories at all and still have synthetic representations or reason.

¹¹⁷ Cf. Kant, *CPR*, B72—““It is also not necessary for us to limit the kind of intuition in space and time to the sensibility of human beings; it may well be that all finite thinking beings must necessarily agree with human beings in this regard (though we cannot decide this).” Also see B270-271.

Yet, despite this, Kant makes some intriguing claims regarding non-human animals, or “brutes,” as he often calls them. Kant rarely discusses cognition in brutes explicitly, and even then mostly in comparison to humans, but Steve Naragon has painstakingly extracted a number of passages and reconstructed what appear to be Kant’s views on animal cognition.¹¹⁸ Essentially, Naragon shows that Kant admits that brutes are not machines (like Descartes thought) in that they have souls, representations, apprehension, feeling, and desires, but do not have apperception, universalization, concepts, or the understanding.¹¹⁹ What Kant grants to non-human animals explains Kant’s view (as seen in Chapter One) that we ought not be cruel to animals. But, what is of interest here is that what he denies non-human animals seems to deny necessary features of his own theory, specifically apperception and concepts of the understanding (or even the faculty of understanding at all). These denials have massive implications. For instance, as we saw above, one of Kant’s definitions of the understanding is “the faculty for bringing forth representations itself.”¹²⁰ But, if this is the case, how can brutes have representations but not the understanding? Further, as we saw in the explanation of the TSI above (as well as the previous paragraph), if synthesis is required for representational cognition, and the synthesis itself requires the categories of the understanding, it doesn’t seem possible that any being could have representations while not possessing *any* categories of the understanding.

¹¹⁸ Professor Naragon’s extensive research on Kant’s views regarding animals utilizes several resources that, unfortunately, I do not have access to, primarily because I do not read German and the English translations I can access often have edited out large sections of, for instance, Kant’s lectures—sections which contain many of the passages Professor Naragon points to. As such, I have chosen to reference his points, rather than the passages from which his points derive.

¹¹⁹ Steve Naragon, “Kant on Descartes and the Brutes,” *Kant-Studien* 81, no. 1 (January 1990): 4-5.

¹²⁰ Kant, *CPR*, B75.

But, what about Kant's denial of apperception in brutes? As Naragon points out, "there is no question that Kant attributed representations to [animals], there is also no question that he denied them self-consciousness [apperception]." ¹²¹ This is why, above, I said it was misleading to understand the TUA as meaning that consciousness entails self-consciousness. The reason this is misleading is because self-consciousness and consciousness can mean vastly different things depending on the context of a discussion, and each meaning has different implications. For instance, "if we look back at some of the passages wherein Kant denies brutes consciousness, what we find him saying is that brutes lack this ability to have representations of their representations." ¹²² This denial seems to be of a sort of metacognitive, reflective ability to make judgments regarding oneself. Further, "it is yet an additional ability to be self-conscious in the sense of having a representation of the "I" and the ability to attach this representation to all the rest." ¹²³ Denying this sort of self-consciousness would seem to deny the explicit ability to think of oneself as a thinker—as an owner and judge over one's thoughts, desires, and feelings. Both of these are crucial to the sort of moral self-consciousness discussed above.

However, there is another sense of self-consciousness that it seems Kant could not possibly deny—a sort of implicit self-consciousness necessary for spatiotemporal navigation in the world. For instance, it is not uncommon to see a dog catch a Frisbee, but in order to do so the dog must track not only the Frisbee over time as it flies through the air, but also itself as it moves and leaps and closes its jaws, and itself relative to the Frisbee (in order to intercept it). If the dog intuits the Frisbee, then, at least in the example, it must also have the ability to intuit itself, at

¹²¹ Naragon, "Kant on Descartes and the Brutes," 6-7.

¹²² Naragon, "Kant on Descartes and the Brutes," 12-13.

¹²³ *Ibid.*, 13.

least spatiotemporally. This, of course, does not mean the dog explicitly thinks of itself as a Frisbee-catcher, or even that it could possibly do so. However, the point is that, by Kant's own words, apperception is "an objective condition of all cognition...something under which every intuition must stand in order to become an object for me."¹²⁴ In other words, the dog's intuition must be united under one consciousness if it is to catch the Frisbee, just like representational consciousness must be united under the TUA if synthesis is to occur.

Thus, Kant's denial of these important capacities in brutes means either these features of his theory are not necessary for cognition, or that Kant owes us an alternative account of how his theory applies to animals. In this sense, it is not clear how non-human animals can intuit anything, if Kant wants to hold to this denial. In this sense, if Kant's theory is to have much explanatory value at all—in regards to humans or non-human animals—and remain in tact, it seems that a developmental account that admits of degrees might be far more valuable. As I will show in Chapters Three and Four, we have good reason to assume that non-human animals intuit and cognize much like humans do. In fact, we already assume this in our scientific research when we use animals as models for humans. For instance, much of our knowledge of visual pathways is derived from the visual pathways of primates, including the rhesus monkey.¹²⁵ And, for this reason, I will argue in the next two chapters that a model of degrees has greater scientific explanatory value than a sort of "on-off" model of cognitive capacities.

Ultimately, I find it far more likely that what non-human animals lack is something more robust—something like the metacognitive reflection Naragon points to—rather than some feature

¹²⁴ Kant, *CPR*, B138.

¹²⁵ Cf. Barry B. Lee, "Visual Pathways and Psychophysical Channels in the Primate," in *J Physiol*, 589.1 (2011), 41-47, accessed June 13, 2013, <http://onlinelibrary.wiley.com/doi/10.1113/jphysiol.2010.192658/pdf>.

necessary for Kant's theory. In fact, in a passage where Kant argues against positions that attempt to establish more about the soul (here mind, or the thing which thinks) than can be known, he admits that "consciousness always has a degree, which can always be diminished; consequently, so does the faculty of being conscious of oneself."¹²⁶ As I take it, this passage provides one of the deepest clues that Kant gives us regarding investigating minds, self-consciousness, and the nature of experience, as it opens the door for a model that maps degrees of self-consciousness to beings based upon their cognitive faculties. Thus, in the next two chapters, I will try to follow this clue while maintaining the necessary features of Kant's cognitive theory. As we have seen, memory, self-awareness and reason play key roles in Kant's theory of cognition, and metacognition plays a key role in his ethical theory. I suggest that greater capacities for memory and metacognition entail a greater degree of self-awareness and, thus, self-consciousness. This begins from a minimal self that is necessary for all perceptual cognition and is most likely present in many, if not most, non-human animal species, and culminates in the self-conscious narratives and mental and moral lives of human beings. Though I can only sketch this model, I will try and connect each level of development to moral considerability.

¹²⁶ Kant, *CPR*, B415.

CHAPTER THREE

DEGREES OF SELF-CONSCIOUSNESS AND COGNITION:

Part I: From Transcendental Conditions to Self-Conscious Reflection

“Yes, a violent quarrel was in progress. There were shoutings, bangings on the table, sharp suspicious glances, furious denials... Twelve voices were shouting in anger, and they were all alike. No question, now, what had happened to the faces of the pigs. The creatures outside looked from pig to man, and from man to pig, and pig to man again; but already it was impossible to say which was which.”

--George Orwell, *Animal Farm*

§3.1 Introduction

In Chapter One, we saw that Kant’s ethical theory is tied to his cognitive theory due to the stress placed upon rationality, autonomy, and self-consciousness as the ground of moral considerability. In Chapter Two, we saw that Kant’s cognitive theory requires a developmental account of cognitive capacities in order to be complete in terms of his own ethical project and how humans might develop as rational and moral beings, and in terms of my project in this thesis and grounding the place for non-human animals among Kant’s theories and principles. In the present chapter, I will begin to construct this developmental account by building off of the principles and necessary conditions established in Kant’s own theory and utilizing important distinctions and evidence regarding memory, cognition, and self-consciousness as found in modern theories of biology, cognitive psychology, ethology, and philosophy of mind. Part of the purpose of this chapter, then, is to isolate and clarify the structure of self-conscious experience, memory, and the conscious deployment of concepts (or reason) in experience. As such, by the end of this chapter, I will have placed many important features of cognition, including all the necessary criteria for moral considerability discussed in the previous two chapters. The results of

the efforts in this chapter will show that most non-human animals do, in fact, possess relevant degrees of the cognitive capacities that comprise moral considerability and, thus, that our moral duties ought to extend directly to those non-human animals.

Unfortunately, I cannot fully establish this model simply by citing evidence of cognitive capacities in non-human animals. To place non-human animals on the same scale as humans I must also provide a framework capable of explaining the evidence relative to important robust features of human cognition and necessary features of cognition. To fail in this regard would make my model both implausible and fundamentally incomplete. However, fleshing out all aspects of this model would be far too exhaustive and controversial for my purposes here, as it would require endorsing many assumptions that are not generally agreed upon in the philosophy of mind or the cognitive sciences. Instead, I will outline a consistent model that either explains important features of conscious cognitive lives on the same evaluative scale, or solidifies the resources for doing so in the future. The approaches I bring to this model center on first-personal aspects of experience as necessary features of experience and, thus, help to extend the explanatory value and breadth of Kant's theory. Many of these theories, especially those from phenomenology, either directly borrow from, or have important features in common with, Kant's theory of cognition. Of course, it is not obvious that Kant would necessarily agree with all the distinctions and theories I will incorporate into this model, but it is not my argument that he would agree, only that he could, indeed should, agree with these theories and distinctions based upon principled extensions of his own theories and commitments.

Nevertheless, this project requires treading on controversial ground. Specifically, I will presume that the proper way to understand the structure of cognition and conscious experience, especially on Kant's account, is as fundamentally subjective—in other words, that there are

qualitative aspects to experience and cognition that are necessary components of the structure of experience and cognition. We saw the basis of this assumption in Chapter Two. However, for most of the twentieth century, the assumption of a subjective quality to experience was explicitly denied by many dominant views and still is today by reductionist and eliminativist views that seek to eliminate the language of subjectivity in favor of purely third-personal scientific explanations.¹²⁷ Among these theories is a persistent and pervasive worry about what is scientifically investigable—a worry that has, historically, led to rejecting subjective accounts. I will take on objections extracted from these views at the end of Chapter Four, where I will argue for the scientific, explanatory, and parsimonious superiority of my model over both reductionist views and Kant’s argument from parsimony seen in Chapter One. However, fully addressing these objections requires the resources of the model I will develop over the next two chapters.

In the meantime, it should be noted that I am not alone in my endeavor to give a subjective model of the features of experience and self-conscious experience. Most of the theories I will incorporate into this model reject reductionist and eliminativist views in some important regards, and hold that subjectivity either is scientifically investigable or need not be due to the demonstrable logical necessity of subjective aspects of experience. Discussing these aspects of these views, however, would take my discussion way too far afield. Thankfully, I am not the first to try and construct a phenomenological model that includes the cognitive lives of non-human animals, and the moral duties that result from them, on Kantian grounds.

¹²⁷ See, for instance, the views of Paul Feyerabend, W.V.O. Quine, Richard Rorty, Wilfred Sellars, or, more recently, Patricia and Paul Churchland, and Daniel Dennett. Many of these views have common roots in the logical positivists and behaviorism. For a historical overview of the thinking which informed these thinkers, see Bernard E. Rollin, *The Unheeded Cry: Animal Consciousness, Animal Pain, and Science*, expanded ed. (Hoboken, NJ: John Wiley & Sons, 1998) specifically chapter four. Hereafter, *UC*.

§3.2 A Partner in Crime: Gennaro's Model of Degrees of Self-Concept

As my account cannot be fully fleshed out and established here, I ought to note, again, that I am treading on controversial ground both with this model and with the way in which I will flesh it out—that is, by appealing to memory, metacognition, and theories regarding the inherently subjective structure of consciousness. But, as noted, I am not the first to try a model like this, nor the first to attempt to ground such a model in Kant's theory of cognition. Rocco J. Gennaro has developed a degrees of self-concept model by working with Kant's Transcendental Unity of Apperception (TUA) as a base.¹²⁸ Though I shall use the same base, and many of the same distinctions and appeals to other theories—including the work of Endel Tulving regarding episodic memory and a biographical mental life—I do not fully agree with Gennaro's model. Primarily, this is because the highest level of his model fails to account for the robust self-conscious narrative constructions which characterize human mental life, while the lowest level of his model fails to account for the structure of consciousness in a way that captures what I believe Kant is after with his TUA. I will take up this latter point first as the base of the model and the former point last after building up to that level of the model (in Chapter Four). As I build this

¹²⁸ Additionally, Gary E. Varner has developed a model based partially on Kant's TUA (though not on Kant's ethics). Not coincidentally, he concludes, like Gennaro, that non-human animals have robust cognitive lives and, like me, that we owe ethical obligations to non-human animals based on those cognitive lives. For the purposes of this thesis, and as far as I can tell, Varner's model does not substantially differ from Gennaro's model, or from what I will argue for, except in the ways in which I disagree with the conclusions he draws from the model and in my explicit tie to Kantian ethics. Varner believes we have moral duties to reduce pain in non-human animals that can feel pain, but, more importantly, here, that auto-noetic consciousness yields a special moral status for any being that has such a robust sense of its own past and future. However, this criteria would only award that status to the most cognitively developed animals, such as elephants, dolphins, and the great apes. Further, Varner constructs a moral hierarchy that includes these animals as "near-persons" somewhere in between merely sentient creatures and full persons. But, to me, this status of "near-persons" seems to give too easy a pass to human concerns—that is, to humans overriding what obligations they would have to non-human animals in favor of distinctly human concerns. Neither leaving animals as potential second-class citizens, nor only letting the smartest of animals into that class accomplishes the project I have in mind and the concerns discussed in Chapter One, though I will utilize a similar model and similar distinctions. For Varner's view see his *Personhood, Ethics, and Animal Cognition: Situating Animals in Hare's Two Level Utilitarianism* (New York: Oxford University Press, 2012), and "Personhood, Memory, and Elephant Management," in *Elephants and Ethics: The Morality of Coexistence*, ed. Christen Wemmer and Catherine Christen (Johns Hopkins University Press, 2008).

model, I will use Gennaro's model as a sort of scaffolding to help support my project. On the way, I will present some other issues with Gennaro's model, sketch the distinctions that yield the middle levels, and discuss the rough ethical issues each level presents as well as the evidence that non-human animals possess cognitive capacities at these levels.

To begin, at his lowest level, Gennaro uses what he refers to as a Kantian-style argument, based on the TUA, to suggest that an organism's re-identification of objects via a purely physical, basic subject-object bodily self-consciousness entails that "it implicitly understands itself as something which endures through time."¹²⁹ While I think this is a correct and important account of bodily self-consciousness, if we are to follow Kant's actual arguments for the TUA, I believe this is too rich to be the ground floor of a model for degrees of self-consciousness (or self-concept). This is because Gennaro's formulation depends upon an implicit self-reference *in experience*, while, as we saw in Chapter Two, Kant is after transcendental *a priori* conditions for the possibility of experience. That is, Kant's TUA is structural and not only prior to experience, but a condition *for* experience. This suggests that a more basic level than Gennaro has in mind must ground his first level. I will call this more basic level "Level Zero." But, it also means that we must inquire as to what all must be present at Level Zero in order to account for the possibility of Level One—or Gennaro's lowest level.

¹²⁹ Rocco J. Gennaro, "Animals, Consciousness, and I-thoughts," in *The Philosophy of Animal Minds*, ed. Robert W. Lurz (Cambridge: Cambridge University Press, 2009) 189.

§3.3 Level Zero: The Non-Conceptual Kantian Base of Self-Conscious Experiences

On this level of the model, we find all the transcendental *a priori* aspects of Kant's cognitive model that were discussed above. The Transcendental Unity of Apperception, the Transcendental Synthesis of the Imagination, the categories of the understanding, and all the capacities for memory and cognition these aspects require to function exist at this level as conditions of the possibility of all the levels to come. They are the foundation on which those further levels are built. However, since we are now explicitly attempting to account for how a being cognizes *itself* in experience (rather than just what makes cognition and experience possible), we must examine what further implications may be drawn that make such self-conscious cognition possible. There are two issues to examine here, the first regarding what kinds of experiences will count at the subsequent levels, and the second regarding what those experiences imply about the necessary structure of a subjective experience.

First, Gennaro references a bodily self-consciousness at the minimal levels. Though a subject of experience may need to be embodied to be conscious, we can imagine a scenario where an embodied subject lacks bodily awareness, but is nevertheless conscious. For instance, perhaps this being is a *Johnny Got His Gun* type—once capable of perceptual experiences but now, for all intents and purposes, comatose to any observer. Or, perhaps the being is dreaming—as in dreaming we do not have perceptual experiences of external objects and events. The point is that the perceptual aspects of Gennaro's model would be missing, and only thought and pure consciousness would remain. However, for Kant, even introspection and thoughts are synthesized through the transcendental conditions of experience (since they must be cognized).

This means that thought and introspection must count as possible experiences at the various levels of this model, which has important implications.¹³⁰

Interestingly, the being under consideration, since it once had perceptual experiences may think or introspect with conceptual content. Think of Johnny, for instance, laying in his apparently comatose state doing math problems or thinking of his past experiences. However, we can also imagine a being that was born without perceptual capacities, or, perhaps, a brain-in-a-vat “hooked-up” to no perceptual apparatus (or a perceptual apparatus that is not currently, and never was, “on”). This being has no such empirical or learned concepts. However, it would still possess the categories of the understanding, since it is possible this being can have introspective experiences and the categories, as well as the other transcendental conditions, are required for those possible experiences. But, if the proposed being has no concepts other than the categories, what could it possibly think about, or introspect on? Simply put, there would be no content, there would only be pure subjectivity. Additionally, this subjectivity would be a transcendental condition of all possible future subjectivity—that is, if later experiences are experienced subjectively, then the possibility of experiencing subjectively must be part of the transcendental structure. This brings us to the second issue.

We know from Kant’s theory of cognition that a being’s perceptual experience is filtered and conceptualized by it for it, and, therefore, implicates a self as the subject of the experience. Unfortunately, Kant does not have much to say about what this self is, only that “in the synthetic original unity of apperception, I am not conscious of myself as I appear to myself, nor as I am in

¹³⁰ This does not mean that all beings that are capable of perceptual experiences are also capable of conscious introspection or conscious thought. These capacities may not necessarily mutually imply each other. However, in the example, the being presumably had each of these capacities and lost all but the capacity for introspection and thought. In this sense, if whatever is experienced through those capacities counts as experience, then they must also be subject to Kant’s transcendental model.

myself, but am conscious only that I am.”¹³¹ Additionally, as we saw in Chapter Two, Kant says, “The *I think* must be *capable* of accompanying all my representations.”¹³² Though “I think” is a propositional stance toward one’s own cognition that would require language, developed concepts, and higher-order thought, the operative concept here is that it is *capable* of accompanying the perception, which, I suggest, does not imply a certain level of cognitive functionality (i.e. a typical, adult human), but, instead, implies a persistent availability. This availability can be understood in terms of reflexivity and perspectival ownership.

Reflexivity is a way to ground what David Rosenthal calls, “The Transitivity Principle”—or “the observation that mental states are conscious only if one is in some way conscious of them.”¹³³ The principle itself seems obvious, as if one has a perceptual experience in the present, then one is conscious of that experience. This is required for the bodily self-consciousness we are after in Level One, but this principle must itself be grounded in our *a priori* structure in that we must account for how a mental or perceptual state goes from being unconscious to being conscious. Gennaro utilizes a Higher-Order Thought (HOT) theory to ground the answer to this question. On a HOT theory a mental state becomes a conscious mental state when it becomes the target of a higher-order mental state, or a kind of metacognition.¹³⁴ This is problematic as it requires metacognitive capacities in order to have conscious mental and

¹³¹ Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998) B157. Hereafter, *CPR*. Also, the pagination of the *CPR* varies slightly in the Academy edition from Kant’s other works, in that its format is edition (A or B) followed by section number. I have chosen to follow that pagination here. Further, all emphasis from Kant’s writings should be assumed as his unless otherwise noted.

¹³² *Ibid.*, B132.

¹³³ David Rosenthal, *Consciousness and Mind* (New York: Oxford University Press, 2005), 4.

¹³⁴ Cf. Gennaro, 184.

perceptual states.¹³⁵ A much simpler thesis, which is still compatible with the existence of higher-order thoughts, is what Matthew D. MacKenzie calls “The Reflexivity Thesis” or that “an awareness of an object reveals *itself* at the same time as it reveals its object.”¹³⁶ In other words, a perceptual experience contains within it a pre-reflective awareness of itself—that is, a subjective nature that one might not always give attention to, but, nevertheless, could upon reflection. This pre-reflective subjective nature is constituted by a subjective aspect of the experience that accompanies the objective aspect (what the state is directed toward). For instance, in perceiving a red chair one may pay attention to the chair (the objective aspect) or to the fact that one is perceiving the chair (the subjective aspect). On a HOT theory the subjective aspect is the target of the higher-order state, but was not in the original state, raising the question as to where it came from and how it became available. However, on The Reflexivity Thesis the subjective aspect is still the target of the higher-order state, but this targeting is possible because the subjective aspect was already there as part of the original mental state. Thus, the subjective aspect is always available as a necessary feature of experience, whether or not one is cognitively capable of reflecting upon it, much like Kant’s “I think.”¹³⁷

¹³⁵ Additionally, there is the issue that a HOT theory opens up an infinite regress. This is because each conscious mental state, to be a conscious mental state, requires a further state that is not itself conscious unless it is the target of a yet further state, and so on, *ad infinitum*. Gennaro asserts that the regress is stopped because the higher-order state need not itself be conscious to make the lower-order state conscious. But, this solution seems somewhat hasty to me, as it still means that a HOT theory piles up higher-order mental states in order to explain how mental states become conscious.

¹³⁶ Matthew D. MacKenzie, “The Illumination of Consciousness: Approaches to Self-Awareness in the Indian and Western Traditions,” *Philosophy East and West* 57, no. 1 (January 2007): 41, emphasis in text.

¹³⁷ This move is not unprecedented. Similar views can be found in the philosophies of Antonio Damasio, Jean-Paul Sartre, Edmund Husserl, and others—see Dan Zahavi, “The Experiential Self: Objections and Clarifications,” in *Self, No Self?: Perspectives from Analytical, Phenomenological, and Indian Traditions* ed. Mark Siderits, Evan Thompson, and Dan Zahavi (New York: Oxford University Press, 2011), 56-78. Moreover, work in bridging the traditions of Eastern and Western philosophy often focuses on this debate and the reflexivity aspects of the Buddhist philosopher Dharmakīrti’s notion of *svasaṃvedana*—see, among others, the work of Jonardon Ganeri, Matthew D. MacKenzie, and Dan Zahavi cited in this chapter and the next one.

But, what is the structure of this subjective aspect of subjectivity? For one, phenomenologists and philosophers of mind often refer to what is known as “perspectival ownership.” What this implies is that your experiences are, fundamentally, given to you in a unique first-person point of view, or perspective, as a necessary part of the experience. This point of view is a common feature of all your experiences, and marks these experiences as yours, as they show up in your experience. Interesting pathology cases aside, when you see a red chair, for instance, you do not think that the mental representation of your perception actually belongs to someone else, or comes from ten feet away from where you are. In addition, and critical to the bodily self-consciousness of Level One, when you spatiotemporally cognize an object you also, at the same time, spatiotemporally locate yourself in relation to that object. But there must be a condition of this possibility. One part of this is the TUA of the TSI, discussed in Chapter Two.

Another part of this perspectival ownership is what a phenomenologist might call the necessary subject of experience, or the “for-whom-it-is-like” for Thomas Nagel’s “what-it-is-like-to-be,” when he says, “fundamentally an organism has conscious mental states if and only if there is something that it is like to *be* that organism—something it is like *for* the organism.”¹³⁸ The notion here is what we earlier established with Kant’s TSI—that your experiences are filtered and conceptualized by you *for* you, and this implies a you. Dan Zahavi calls this “first-personal givenness” and considers it constitutive of an “experiential core self”—or a minimal self that “is defined as the very subjectivity of experience, and is not taken to be something that exists independently of, or in separation from, the experiential flow.”¹³⁹ In other words, this

¹³⁸ Thomas Nagel, “What Is It Like to Be a Bat?,” in *The Nature of Consciousness: Philosophical Debates*, ed. Ned Block, Owen Flanagan, and Güven Güzeldere (Cambridge: MIT Press, 1997), 519, emphasis in text.

¹³⁹ Zahavi, “The Experiential Self: Objections and Clarifications,” 60.

minimal self is purely formal. It is not some full-blown self, nor a substantial self, but is merely a logically implied subjective referent of experience that is not separable from experience, but also not reducible to experience (as experiences change while the owner does not).¹⁴⁰ Like Kant's TUA and "I think," it is part of the transcendental base of all our experiences, as it is the necessary subject of the subjectivity of an experience. Thus, I suggest this minimal self, as a purely formal aspect of experience, as a ground for this model of degrees of self-consciousness.¹⁴¹ I also suggest that any being capable of conscious experience must function similarly. In fact, the reactions of human and non-human animals to sensory stimulation requires it.

Of course, these distinctions just discussed are not made by Kant. However, my aim here is not to assert what I think Kant meant, or would necessarily agree with. My aim is to flesh out Kant's theory in ways he *could* agree with based upon his own principles and argumentation, and which also provide solutions to the issues we saw in Chapter Two. For clarity, it should be stressed that, at this level, none of the features just discussed are linguistic, learned, or explicit in the mind of the experiencer. Instead, the features just discussed flesh out the *transcendental* conditions of experience and self-consciousness in ways that allow us to explain the more robust cognition and senses of self experienced with advanced cognitive capacities, which build upon and enhance the subjectivity of an experience. Of course, it may be the case that a being that could, somehow, have experience without subjectivity would not possess all of these transcendental conditions. However, we are concerned with conscious experiences and if these

¹⁴⁰ Again, similar views can be found in Damasio, Sartre, Husserl, and others. See Zahavi.

¹⁴¹ A minimal, formal self need not commit us to a view about a substantial self, such as a Cartesian enduring entity, nor must such a notion come into play on the model at all. Indeed, Hindu and Buddhist traditions manage to speak of experiential selves in their debates about the existence of a self, even while the bulk of those traditions deny a Cartesian self.

considerations have been correct regarding the necessary structure of consciousness, then a being could not have a conscious experience without also having a subjective experience.

§3.4 Level One: Bodily Self-Consciousness

In laying the foundation for this level of the model, we have already seen much of what occurs at Level One. As such, this level requires little argument, as all one must accept is that the conditions of perceptual and self-conscious experience at Level Zero have now become actualized in perceptual experience. It would be absurd to deny that at least some beings have perceptual experiences. Moreover, as established in Level Zero, we cannot deny that a perceptual experience of an external object implicitly references one's own body or spatiotemporal point-of-view—it is hard to imagine how any being would gain any utility out of a perceptual experience if this was not the case. This includes anything from basic sensory information to simple and complex diachronic tasks like tracking prey, eating food, observing a sunset, or, as in an example from Chapter Two, tracking and intercepting Frisbees. Thus, bodily self-consciousness entails an *implicit* diachronic and synchronic awareness of oneself and objects of awareness. Further, as Bernard E. Rollin has argued, this includes all non-human animals, as the transcendental conditions of perceptual experience necessarily extend to non-human animals once we grant that they can perceive anything.¹⁴²

However, the implicit nature of a Level One awareness must be stressed, as though perceptual and bodily awareness entails implicit concepts of self and other, there is no

¹⁴² See Bernard E. Rollin, *The Unheeded Cry: Animal Consciousness, Animal Pain, and Science*, expanded ed. (Hoboken, NJ: John Wiley & Sons, 1998) specifically, page 45 and Chapter 4. Hereafter, *UC*.

metacognitive judgment made on this level. There is only experience in the present moment—whatever is happening, here, now. As discussed above, it is these present moment subjective experiences that will be the target of higher-order mental states and reflection. Of course, present moment experiences require the TSI, TUA, and a temporal consciousness. But, this level does not include any higher-order metacognitive capacities, nor does it imply any *conscious* awareness of the transcendental conditions or even of previous experiences, just that a conscious perceptual experience is happening, now.

Yet, this level is more robust than it may seem, as this is also the level of learning by conditioning (or, indeed, the beginning of all possible learning). What is often called procedural memory, or “remembering how” or “habit memory,” is linked with learned skills, though it can be more flexible than automated conditioning systems.¹⁴³ The paradigm example of procedural memory is remembering how to ride a bike without having to think about how you do it, even if you have not ridden a bike for a long time—that is, the memory stays with you and guides you at an unconscious level. Alternatively, one might think of Pavlov’s dogs, or B.F. Skinner’s work in conditioning by reinforcement. In Pavlov’s case we might think of the process as “stimulus-food, stimulus-food,” while in Skinner’s case we might think of the process as “stimulus-action-reward” or “stimulus-action-pain.” In either case, any conceptual content to the experience remains implicit. However, as Skinner’s experiments show that pain can reinforce learning by conditioning, this level would include experiences of pain. As a result, any being which can learn by conditioning or experience pain has at least a Level One self-consciousness.

¹⁴³ John Sutton, “Memory,” *Stanford Encyclopedia of Philosophy*, last modified February 3, 2010, accessed March 19, 2013, <http://plato.stanford.edu/entries/memory/>, §1.1.

As an ethical matter, pain ought not be confused with suffering. Presumably, suffering involves enduring pain (of some sort), and since enduring implies a diachronic element, then suffering would seem to imply a conscious awareness of that diachronic element. Since the experience is entirely present moment at Level One, any past experience of pain would be, effectively, forgotten—that is, it only remains at an unconscious level—and, thus, would not constitute suffering. This lack of a conscious diachronic awareness goes unnoticed by Gennaro in his model, as he takes his Kantian-style argument to show that “at least some form of episodic memory is necessary for even being a conscious creature.”¹⁴⁴ But, episodic memory does not just mean having an implicit understanding of oneself as existing over time, but involves the capability to project one’s current self into the phenomenal character of one’s past or future. I will return to this point shortly (transitioning from Level Two to Level Three), but, for now, the point is that at Level One a being does not have a conscious awareness of its past or future, so it is not clear how that being would project itself into its own past or future. Instead, all past and future references to experience (such as that needed for the TSI) are implicit and unconscious.

§3.5 Level Two: Concepts, Facts, and “I” as Experiencer of Mental States

Another kind of memory, however, may come in to play between Level One and Level Two, or at Level Two. Semantic memory is often termed “remembering that” and is linked with remembering facts and recognizing objects. It is also the “vast network of conceptual information underlying our general knowledge of the world.”¹⁴⁵ In this sense, the mental

¹⁴⁴ Gennaro, “Animals, Consciousness, and I-thoughts,” 189.

¹⁴⁵ Sutton, “Memory,” §1.1.

concepts used in making distinctions and in re-identification of an object come in at this level. Crucially, it is unclear whether or not this type of memory (and the concepts deployed) can be accessed unconsciously. On the one hand, recall the role of concepts in Kant's TUA—that is, that conceptual filtration occurs without our awareness. On the other hand, semantic memory (along with episodic memory) is often termed “declarative memory” or “explicit memory” as opposed to procedural memory which is “non-declarative,” “implicit,” or memory without awareness.¹⁴⁶ The point is that conscious discrimination and recall become possible by utilizing the memory of previous experiences and concepts. One can, for instance, consciously remember that the last time one pushed a button, one got an electric shock. Further, conscious recognition of distinctions such as “friend-or-foe,” or “this is not a stranger” become possible at this level. As a matter of common sense, most non-human animals seem capable of this sort of recall and these sorts of distinctions—every pet would attack its owner, or every mother would attack its own children if this were not the case, as they would not consciously recognize any previous feeling of interaction or relation with the other being even in the vaguest sense.

However, the memory and concepts at this level may not be quite as robust as this sounds, as they could still be implicit, and, as such, so could the self-reference entailed by reflexivity. In other words, one may recall *that* one knows a fact without recalling *how*, or how well, one recalls that fact. For instance, a study involving rhesus monkeys placed them in “betting” trials meant to test their metacognitive confidence in matching stimuli by allowing them to risk a reward for either no reward or an increased reward.¹⁴⁷ The trial not only tests the

¹⁴⁶ Ibid.

¹⁴⁷ See, Robert R. Hampton, “Can Rhesus Monkeys Discriminate Between Remembering and Forgetting,” in *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, ed. Herbert S. Terrace and Janet Metcalfe (New York: Oxford University Press, 2005).

monkey's ability to semantically recall the stimuli, but also to make a judgment about how strong its recollected knowledge is. There is, however, some dispute as to whether or not these betting trials involve unconscious metacognitive judgments or "a judgment about the world as perceived...[which] does not necessarily involve the self or anything self-referential."¹⁴⁸ The thinking behind this debate involves a distinction between semantically recalling some facts with a strong accompanying sense of confidence or a degree of uncertainty and asking myself "do I really know this?"

I think the best way to navigate this debate is to acknowledge that at this level the being's mental states do become explicitly self-referential in the form of thoughts like "I am hungry," even if the I-thought part of the content is not itself reflected upon.¹⁴⁹ That is, the organism may not reflect upon, or have the cognitive capacity to be able to reflect upon, the subjective aspect of the metacognitive state in such a way as to explicitly think of itself as a self.¹⁵⁰ In other words, it is a conscious metacognition, but the organism may not be aware that it is a conscious metacognition. This sort of structure seems necessary for semantic recall of facts and objects, as to recall a fact from the past, the present recollection must occur as one's own (just like one's mental representations must). If I remember a fact, it does me no good if I think it is not me that remembers, even if this only means that recall is accompanied by only a purely reflexive sense of owning that recall. If this is correct, then the being now reflexively understands itself as

¹⁴⁸ Janet Metcalfe and Lisa Son, "Anoetic, Noetic, and Auto-noetic Metacognition" in *The Foundations of Metacognition*, ed. Michael J. Beran, Johannes Brandl, Josef Perner, and Joëlle Proust (Oxford: Oxford University Press, 2012), 292.

¹⁴⁹ i.e. "I am aware that I am hungry."

¹⁵⁰ Recall that a higher-order thought takes a lower-order thought as its object. Here, the organism would not have the ability to introspect in Gennaro's sense (as a third-order mental state that takes a second-order mental state as its object), though there would be a reflexive subjective aspect to the metacognition.

something which endures through time and has experiences over time—or, as Gennaro puts it, “I *qua* experiencer of mental states.”¹⁵¹ In this sense, suffering would now come into play and, because most non-human animals are capable of this sort of recall and conceptual deployment, most animals could be said to be capable of suffering.

§3.6 Level Three: Reflection, Episodic Memory, and “I” as Enduring Thinking Thing

The answer to the question of “do I really know this?” or perhaps the ability to ask it, would come at Level Three, the level of thinking of oneself as an “enduring thinking thing.” In other words, the being becomes aware not only that it experiences mental states now, but also that it has experienced other mental states in the past that are related to the present experience, and that those experiences are related to the present *by* it having experienced them. At this level, the being’s sense of itself and objects as related, diachronically enduring entities becomes explicit. Perhaps it is too strong to say these are objects of their knowledge, but they are available for reflection as this is the level of introspection. Where the betting monkeys may have been unconsciously making metacognitive judgments, at this level it becomes apparent that “I know that I know this.” On Gennaro’s model “most animals are at least capable of level 3 self-concept and perhaps even a level 4 concept.”¹⁵²

However, we cannot follow Gennaro quite that easily. Since the diachronic sense of oneself is now explicit, episodic memory should be placed at this level (lack of explicitness is

¹⁵¹ Gennaro, “Animals, Consciousness, and I-thoughts,” 189.

¹⁵² Gennaro, 190. To prevent confusion, it should be noted that, because I built in a Level Zero beneath Gennaro’s level 1, his level 3 and my Level Three now correspond, for the most part (if one includes the distinctions I have made and will continue to make).

part of why it could not be placed at Level One). But, there are complications as to how much is built into the idea of episodic memory, especially at this level. Episodic memory is often termed “direct memory,” “personal memory,” or “recollective memory.” These are conscious memories tied to our personal identities and linked to memory of particular past experiences and events—such as the day you first learned to ride a bike—such that “some notion of self or “I” seems necessary to have a genuine episodic memory.”¹⁵³ Endel Tulving calls this kind of memory “autonoetic” and links it to mental time travel (i.e. re-experiencing the past and projecting into the future) as well as autobiographical memory.¹⁵⁴ Autobiographical memory is memory that yields a sense of continuity between past and present self as a narrative construction of self-concept and personal identity—the explicit sense, for instance, that I am the same person that learned to ride a bike so many years ago on that particular day.

Yet, autobiographical memory may be distinct from episodic memory as “Not all autobiographical memories...are episodic...[and] the converse question, whether all episodic memories are autobiographical, remains open.”¹⁵⁵ Indeed, long-term episodic memory occurs

¹⁵³ Gennaro, 187.

¹⁵⁴ Tulving’s association of episodic memory with autonoetic consciousness is important, as he distinguishes between three types of consciousness: anoetic, noetic, and autonoetic. Anoetic consciousness is bound to the present moment, makes no reference to the self, and does not involve an explicit mental representation, but is simply “a judgment about the world as perceived.” Noetic consciousness is associated with semantic memory, judgments about internal representations, and can be, but is not necessarily, self-referential. Finally, autonoetic consciousness is tied to episodic memory, mental time travel, is necessarily self-aware, and is, according to Tulving, unique to humans. However, recalling the discussion above about Kant’s TUA, Tulving’s distinctions are problematic as anoetic consciousness must involve mental representation if it is about the world as perceived and not about the world itself. In other words, if anoetic consciousness is as Tulving describes, then he must either commit to a metaphysical realism about the external world—and our ability to interact with it—or admit that perception entails mental representation. For Tulving’s views cited here, see his “Memory and Consciousness,” *Canadian Psychology* 26, no. 1 (January 1985): 1-12 and “Episodic Memory and Autonoesis,” in *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, ed. Herbert S. Terrace and Janet Metcalfe (New York: Oxford University Press, 2005), 15. Additionally, cf. Gennaro, Metcalfe, and Sutton, cited above.

¹⁵⁵ Sutton, “Memory,” §1.2.

previously to autobiographical memory in infants and young children,¹⁵⁶ autobiographical memory “development in school-age children is essentially bound up with the increasing efficiency of cognitive abilities,”¹⁵⁷ and there is often “dissociation in autobiographical memory in aging” even when long-term episodic recollection is unaffected.¹⁵⁸ The idea here is that while episodic memory seems to be a capacity, autobiographical memory seems to be developed by building on the capacity for long-term episodic memory. Thus, we cannot view episodic memory as autobiographical by default, though autobiographical memory will become important at higher levels of this model (implicitly at Level Four, and explicitly at Level Five).

The distinction between these two types of episodic memory might be put like this: a sense of one’s future or past, and that one will exist in the future or has existed in the past, is not the same as mentally inhabiting the future or past.¹⁵⁹ Jonardon Ganeri makes a similar distinction in his discussion of *svasaṃvedana* (a reflexive theory of consciousness put forth by the Buddhist philosopher Dharmakīrti), which might be useful in understanding the distinction I have in mind. Ganeri notes a difference between a first-person perspective and a first-person stance. “Having a first-person perspective is...a matter of accessing the subject-aspect of one’s own thought...one can do this even with respect to one’s thought in the past.”¹⁶⁰ What distinguishes a first-person

¹⁵⁶ Ibid.

¹⁵⁷ Laurence Picard, Isméry Reffuveille, Frances Eustache, and Pascale Piolino, “Development of Autonoetic Autobiographical Memory in School-Age Children: Genuine Age Effect or Development of Basic Cognitive Abilities?,” abstract, *Consciousness & Cognition* 18, no. 4 (December 2009).

¹⁵⁸ Pascale Piolino, Béatrice Desgranges, David Clarys, Bérengère Guillery-Girard, Laurence Taconnat, Michel Isingrini, and Francis Eustache, “Autobiographical Memory, Autonoetic Consciousness, and Self-Perspective in Aging,” abstract, *Psychology & Aging* 21, no. 3 (October 2006).

¹⁵⁹ Tulving is not very clear about this distinction in his 1985 paper (cited above) and his 2005 elaboration (cited above) seems to establish too robust of a concept—one that is truly unique to humans.

¹⁶⁰ Jonardon Ganeri, *The Self: Naturalism, Consciousness, and the First-Person Stance* (Oxford: Oxford University Press, 2012), 174. Hereafter, *The Self*.

stance is “that one occupies and endorses one’s states of mind and is not merely a spectator of them.”¹⁶¹ The important notion here is the work that the subject-aspect does in remembering a past event as opposed to “remembering *experiencing* that event.”¹⁶² In episodic memory, one recalls the subjective aspect of the experience, while in autobiographical memory one remembers experiencing the subjective aspect of the experience. For instance, I have the memory that my father died with his head in my hands. It is my memory, and I know that the “I” that recalls that memory is the same “I” that experienced that event a little over a year ago. This is an episodic memory. As an autobiographical memory, I remember what that moment felt like, from the crushing emotions in the room, to the vivid feel of the soft grey stubble on my father’s cheek, to the way how I thought of myself as an “I” fundamentally and irrevocably changed in that moment.

It is this last aspect—how one thinks of oneself as an “I”, as well as an “I” related to other objects—that further distinguishes episodic and autobiographical memory. An organism understanding that it will be in the future, or has been in the past, is not the same as the organism mentally projecting itself into the future or past as it can with autobiographical memory. This mental projection will come into play, with important implications, at higher levels. For now, at Level Three, the level of short and long-term episodic memory, any being at this level cognizes itself as a self that persists and an “I” that has, has had, and will have its own experiences. Similarly, a being at this level can track others and objects across time and understand them as persisting as well. Thus, complex interactions between the being and objects of experience now become possible. Moreover, as conceptual distinctions are a hallmark of developed reason, it

¹⁶¹ Ibid., 12.

¹⁶² Ibid., 172, emphasis in text.

seems that while Level Two might be the beginning of reason insofar as one categorizes others and objects (such as “friend-or-foe” discussed above), at Level Three one seems to become able to generalize those distinctions and utilize them in long-term planning.

For instance, typically, western scrub jays cache food at a place where they have learned they will be the following day—an inherent capacity to future plan that itself seems to involve projecting some degree of self-concept into the future. In a set of fascinating studies, scrub jays were given preferred, but perishable food as well as less preferred, but less perishable food and were shown “even days after caching, to know not only *what* kind of food was *where* but also *when* they cached it.”¹⁶³ As the authors of the studies note, “the information about a specific caching event is represented in a what-where-when memory with episodic-like properties, whereas general knowledge about the relative desirability of difference [sic] food types and the way in which desirability changes with the time in the cache, is represented by semantic-like memories.”¹⁶⁴ The authors do not say why these capacities are merely “episodic-like” and “semantic-like,” leaving one to speculate if these terms are the result of a cautious predilection to reserve the actual capacities for humans—a point I shall return to in the last section of Chapter Four in my arguments from parsimony.

If the scrub jays are not convincing enough, in a more general example Janet Metcalfe notes that action monitoring skills presume future planning—such as primates climbing trees. As she says, “Presumably to accomplish such acrobatic feats so common in the wild, our primate

¹⁶³ Gennaro, “Animals, Consciousness, and I-thoughts,” 188, emphasis in text.

¹⁶⁴ Nicky Clayton, Nathan Emery, and Anthony Dickinson, “The Rationality of Animal Memory: Complex Caching Strategies of Western Scrub Jays,” in *Rational Animals*, ed. by Susan Hurley and Matthew Nudds (New York: Oxford University Press, 2006), 213.

ancestors would need to have a finely tuned action monitoring system.”¹⁶⁵ In other words, climbing a tree involves a strategy and, more importantly, reference to one’s own spatial positioning and motor functions over a span of time—even if you’re making it up as you go along. Such action monitoring, or in the case of scrub jays, monitoring one’s own survival strategies, is prevalent in non-human animals and does seem to exhibit cognitive functioning, metacognitive functioning, episodic memory, and self-reference while being widespread in the phylogenetic community. Indeed, there is an expanding body of literature “strongly suggesting episodic memory in a variety of animals, including dolphins, birds, and rodents (such as mice and rats)”.¹⁶⁶

Since there is reason present here (in the form of concepts, and deployment of concepts in strategizing), there is presumably, by Kant’s view discussed in Chapter One, a degree of autonomy present at this level as well. If one needs a strategy to survive or climb a tree, then, presumably, one chooses when to enact that strategy to achieve an end, perhaps even when to adapt and improve upon that strategy (remembering that learning has entered the picture at Level Two). Perhaps that end is climbing a tree to get food, or to escape danger, or to play; perhaps that end is to save perishable food for times of scarcity. The point is that one has a motivating reason, a rationale, to do what one does in a given situation, and one explicitly deploys a strategy to meet those motivations. Thus, it appears that at Level Three there are all the criteria for ethical considerability: self-consciousness, reason, and autonomy, including a limited ability to set one’s own ends (in terms of strategy). Moreover, it ought to be noted that the features of Level Three reference the phenomenal aspects of first-personal givenness and the subject-aspect from Level

¹⁶⁵ Metcalfe and Son, “Anoetic, Noetic, and Auto-noetic Metacognition,” 298.

¹⁶⁶ Gennaro, “Animals, Consciousness, and I-thoughts,” 188.

Zero, as well as Kant's discussions of the TUA and TSI insofar as these more robust experiences still require synthesis, and the presence of the "I-think" has started to become explicit. Thus, it appears the Kantian base of this present model is still very much in play.

§3.7 Interlude: Recapitulation and Clarification

That last point about ethical considerability deserves some clarification as it is the premise I have tried to reach and demonstrate—indeed, promised—starting in Chapter One. As I argued in §1.6, moral considerability is based upon the same three cognitive faculties that are also necessary for the ability to set one's ends: self-consciousness, reason, and autonomy. However, even though Kant places much emphasis on the ability to set and govern one's ends, as we saw in Chapter Two, we are in need of a developmental account of when these capacities develop in beings, and, thus, an account of which beings we may have direct moral obligations to. It is this account that I have been trying to develop into a model for degrees of self-consciousness here in this chapter.

So far, we have placed pain, suffering, learning, empirical concepts, generalization, categorization, self-consciousness, metacognition, memory, and autonomy, all based upon a model that takes subjectivity and Kant's Transcendental Unity of Apperception as basic. Further, we have noted that all animals which can perceive belong at least at Level One, where pain and conditioning are possible; that most animals exhibit the basic recall and conceptual deployment of Level Two, where suffering is possible; and that, by example, dolphins, scrub jays, primates (including humans), mice, and rats belong at least at Level Three, the level which exhibits all the criteria of moral considerability, at least to some degree. Based upon the arguments in §1.3 and

§1.4, this means we would have *indirect* moral duties (or duties with regard to) to creatures exhibiting the cognitive capacities of either Level One or Level Two. This includes most animals and we ought not be cruel to these animals that can experience pain and suffering, as doing so prevents us from developing our own moral virtue. Further, since beings at Level Three possess all the necessary conditions for moral considerability, we would have *direct* moral duties (or duties to) to these animals and, thus, must respect and take their ends into account in our moral deliberations. But, is this all that need be established? No, for three reasons: I must make clarifications about the coherency of this model, I must anticipate the practical difficulties with deploying this model, and, finally, to ground this conception of moral considerability I must still map humans to this model (as well as any other non-human animals with advanced capacities).

First, the model I am arguing for here is not as clear-cut as I have made it out to be, but is more like a continuum, which has important implications for both the classification and the ethical treatment of these beings. For instance, not every being at Level One necessarily experiences pain. We can imagine that there are (and there probably are) beings which can perceive and which have bodily self-consciousness, but which do not have the requisite neural structures to experience pain (whatever those may be). Because of this, it may also be possible that at least some animals at Level Two do not actually experience suffering, but are, nevertheless, conscious. This is a serious problem for grounding ethical duties if suffering and pain are the primary criteria for moral considerability and the avoidance of unnecessary pain and suffering is our ethical goal, especially since, as far as I am aware, there is no general consensus as to what is required to experience pain. However, these are not the criteria we have been after. We have been pursuing the level at which a being can set its own ends, thereby exhibiting reason, autonomy, and self-consciousness, which changes the issue considerably.

The difference can be illustrated by Bernard E. Rollin's distinction between needs and interests in his *telos*-driven ethic of moral considerability. For Rollin, we can say that a car has needs (such as gas or oil), or that a lawn has needs (such as water), but an *interest* means a need plus the awareness of that need such that the fulfillment or thwarting of the need matters *to the organism* insofar as it is "aware of its struggle to live its life."¹⁶⁷ But, pain is not the only indicator of struggle, "frustration, anxiety, malaise, listlessness, boredom, and anger" also count as indicators that the being takes the state it is in as to be avoided, or as undesirable.¹⁶⁸ In this sense, in the absence of clear criteria for the ability to feel pain, behavior that indicates a to-be-avoided-ness response in non-human animals may help place animals on the levels of this model. Of course, some to-be-avoided-ness responses are likely unconscious reflexes. Daniel Dennett points to an example of this type of response regarding unconscious pains which prevent us from assuming "awkward, joint-damaging positions while we sleep."¹⁶⁹ But, at least some to-be-avoided-ness responses would be accompanied with a conscious fight-or-flight survival response, such as when a spider frantically flees your attempts to capture it. Further, some such responses may be accompanied by the conscious deployment of a strategy to avoid the undesirable stimulus, such as when a rabbit freezes and tries to blend in with the environment in order to escape the attention of a potential predator. Each of these capacities corresponds to Levels One, Two, and Three, respectively, and is why in addition to pain and suffering I have included criteria such as conditioned learning and conceptual deployment.

¹⁶⁷ Bernard E. Rollin, *Animal Rights and Human Morality*, 3rd ed. (Amherst, NY: Prometheus Books, 2006), 100 and 102. Hereafter, *ARHM*.

¹⁶⁸ *Ibid.*, 102.

¹⁶⁹ Daniel Dennett, "Animal Consciousness: What Matters and Why," in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 118.

As such, Level Three still serves as a successful cut-off point between two capabilities observable in behavior. On the one hand, there is the conscious deployment of reason, autonomy, and self-consciousness which implies interests and ends even if the strategies deployed are not themselves explicitly reflected upon (which may require higher levels on the scale). On the other hand, there is the implicit or unconscious “to be avoided” or “to be defended from” behaviors exhibited in the vast majority of non-human animal species. In this sense, in the absence of consensus of what structures are necessary for an organism to suffer or feel pain, we can still place many non-human animals on the suggested model. But, it must be further noted that the model I am arguing for here is not meant to concretely place every animal onto the model at a particular level, but is meant as more of a conceptual map by which we could place animals on the scale given neurological, behavioral, biological, and psychological evidence of their cognitive lives. The acknowledged incompleteness of this model, and the difficulty in mapping each and every non-human animal to the model, bring us to the second and third reasons why establishing Level Three is not enough.

Second, there is the absolutely crucial practical question as to how we are to fulfill the perfect and imperfect duties we have directly to the non-human animals that would place at Level Three. This level seems to include most non-human animals that concern most animal rights activists. In other words, while insects and earthworms probably do not place at Level Three due to a lack of demonstrating the conscious metacognitive functioning and strategizing which characterizes that level, livestock (including poultry and fish), most research animals (including rats and mice), primates, companion animals, elephants, dolphins, whales, and many others do place at Level Three and do exhibit the corresponding capacities. How are we to reconcile that perfect duties entail rights, and imperfect duties entail respecting ends, when it is

not clear that any of these non-human animals have clear long-term projects that we could adopt as ends, and, further, when it is difficult to see how and why we would grant them rights since they would not understand these rights?¹⁷⁰

I do not claim to have ready answers to these important questions, but I can suggest two directions. One, in terms of rights, there is a difference between societal rights—or the rights conferred upon citizens and protected by governments—and basic moral rights. With the help of Rollin’s distinction noted above, we can understand basic moral rights as *needs* that ought not be unduly frustrated—needs as basic as access to potable water, air, food, shelter, a reasonable degree of freedom from confinement, and protection from unnecessary and needless pain or suffering. What exactly this entails needs more attention than I can give it here, but the basis of such basic moral rights is well-grounded in the Kantian ethic I have been discussing, insofar as they meet the guidelines of respecting the inherent dignity of others. Further, such rights could be respected without necessarily giving up many of our agricultural, dietary, and research practices, though how we go about practicing them would, most definitely, require massive overhauling.¹⁷¹ Two, respecting the ends of non-human animals is, perhaps, a little more difficult, but can be grounded similarly and in a way that coheres with basic moral rights—as a basic respect for the *interests* of the being. For instance, when possible I will capture spiders that are in my home and place them outside where, as far as I can tell, they stand a better chance at flourishing than they do in my home, where they might be killed by my wife. Such practices would not be unproblematic with all animals and all species that deserve respect, but the

¹⁷⁰ This same point has been noted by others, including Peter Singer in his “All Animals Are Equal,” when he says “Since a pig can’t vote, it is meaningless to talk of its right to vote.” See Peter Singer, “All Animals are Equal,” in *Ethics in Practice*, 3rd ed., ed. Hugh LaFollette (Malden, MA: Blackwell Publishing, 2007), 172.

¹⁷¹ In other words, we could still eat meat, use animals for labor, and perform research using animals, but under heavy restrictions that would need to cohere with the implications of this model.

following principal should work: that all beings at this level deserve the *opportunity* to flourish as much as reasonably possible. In my eyes, this principle and the idea of basic moral rights can be coupled into an ethic applicable to most non-human animals and inline with Kant's ethics. Such a coupling could guide the practical application of this model, though to be fully inline with Kant's ethics we must place what it means to participate in the ends of others more directly than just allowing for their possibility. This brings us to the final clarification in this interlude.

Third, Level Three is not enough, because, simply put, there are many non-human animals with further developed cognitive capacities, including non-human animals that seem to exhibit the foundations of moral reasoning. Though Kant ties self-consciousness to rationality, there is a huge difference between a rational agent—which is capable of free choice and a level of cognition to understand one's choice—and a being worthy of moral consideration. We have established the latter, but not the former. Further, there is also a difference between a rational agent and an ethical agent—a being that can understand its choice, as well as understand that it ought to choose in certain ways. I believe it is important to establish just when these senses of agency might develop, as they distinguish between beings which exhibit moral deliberation and owe duties to other beings, and non-human animals which exhibit advanced cognitive abilities that might entail even further moral duties than what the framework of basic moral rights coupled with a principle respecting opportunities to flourish might provide. If there are such non-human animals, then my model is fundamentally incomplete—because it fails to place those animals as well as humans on the same scale of moral considerability (which was the intent of this model). To complete this model, then, I need an account of ethical agency. To achieve this, I must discuss the further considerations of a theory of mind, ethical behavior, and implicit and explicit autobiographical memory. This will be the project of Chapter Four.

CHAPTER FOUR

DEGREES OF SELF-CONSCIOUSNESS AND COGNITION:

Part II: From Theory of Mind to Moral and Parsimonious Theorizing

George went on. “With us it ain’t like that. We got a future. We got somebody to talk to that gives a damn about us...”

Lennie broke in. “*But not us! An’ why? Because...because I got you to look after me, and you got me to look after you, and that’s why.*”

--John Steinbeck, *Of Mice and Men*

§4.1 Introduction

As this chapter is a continuation of the previous one, the project remains much the same. To complete Kant’s ethical project, as well as my own project to ground moral considerability for non-human animals in Kant’s theories based upon their cognitive lives, a model is needed that provides an account of how reason, autonomy, and self-consciousness develop in beings. As with the previous chapter, I will continue to develop this model based upon Rocco J. Gennaro’s similar model, as well as distinctions and leading theories from biology, cognitive psychology, ethology, and phenomenological theories in the philosophy of mind. I believe all of these theories and distinctions are consonant with Kant’s own theory, principles, and argumentation once the anachronisms and denials of crucial cognitive capacities in non-human animals are rightly ejected in a principled way, as detailed in Chapters One and Two.

However, this chapter will take on a decidedly more evolutionary and ethological approach than the previous chapter, due, specifically, to the subject matter and the possibility of any scientific investigation into the rich and robust cognitive lives at these levels of the model.

Moreover, this chapter will complete the development of the resources needed to respond to some important objections to my project and model, as well as to Kant's argument from parsimony detailed in Chapter One. The point of this disclaimer is this: that there are, perhaps, persistent objections and questions that may have led readers to, so far, resist some of the steps, implications, and conclusions of the model I am developing. Once the model is complete, I will take on such doubts and objections in the last section of this chapter.

In the meantime, in the previous chapter, I placed the construction of this model on pause at Level Three—the level at which, as promised, a being can be said to possess all of the relevant criteria for moral consideration (reason, autonomy, and self-consciousness) to a sufficient degree (a degree to which it can be said to set and deploy its own ends). However, it was noted that the model must go further; it must place humans, rational agency, ethical agency, and any more advanced non-human animals onto the model. As I put a concern in Chapter One, a model that places humans and non-human animals on the same scale of moral consideration, as different in degree rather than in kind, must account for important differences without driving too wide of a gulf between similarities. If my model is successful, this is exactly what we will see in this chapter. As we will see, many non-human animals may possess advanced cognitive capacities, even to the extent of possessing rudimentary, and perhaps not so rudimentary, forms of rational and moral reasoning. Establishing this will also establish the conclusion of this thesis: that because some non-human animals do, in fact, possess relevant, even robust, degrees of the cognitive capacities required for direct moral consideration, then our moral duties do extend directly to those non-human animals as an implication of Kant's arguments, theories, and principles and the cognitive lives of those non-human animals.

§4.2 Level Four: Implicit Autobiographical Memory and a Theory of Mind

At Level Three, I distinguished between episodic memory and autobiographical memory, noting that the latter builds upon the long-term form of episodic memory, as evidenced by its development in human children, as well as its loss in aging. Presumably, as episodic memory develops into long-term episodic memory and, with further cognitive development, into autobiographical memory a being gains an implicit sense of its own biography along with the ability to mentally inhabit the future or past. This explains Jonardon Ganeri's use of the words "occupy" and "endorse" as opposed to mere spectatorship. To follow Ganeri's distinction a little further, in the first-person stance that I believe characterizes autobiographical memory there are the "relations of involvement, participation, and endorsement that sustain autonomy."¹⁷² In other words, the being implicitly understands itself as a being which makes choices, does things better or worse, and can have a sense of approval or disapproval of its own actions.

For instance, to return to the studies of the complex caching strategies of scrub jays, in addition to the long-term episodic memory exhibited by the scrub jays, in the studies many "scrub jays return alone to caches they had hidden in the presence of others and re-cache them in new places."¹⁷³ Perhaps the thinking for the scrub jays may take the a form of "I need to re-cache that food" or "I was prevented from adequately caching that food before." This means the being can, in some way, reflect upon its deployment of strategies, and the motivations and rationales that cause it to deploy those strategies, even if it cannot yet reflect on the strategy itself (e.g. to improve upon it). The point is that the being understands itself as a being with a

¹⁷² Jonardon Ganeri, *The Self: Naturalism, Consciousness, and the First-Person Stance* (Oxford: Oxford University Press, 2012), 14.

¹⁷³ Rocco J. Gennaro, "Animals, Consciousness, and I-thoughts," in *The Philosophy of Animal Minds*, ed. Robert W. Lurz (Cambridge: Cambridge University Press, 2009), 193.

biography—a narrative construction which builds upon the subjective aspect of experience as experienced and ties previous experience to present experience, as well as to potential future experiences. This can be further seen by reflecting on why the scrub jays returned to re-cache the food: the fact of the presence of others during the original caching. This suggests the deployment of a theory of mind and a sort of explicit self-reference (i.e. “*others* know where *I* have cached food and, because of that fact, I need to re-cache that food”). In other words, the scrub jays not only tracked what food they buried where and when, but also that other beings that might be interested in their food knew where they buried that food and that, as a result, that food would not be safely cached for future use. This reveals a complex interaction, and the ability to track that interaction, between the being’s sense of its own biography and the presence, desires, and intentions of other beings.

To continue with Gennaro’s model, at his Level 4 Gennaro places what I have just called a theory of mind, which can be expressed as “*I qua* thinker among other thinkers.”¹⁷⁴ In other words, a being at this level can think of itself as one mind among many minds, because it can recognize that other beings have minds, in the form of interests, desires, and intentions. Whereas at Level Three one thinks of oneself as diachronically related to objects (some of which have generalized labels attached to them, such as “dangerous” or “desirable”), at Level Four one thinks of oneself as related to other *beings*. The primary reason for placing autobiographical memory at this level is due to the very fact that one now thinks of oneself as a thinker among other thinkers, an agent among other agents, a being among other beings. A self-other distinction has now been fleshed out with enough concepts that one understands oneself as a being with a biography—even if that biography is mostly comprised of what Bernard E. Rollin calls the

¹⁷⁴ Gennaro, “Animals, Consciousness, and I-thoughts,” 189.

awareness of “its struggle to live its life.”¹⁷⁵ This understanding of oneself may be minimal and implicit rather than the kind of explicit narrative self-identity that characterizes adult human life and future planning—the kind of “I want to be the kind of person I envision myself being in the future.” Such explicit reflection on one’s biography—indeed, the conscious capacity to construct it as an *auto*-biography—will come at Level Five and, like a theory of mind, must be developed, even in humans.¹⁷⁶ Nevertheless, this implicit understanding of oneself is a robust narrative construction utilized in the being’s self-conscious life.

Consider the case of a female chimp named Panzee.¹⁷⁷ Having watched a trainer hide 34 objects, after a time delay and using a lexigram board Panzee was not only able to guide a different trainer (who didn’t even know about the experiment) to all 34 objects, indicate the object on her lexigram board 84% of the time, and identify the objects over 90 hours later, but she also “showed intensified vocalization, shook her arm, and bobbed her head or body as the caretaker got closer to the site,” even going to pains to establish joint attention between herself and the trainer.¹⁷⁸ In addition, Panzee seemed “highly certain, and highly keyed up, [about] her own knowledge.”¹⁷⁹ Such amazing results clearly establish metacognition, long-term episodic

¹⁷⁵ Bernard E. Rollin, *Animal Rights and Human Morality*, 3rd ed. (Amherst, NY: Prometheus Books, 2006), 100 and 102. Hereafter, *ARHM*.

¹⁷⁶ For an account of the development of a theory of mind in humans, see Valerie Stone, Simon Baron-Cohen, and Robert T. Knight, “Frontal Lobe Contributions to the Theory of Mind,” *Journal of Cognitive Neuroscience* 10, no. 5 (September 1998), 640-656.

¹⁷⁷ See Charles Menzel, “Progress in the study of chimpanzee recall and episodic memory,” in *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, ed. Herbert S. Terrace and Janet Metcalfe (New York: Oxford University Press, 2005). Both Metcalfe and Gennaro cite this study in their articles cited in this chapter.

¹⁷⁸ Janet Metcalfe and Lisa Son, “Anoetic, Noetic, and Auto-noetic Metacognition,” in *The Foundations of Metacognition*, ed. Michael J. Beran, Johannes Brandl, Josef Perner, and Joëlle Proust (Oxford: Oxford University Press, 2012), 294-295.

¹⁷⁹ *Ibid.*

memory and what looks like auto-noetic consciousness, or the “time travel” so important to Tulving’s distinctions (as seen in §3.6). But, what should really stand out here is Panzee’s attempts not only to communicate with the trainer by more intense gesturing and vocalization—showing both an awareness of the task of guiding the trainer to the object as well as that the trainer would respond in more desirable ways to these more intense signals—but also her attempts to establish joint attention coupled with her awareness of her own knowledge. This suggests an awareness of her complex biography that is accompanied with spatial facts, facts regarding the identification of objects, facts regarding the symbols on the lexigram board, and certainty regarding all these aspects of her previous experiences and learning all related to a present task oriented around a goal in the future (i.e. completion of the task). Perhaps more importantly, Panzee’s attempts to establish joint attention with the trainer indicate an awareness that the trainer’s attention can be directed and that the trainer can be directed toward what she, Panzee, is attending to. This is the theory of mind at work—attributing mental states, attention, intention, and agency to others.

To bring these ideas together, as perspectival ownership was noted above to be a necessary feature of subjectivity, at Level Four perspectival ownership over and awareness of one’s past, present, and future actions becomes part of the phenomenal experience of oneself as one lives one’s life. Awareness of these events and actions are tied together by an implicit narrative construction—a story of how one understands oneself as an “I”—that is continually deployed in one’s experiences and is, likely, comprised of interwoven threads of the subjective aspects of one’s experiences over time. I have called this narrative construction a being’s biography. With a theory of mind, a being gains the capacity to recognize the agency of other beings. One sees oneself, and one’s biography, as interlaced with the actions, intentions, desires,

and motivations of others. In this sense, one begins to see oneself and others as agents with reasons (loosely construed), which one understands in terms of ascribable motivations and desires, and, as we will now see, restrictions on actions.

Just as the previous levels have been grounded on Kant's Transcendental Unity of Apperception (TUA) and each lower level has served as the ground of the level above it, so these two capacities brought together form a new sort of ground for mental life—one about personal narrative biographies, reasons, intentions, and meaningful interactions with others. This is not abstract, linguistic, or symbolic reasoning, but is more like a conscious practical sense deployable in relations with others—a sort of push and pull regarding what social actions are appropriate. Kant proposes something similar at the base of human morality, in that our faculty of reason pronounces an ought as “a measure and goal, indeed, a prohibition and authorization” that takes the form of “imperatives that we propose as rules to our powers of execution in everything practical.”¹⁸⁰ Moreover, “Practical freedom... is *not merely* that which stimulates the senses... that determines human choice, but we have a capacity to overcome impression... by representations of that which is *useful or injurious*... but these considerations... depend on reason.”¹⁸¹ Kant goes on to say that these considerations yield laws as imperatives, as *what ought to happen*. The only real difference between what Kant suggests and what I suggest here at Level Four is that the being cannot yet reflect upon the law, but would feel the pull of it. As such, I suggest that when coupled with a theory of mind the phenomenal awareness of one's

¹⁸⁰ Immanuel Kant, *Critique of Pure Reason*, ed. and trans. Paul Guyer and Allen W. Wood (Cambridge: Cambridge University Press, 1998) B575-576. Hereafter, *CPR*. Also, the pagination of the *CPR* varies slightly in the Academy edition from Kant's other works, in that its format is edition (A or B) followed by section number. I have chosen to follow that pagination here, as well as follow the Academy edition in references to Kant's other works. Further, all emphasis from Kant's writings should be assumed as his unless otherwise noted.

¹⁸¹ Kant, *CPR*, B830, my emphasis.

actions is the seed for the capacity for the explicit rational and ethical agency that will characterize Level Five. To further understand why, we must look at how a theory of mind gets deployed in inter-subjective relationships at Level Four.

Obviously, many non-human animals are social creatures, whether in packs, herds, troops, colonies, flocks, murders, prides or whatever we might call their social groupings. In the case of Panzee, we have already seen that complex communication involving mutual attention, gestures, and changes in tones of vocalization is possible in non-human animals. However, due to the highly publicized work of ethologists, such as Louis Leakey's "Trimates" (Dian Fossey, Birutė Galdikas, and Jane Goodall), many people readily grant that primates are smarter than most other non-human animals and have complex cognitive and social lives. Likewise, the complex communication strategies of dolphins and whales is well-known and also leads to ready ascriptions that they possess greater cognitive capacities than other non-human animals.¹⁸² Elephants also tend to get an easy pass due to their capacities for memory and deployment of concepts requiring a theory of mind.¹⁸³ In fact, it is well-established that these non-human animals (primates, cetaceans, and elephants) are often on par with eighteen-month-old to four-year-old human children in their cognitive development and deployment of a theory of mind.¹⁸⁴ What is often ignored is the inter-subjective framework necessary for human and non-human animal social interaction, and how widespread that framework is in other non-human animals.

¹⁸² Cf. Hal Whitehead, Luke Rendell, Richard W. Osborne, and Bernd Würsig, "Culture and Conservation of Non-humans with Reference to Whales and Dolphins: Review and New Directions," in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008).

¹⁸³ Cf. Varner, "Personhood, Memory, and Elephant Management," in *Elephants and Ethics: The Morality of Coexistence*, ed. Christen Wemmer and Catherine Christen (Johns Hopkins University Press, 2008).

¹⁸⁴ Again, Cf. Varner, "Personhood, Memory, and Elephant Management."

Juan Carlos Gómez argues that what is explicitly deployed in a theory of mind depends upon an implicit inter-subjective framework that allows the capability “of engaging with [others] in specific patterns of intersubjective interaction that include emotional and expressive behaviors...but also of representing and understanding [others] as capable of engaging in these interactions.”¹⁸⁵ This sort of mutual-awareness (as Gómez calls it), clearly exists in primates, but also seems required for the vertical (intergenerational) and horizontal (within the same generation) transmission of culture.¹⁸⁶ In other words, one must have this framework to deploy the theory of mind that seems apparent in teaching customs to others and interacting based upon those customs. What I suggest is often ignored is that this same framework is required to explain the tremendous amount of anecdotal evidence regarding cross-species “friendship” and cooperation in non-human animals, as well as basic social hierarchies (such as dominance hierarchies) and group practices (such as grooming) that are prevalent across non-human animal species.

To elaborate, anecdotal evidence of cross-species friendship is abundant and, relatively recently, ethologists have published volumes of literature on the subject.¹⁸⁷ Among my favorite examples are a lion, a tiger, and a bear being raised in the same habitat since they were cubs;¹⁸⁸ a lioness adoptively caring for five different baby Oryx on five separate occasions;¹⁸⁹ and Rollin’s

¹⁸⁵ Juan Carlos Gómez, “Are Apes Persons? The Case for Primate Intersubjectivity,” in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 170.

¹⁸⁶ For this distinction, see Whitehead, cited above.

¹⁸⁷ See, for instance, the work of Frans de Waal and Marc Bekoff, especially Bekoff and Jessica Pierce’s *Wild Justice: The Moral Lives of Animals* (Chicago: University of Chicago Press, 2009).

¹⁸⁸ See Jennifer S. Holland, *Unlikely Friendships: 47 Remarkable Stories from the Animal Kingdom* (New York: Workman Publishing Company, 2011), 76-81.

¹⁸⁹ *Ibid.*, 82-85.

story of an “non-domesticable” German shepherd attack dog befriending and defending a male turkey.¹⁹⁰ Such relationships defy explanations which appeal to instinctual programming. Lions, tigers, and bears are all apex predators and would have little reason to get along in the wild by instinct, especially since they come from different continents; the lioness naturally preys on antelopes, including their young; and while aggressive attack dogs may not consider a relatively docile turkey an enemy, it certainly would be a stretch to say the German shepherd befriends the turkey by instinct. While we can make sense of human friendships with pet animals since we have bred pets to be more domesticable, we cannot utilize the same explanation for why they and wild animals might get along with entirely different species, especially ones they usually do not interact with, or interact with entirely differently under normal circumstances.

Of course, many accuse such anecdotal evidence of anthropomorphism, and, indeed, we must be careful not to ascribe too much to these examples, especially since they are not the norm. However, the mutual-awareness framework in terms of deploying a theory of mind does seem a far more likely explanation than mere instinct, as it reveals that in such examples the non-human animals see the other as not a target of usual behavior, but, instead, a sort of partner in another behavioral pattern. This explanation is strengthened if we consider the grooming practices prevalent in many species, such as jointly bonding by removing insects or dirt from each other’s fur. Such activities depend on mutual social agency and the bonds that are created and strengthened by this behavior tend to reinforce group cohesion.

Moreover, the mutual-awareness/theory of mind explanation is even further strengthened by analyzing social hierarchies, especially dominance hierarchies. Such social structures are

¹⁹⁰ Rollin, *ARHM*, 55-56.

extremely prevalent and the sort of challenges which establish pecking order seem readily explainable in terms of a being understanding its own biography amid other beings. For instance, a subordinate wolf must understand that the alpha is dominant, though the subordinate wolf may itself be dominant over others. In this sense, the wolf must track its relations with each member of the pack. Moreover, in many species, a subordinate may challenge a more dominant conspecific to win rewards or even establish its own dominance. Such interactions seem highly dependent upon both the idea that one stands in relation to others and that one remembers how one stands in relation to others—thus referencing both the capacity to see others as beings with minds, and the capacity to track one’s own biography and status relative to others. Finally, in novel conflict situations with other species, such as bears and dogs in the wilderness, non-human animals often make dominance displays as if these other beings ought to respond to these displays. Of course, the animal friendship examples suggest a degree of flexibility in these encounters, which is exactly the point as these interactions exhibit learned behaviors, customs, and facts about oneself (biography) as well as ascriptions of intent and judgments of others (theory of mind).

One could, of course, object that social hierarchies also exist at what seem to be lower levels on the model, such as with hives of insects. However, what is not clear in the case of insects is clear at Level Four—the ability to consciously determine when to act or refrain from acting given a particular encounter. At Level Four, the behaviors are not automatic, but seem to be based upon complex conscious discriminations and judgments regarding incoming information about the situation and the beings involved. Recall that my solution to the Dilemma of Moral Considerability in §1.5 involved a middle ground consisting of the coupling of understanding oneself and one’s own capacities and recognizing the capacities of others such

that one sees the other as a being one can be obligated to. It is because of this that I suggest that Level Four is the seed for ethical agency. At this level, a being can be reasoned with, at least to some degree, such as bargaining for food, toys, or social and reproductive status. Thus, the being possesses some advanced degree of rational conceptual deployment and autonomy, as well as a robust self-conscious life that includes seeing itself as one among other minded beings. Since these traits seem required for social interaction, most non-human animals which engage in social behavior are, thus, candidates for Level Four.

§4.3 Level Five: Autobiographical Narrative Enhancement and Moral Agency

As I have been loosely following Gennaro's model, it is worthy of note that his highest level is his Level 4—a theory of mind. Gennaro holds that “most animals are at least capable of a level 3...and perhaps even a level 4” self-consciousness.¹⁹¹ As we have seen, with some further distinctions and explication he may be correct. However, as noted at the beginning of this model (in Chapter 3), I believe his Level 4 is not rich enough, because it fails to account for the enriched narrative self-concepts that build on the theory of mind with social identity and long-term goals and projects involving oneself and others. These narrative constructions characterize adult human self-consciousness and self-concept and the mental time travel that Tulving is after with his distinctions regarding episodic and autobiographical memory. Thus, accounting for these constructions would place humans on the model and complete the model. However, some further distinctions must be made for the model to fully cohere with the moral and developmental aspects of the model I have proposed.

¹⁹¹ Gennaro, “Animals, Consciousness, and I-thoughts,” 190.

First, each level of this model has, so far, followed a pattern. From the necessary features for any experience at all, what was implicit becomes explicit with the cognitive development at higher levels, where new implicit capacities develop, which, in turn, become explicit at yet higher levels. Thus, at Level Five, a being's biography, theory of mind, and autobiographical memory become explicit. How exactly this is achieved is a matter for considerable debate. In truth, each level of this model represents explosions of complexity and diversity over millions of years of evolution—complexity and diversity so vast that it should not at all be surprising that species of the present era exhibit vastly different capabilities and have developed very differently from each other. But, it should also not be surprising that certain features have been extraordinarily advantageous and, thus, found in many species—features such as mindedness, episodic memory, the ability to learn, to feel pain, and the ability to see others as beings like oneself. Thus, the transition from Level Four to Level Five also represents an explosion of complexity and diversity, but one that is, perhaps, far more difficult to track.

This difficulty stems from the fact that if adult humans are placed at Level Five, then we must also place a myriad of new abilities at this level. For instance, we began this model with Kant's categories of the understanding, TUA, and Transcendental Synthesis of the Imagination, which are necessary features of experience, but which operate largely unconsciously. At this level, however, we humans are now able to reflect on those necessary features. Further, we can analyze and contemplate each capacity at each level of this model, we can analyze and reflect upon our own survival strategies, our personal histories, and, crucial to Kant, the maxims which guide our actions in the world, whether directed at the world or at others. Simply put, at some point between Level Four and Level Five language, reason, abstraction, the use of symbols, and reflection on deep aspects of ourselves and the world becomes possible, or deeply enriched. But

what makes this possible? I have no great uncontroversial answer to this. I can only say that the most probable explanation is language—that the ability to use systematized symbols to communicate with others allowed for exponential advances in cultural transmission, social planning and cohesion, abstract thought, and reasoning, much like changes in technology have caused exponential growth and change over comparatively short spans of time.

However, if this is right, these cognitive advances were built upon the capacities represented at the lower levels of this model. For instance, in discussing self-concept in humans, Joel W. Krueger distinguishes between Narrative Constitution Accounts (NCA) and Narrative Enhancement Accounts (NEA). NCAs, such as the views of Daniel Dennett and Marya Schechtman, advance that “The self is ultimately nothing but a dense constellation of interwoven narratives, an emergent entity that gradually unfurls from (and is thus constituted by) the stories we tell and have told about us.”¹⁹² This is the rich biographical mental life which is central to adult human self-concept. By contrast, NEA views claim that “some, but indeed not all, aspects or parts of the self are at least potentially enhanced or explicated by narratives.”¹⁹³ The debate between NCA and NEA views is about the ultimate nature of the self we take ourselves to be. As I mentioned at Level Zero, my model need not weigh in on this matter, as rich accounts of our biographical and conscious lives can be given without ever committing to the existence of any sort of substantial self. Further, if we continue to follow Kant’s philosophy, the epistemic limits he places on our cognitive powers—that we cannot fully cognize ourselves and cannot know that we are metaphysically free—means we cannot know exactly what constitutes the self we take

¹⁹² Joel W. Krueger, “The Who and the How of Experience,” in *Self, No Self?: Perspectives from Analytical, Phenomenological, and Indian Traditions*, ed. Mark Siderits, Evan Thompson, and Dan Zahavi (Oxford: UP, 2011), 37.

¹⁹³ *Ibid.*, 36.

ourselves to be. The point I wish to make is that the “dense constellation of interwoven narratives” enhances the self-conscious awareness of oneself to the point that one takes the narratives to be meaningful and actively participates in the generation and social mediation of those narratives, ultimately defining oneself by them.

I suggest that, essentially, a being that understands itself to have a biography can now explicitly reflect upon that biography and their relationships with others, attach meaning to certain subjective experiences and, in so doing, craft a more cohesive narrative for oneself. In this sense, the ascription of meaning is why it is now an *autobiography* and social interactions and immersion help develop this narrative. The difference between this autobiography and the biography at Level Four is that “summarizing, constructing, interpreting, and condensing life experiences, often smoothing over the boundaries between different moments of our lives” is now possible, but matures over time.¹⁹⁴ Indeed, as referenced above, evidence suggests that human babies must develop a theory of mind, and, to add to that, the struggle of figuring out “who am I” and “who do I want to be” is a familiar characterization of human adolescence which does not resolve (if it ever does resolve) until adult human life. This should not be surprising, as the kind of “constructed narrative self” that is a feature of adult human life is not only learned through social interaction, but probably comes via language and advanced abilities to analyze concepts. As Marian Stamp Dawkins points out regarding emotions, “We may have evolved an additional conscious verbal route that is lacking in [other animals]. Indeed the evolution of the vertebrate brain has often involved overlaying existing pathways with new ones rather than

¹⁹⁴ John Sutton, “Memory,” *Stanford Encyclopedia of Philosophy*, last modified February 3, 2010, accessed March 19, 2013, <http://plato.stanford.edu/entries/memory/>, §1.2.

eliminating existing ones.”¹⁹⁵ In this sense, our conscious verbal route allows deeper understanding of concepts, but it was also a development upon the foundation of the lower levels which are present in many non-human animals. Further, this smoothing of the gaps of the narrative is probably made possible by explicit autobiographical memory—or, to build off an example from Chapter Three, that I can not only remember those last moments with my father, but can step back from them, yet keep what I have learned from them, and tie those memories to other memories in an ongoing narrative of, for instance, my development as an adult human.

Second, as mentioned at Level Four, a being’s immersion in social interactions, and the framework necessary for doing so, seems to indicate a sort of practical sense of restriction and permissibility regarding appropriate actions in social settings. Through advanced capabilities to reflect upon concepts and an enhanced myriad of social interactions, it is likely that this practical sense developed into a moral sense—a sense comprised of shoulds and oughts that apply to generalized situations as well as specific instances with complex contexts. Tom Regan makes a distinction between moral patients and moral agents which should be helpful here. For Regan, a moral agent is an individual who has “a variety of sophisticated abilities, including...the ability to bring impartial moral principles to bear on the determination of what, all considered, morally ought to be done and, having made this determination, to freely choose or fail to choose to act as morality, as they conceive it, requires.”¹⁹⁶ The important concept here is the deployment of impartial principles, the sorts of maxims that we saw Kant concerned with in Chapter One (§1.5) and his supreme moral principle, the Categorical Imperative. On the other hand, a moral patient

¹⁹⁵ Marian Stamp Dawkins, “Animal Minds and Animal Emotions,” in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 122.

¹⁹⁶ Tom Regan, “The Case for Animal Rights,” in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 19.

lacks these abilities and, thus, cannot be held morally accountable for their actions.¹⁹⁷ In other words, for Regan, most non-human animals are moral patients while humans (at least the ones that have the requisite sophisticated abilities) are moral agents.

I agree with Regan's distinction and evaluation, to a point. The beings at Level Four or lower are not moral agents, nor are they truly rational agents by human standards. This is because there is no available evidence that these beings can reflect upon their judgments in the requisite ways and to the requisite degrees. For instance, a young silverback gorilla may challenge an older silverback and may wonder if he can successfully do so, or how committed he is to the challenge, but it is not likely that he questions whether or not social challenges are the correct way for gorillas to go about things, or thinks that all gorillas ought to think or act according to the same principles he does. Similarly, a bear that makes its way into someone's mountain home does not know that it is violating that person's rights or ends. In this same sense, we cannot consider female praying mantises barbaric for biting the head off their mate, nor can we consider non-human animals who commit filial infanticide to be doing something immoral. Simply put, there is no reflection on morality itself for moral patients, nor any conscious decision to be moral or immoral, so if non-human animals are moral patients, then we cannot hold them morally accountable for their actions, though we may still owe them moral considerability.

Yet, we cannot say that non-human animals are not moral at all, because, as we have seen, some non-human animals clearly have rules of conduct within their own ingroups. Further, there is ample ethological evidence that those rules get imported into new groups as those groups are formed. In truth, the lack of linguistic reports means we cannot know if non-human animals

¹⁹⁷ Ibid.

actually do reflect on moral principles even if they do seem to follow what look like moral principles at times. In this sense, non-human animals are not moral or rational by human standards—i.e. to the same degree that humans are, or can be. However, recall that in Chapter One (§1.5), we established that Kant thinks we cannot know when a human acts from moral principles either—i.e. we can never be sure about the maxims which guide our actions or the actions of others. We can, of course, ask humans and get cogent responses that we cannot get, or expect to get, from non-human animals due to the lack of language. This is a major historical issue that has often led to the denial of reason, autonomy, morality, self-consciousness, pain experiences, and even consciousness itself in non-human animals.

However, as Gennaro argues, and as I have tried to show in this model, the lack of linguistic ability does not mean we cannot ascribe some concepts to non-human animals, even if advanced concepts such as death, universals, or personal narratives are beyond them. Gennaro says we may reasonably attribute a concept, X, to an organism whenever the organism: 1- “systematically discriminates some Xs from some non-Xs,” 2-it is “capable of detecting some of its own discrimination errors between Xs and non-Xs,” and 3-it is capable of learning to better discriminate Xs from non-Xs as a consequence of its capacity (2).”¹⁹⁸ Recall that the animals in the studies mentioned in this model all were tested on abilities to discriminate, and in some cases to discriminate at metacognitive levels. This means we can comfortably attribute some sort of concept to these non-human animals due to their discrimination. Moreover, as Rollin notes, “the working criterion for mentation is the ability to adapt to new situations and to learn from experience,” in other words, these non-human animals must be said to be conscious and, via

¹⁹⁸ Gennaro, “Animals, Consciousness, and I-thoughts,” 199.

Gennaro, to have concepts if they exhibit learning.¹⁹⁹ Thus, the presence of learning, episodic memory, and discrimination suggests that non-human animals are conscious and do have some degree of significant conceptual understanding. Add to this the social dynamics and personal understanding which characterizes Level Four, and it seems probable that at least some non-human animals have cognitive capacities beyond Level Four, though perhaps not quite as developed as human beings at Level Five, and could, possibly, discriminate between what they think ought to be done and what they think ought not to be done.

The point of all of this is that some non-human animals may be more developed rationally, morally, and self-consciously than we know, or can possibly know. As mentioned in the discussion of Level Four, primates, cetaceans, and elephants tend to get easy passes in that many people already accept that these beings are smarter than most non-human animals and, potentially, close to on-par with at least some human beings. It is also critical to recall a point made in Chapter One (§1.6): that we cannot set the bar of moral considerability too high as that would exclude most humans. Thus, primates, cetaceans, and elephants make good candidates for beings that might be at Level Five along with humans. Further, according to Kant, while all humans have the requisite predispositions for morality, those capacities must be developed and many of us never succeed in that project.²⁰⁰ This shows that being human does not make one ethical by default, but it is through the cultivation of our autonomy and rationality that we become moral. We may further note that the epistemic limits discussed by Kant also mean we

¹⁹⁹ Bernard E. Rollin, *The Unheeded Cry: Animal Consciousness, Animal Pain, and Science*, expanded ed. (Hoboken, NJ: John Wiley & Sons, 1998), 44. Hereafter, *UC*.

²⁰⁰ This point was made in §1.6 of this thesis. For the relevant passages in Kant see his *Metaphysics of Morals*, in *Practical Philosophy*, ed. and trans. Mary J. Gregor (Cambridge: Cambridge University Press, 1996) 6:387 and 6:399. Also see Immanuel Kant, “Universal Natural History and Theory of the Heavens,” in *Natural Science*, ed. Eric Watkins, trans. Olaf Reinhardt (Cambridge: Cambridge University Press, 2012) 1:356.

cannot rule out that there are more rational, more autonomous, more moral beings than human beings—essentially, that there may be yet higher levels than Level Five to this model. If human beings are neither guaranteed to be at Level Five, nor guaranteed to be at the apex of the model, then room is open to place at least some other non-human animals at Level Five as some kind of moral agent. However, even if this idea is just too much, we still must note that it is not necessary for any non-human animals to be moral agents, or to be at Level Five of this model, in order for non-human animals to be worthy of moral considerability and, thus, the target of direct moral obligation from beings which are moral agents.

§4.4 Full Circle: Moral Considerability for Kantian Ethics Based Upon Cognitive Lives

Thus, the arguments for this model for degrees of self-consciousness have come full circle—back around to the arguments in Chapter One which established that such a model was needed. As was shown in Chapter One, on a view of Kant’s ethics we can have moral obligations to beings that cannot or do not have such obligations to us. Further, the criterion for moral obligation is intrinsic to the moral agent, but what beings that agent is obligated to is a matter of the cognitive development of the other being in terms of reason, autonomy, and self-consciousness.

What I have hoped to provide with this model is an account of when this threshold is reached. What begins with perceptual experience and contains implicit self-awareness through bodily self-consciousness and basic self-other distinctions at Level One could be built upon to then include unconscious metacognitive judgments that require unconscious memory functioning, such as certainty or confidence, at Level Two. Advanced cognitive functioning and

memory capacities might then yield conscious metacognitive judgments out of which might grow a more robust sense of self, including the ability to cognize oneself as a self. I have argued that this level, Level Three, is the threshold of moral considerability and direct moral obligations. Further considerations regarding social interactions and a theory of mind which allows one to see others as beings then may develop at Level Four, the same level where the necessary framework for the possibility of being a moral agent develops. Finally, as seen in humans, linguistic capacities and further developments in cognitive functionality might yield the ability to construct personal narratives which meaningfully deepen and enrich one's cognitive life, including a developed sense of moral and rational agency such that one recognizes and chooses to fulfill or not fulfill one's moral obligations to others.

Thus, we can, contrary to the anachronistic denial of self-consciousness in non-human animals by Kant seen in Chapter One, flesh out Kant's model of cognition and morality to include non-human animals. Further, we can place humans and non-human animals on the same scale of cognitive and moral development, rather than anthropocentrically divide the world of conscious beings to humans, angels, aliens, and spirits on one side and everything else on the other. What comes out of this model, then, is that we have direct moral obligations to non-human animals, precisely because they do have rich and meaningful cognitive lives that are worthy of our respect regardless if they can, or ever will be able to, develop into moral agents themselves. Further, we may have even more obligations to non-human animals which exhibit social behavior, as well as the non-human animals typically placed as being closest to the capacities of human beings. The remaining question is how we are to change our practices to fulfill these moral obligations.

§4.5 Subjectivity, Scientific Observation, and Parsimony

There are, of course, some objections that need to be dealt with regarding the suggested model of degrees of self-consciousness. Specifically, three objections are of immediate concern. First, my model imports the view that there are subjective qualities to experience as a fundamental feature of experience. Some either deny that there are subjective qualities to experience, or that they are fundamental, and, thus, would reject my model because I build that notion into its foundation. Second, some would posit the denial of subjective qualities to experience, or even deny the existence of conscious experience in non-human animals, based upon the belief that we cannot scientifically investigate or prove those qualities or the consciousness of non-human animals. Thus, my model purports to posit and explain animal consciousness in non-confirmable ways. Third, many people resist explanations that appeal to these qualities and consciousness in non-human animals and, instead, appeal to the principle of parsimony and posit instinct as an explanation of what appears to be conscious behavior. Thus, my model competes with a historically accepted view and cannot meet the explanatory burden already claimed by instinct. I believe these three objections are closely tied together and my replies to each one in order should be seen as building toward an ultimate reply to all three, both individually and collectively.

Further, and tying into these objections, it should be noted that my suggested model might, at times, seem like a disparate hodgepodge of theories brought together as if they unproblematically form a cohesive whole. For instance, I have referenced theories from phenomenology, ethology, psychology, and biology, ranging from inherent qualitative aspects of subjectivity, to phylogenetic continuity, to distinctions in advanced forms of memory. However, while it is true that these theories seem disparate, and without volumes of argumentation and

investigation they may not fit with a basis in Kant's theory entirely unproblematically, I believe there is a strong and plausible basis for having brought in each theory. I will tackle how I have brought in ethology and biology below, but how, and why, I have brought in theories from psychology should be readily apparent. Kant's theory of cognition seeks to explain how experience and knowledge work, and his explanation is, at its core, at least partially descriptive in regards to human minds. As we have seen, the TSI requires memory. As such, a developmental model that builds upon Kant would need to explain how memory works at other stages of development. Thus, theories from psychology ought to be welcomed by a Kantian base as long as they add to its explanatory value and do not contradict that base. Similarly, for the most part, phenomenology is a historical continuation of Kant's project and seeks to explain the structure of experience. Certainly, Kant would not agree with all phenomenologists. But, my aim in this model is not to advance what Kant *would* agree with, but to flesh out his theory with an account that he *could*, in principle, agree with. Thus, phenomenological theories that expand Kant's explanatory power without contradicting his base ought to be welcome as well. I have tried to develop the suggested model with this sort of cohesion in mind. However, that will, certainly, not be enough for some.

For instance, in discussing whether non-human animals might have consciousness similar to what humans have (i.e. subjective qualitative experience), Daniel Dennett claims that:

In order to be conscious—in order to be the sort of thing it is like something to be—it is necessary to have a certain sort of informational organization that endows that thing with a wide set of cognitive powers (such as the powers of reflection and re-representation). This sort of internal organization does not come automatically with so-called “sentience.” It is not the birthright of mammals or warm-blooded creatures or vertebrates; it is not even the birthright of human beings. It is an organization that is swiftly achieved in one species, ours, and in no other. Other species no doubt achieve *somewhat similar*

organizations, but the differences are so great that most of the speculative translations of imagination from our case to theirs *make no sense*.²⁰¹

The important issues to note in Dennett's position are that he references informational organization, wide and advanced cognitive powers, and analogous organizations in non-human animals that are, nevertheless, vastly different. I suggest that each of these points illustrates at least part of each of the three objections that I am replying to in this section. Thus, answering Dennett's points will lead us into my reply to the second objection and carry us through to the reply to the third objection.

First, as I see it, in singling out humans as capable of subjective experience, and, therefore, excluding the rest of the phylogenetic scale, Dennett's view—and other views that deny a basic subjectivity and quality to subjective experiences—either ignores Kant or evolution, and probably both. Recall that Kant's view takes informational organization and unification as necessary conditions for the possibility of experience. In other words, a being would not have any sort of cohesive experience unless that experience is filtered through the TSA and TUA. Part of this informational organization is spatiotemporal and, thus, seems to imply a certain sort of qualitative aspect that may change based upon the sense that is used. For instance, Thomas Nagel's "what-it-is-like-to-be" advances the notion that there is a subjective quality that is fundamental to experience, such as the necessary point-of-view exhibited in a bat's echolocation.²⁰² In response, critics like Dennett note that Nagel is "at best ill-advised in simply

²⁰¹ Daniel Dennett, "Animal Consciousness: What Matters and Why," in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 116, emphasis in text. Hereafter, "Animal Consciousness."

²⁰² Thomas Nagel, "What Is It Like to Be a Bat?" in *The Nature of Consciousness: Philosophical Debates*, ed. Ned Block, Owen Flanagan, and Güven Güzeldere (Cambridge: MIT Press, 1997).

assuming that a bat *must* have a point of view.”²⁰³ This sort of objection depends on what Dennett and Nagel mean by a point-of-view. On my suggested model, there is a distinction between a robust narrative point-of-view and a fundamental point-of-view inherent to experience. Nagel is after the latter and cannot be ill-advised in that assertion since a bat’s echolocation would simply not work if that bat did not have a point-of-view for the incoming information—otherwise the bat would have no idea where it is in relation to the sonic map it receives. This is the same sort of implicit self-reference, or first-personal givenness (as some phenomenologists call it), argued for in Kant’s TUA as being necessary for a perceptual experience. Further, it seems like common sense can establish that a bat’s “seeing” via echolocation would be qualitatively different than its seeing via its eyes—both in terms of the precise informational synthesis required and the experience the bat has of that represented information (otherwise, we might wonder why it has two senses that achieve the same end). Thus, Nagel’s view seems friendly to a Kantian base.

Moreover, it must be stressed that, for Kant, our experiences of the world are as the world appears to us in our representations. Since experience is fundamentally representational, it seems apparent that if there is anything it is like to experience anything, ever, then Kant seems committed to the idea that this is either captured by or a result of our representations. To put this another way, if we admit that there is a subjective aspect, an accompanying feeling, to witnessing a sunset that makes witnessing sunsets unlike witnessing other things (such as watching paint dry for instance), then this feeling must either be a feature of what we are witnessing that is then represented by our minds in experience, or a feature of how we experience our representations. Either way, the representations have a qualitative element—a “something-it-is-like” to

²⁰³ Ibid., 117, emphasis in text.

experience them—and this qualitative element must somehow be accounted for in the informational organization that makes such experiences possible. Finally, for Kant, such organization *does* come automatically with sentience, since it is required for the possibility of sentience. Thus, Dennett’s objections both to Nagel and to basic informational organization in non-human animals can be readily dismissed on Kantian grounds.

There is, however, still the issue of requiring a wide set of advanced cognitive powers and it is here that I suggest that critics like Dennett ignore evolution. Evolution by natural selection—the fundamental theory in modern biology—explains organizational complexity and diversity through phylogenetic continuity and, typically, capacities found in one species are often found, in more rudimentary form, in the species’ cousins. This continuity includes mental and behavioral capacities, including anticipatory mechanisms, learning, and emotional experiences.²⁰⁴ Scholarship, such as Bernard E. Rollin’s *The Unheeded Cry*, historically tracks the idea of a phylogenetic continuity in terms of consciousness—more specifically, the idea that the presence of conscious experiences in humans suggests the presence of conscious experiences in at least some non-human animal species.²⁰⁵ By ignoring this explanatory theory, critics like Dennett must either propose another explanatory theory, which he does not give an account of, or essentially, by implication, relegate consciousness to an epiphenomenon. As Rollin puts it, “Continuity and small variation constitute the rule in living things. If someone wishes to violate the principle of continuity and assert quantum jumps between animals, while remaining a proponent of evolution, the burden of proof is on him.”²⁰⁶

²⁰⁴ Cf. Marian Stamp Dawkins’s “Animals Minds and Animal Emotions.”

²⁰⁵ Specifically, see *UC*, Chapter 4.

²⁰⁶ Rollin, *UC*, 32.

Basically, if Dennett (or positions like his) wants to claim that subjective experience arises only in humans, then he lacks an explanatory fecundity of similar, though less developed capacities in non-human animals. Further, an evolutionary account explains the presence of the necessary components of human consciousness on Kantian accounts that are known to exist in animals—such as basic spatiotemporal awareness and capacities for memory and learning (regardless of Kant’s having denied some of them). It seems far more likely that human subjective experience builds upon capacities found on the phylogenetic scale, with differences in degree explainable by adaptations selected for to enhance the survival value of awareness—adaptations such as mindedness, increased reflectivity, metacognition, reason and all that flows from these enhanced capacities. All of these “advances” increase conscious capacities and perhaps enhance aspects of them, but are not radically different from consciousness found in other degrees. Thus, those who object to subjectivity and subjective qualities to experience existing in non-human animals seem to ignore the explanatory value of Kant’s cognitive theory and the theory of evolution, and we have no reason to abandon an explanation that works for no explanation at all (a point I shall return to in my reply to the third objection).

To move on to the second objection, some deny subjective qualities to experience or the existence of conscious experience in non-human animals based upon the lack of being able to cogently investigate or confirm them scientifically. As the objection goes, there are only so many ways available to us to investigate subjective experiences. We have no way of directly accessing another being’s mental states and, thus, can only rely upon measurable scientific data, limited observational evidence, and linguistic reports from the subjects of those experiences. Though some scientists have devised ways to attempt to test non-human animals for advanced cognitive abilities (such as mirror tests and betting trials), objectors are often skeptical either because they

see ascriptions of advanced cognitive abilities as anthropomorphizing or because the tests are questionable in what they really establish. This doubt is often supported by the lack of linguistic reports from non-human animals. When coupled with a principle of parsimony (which I will discuss in the next reply) such as Occam's Razor, or Morgan's Canon, this means we ought not attribute conscious experiences to non-human animals, or, at least, ought not attribute any more advanced capabilities than absolutely necessary.

In truth, there is a lot to this objection. As the combination of science and philosophy of mind tell us, we do not really know exactly what consciousness is, nor what causes it, which makes it impossible to test, which, in turn, ought to make us cautious about ascribing it. Further, many people, including some ethologists, do anthropomorphize when it comes to explaining the behavior of non-human animals. However, this sort of objection also proves too much. Recall from §1.5 that by Kant's philosophy one cannot know that one is truly free and, thus, one cannot know which other beings are free either. In philosophy of mind there is a similar problem called the problem of other minds. Essentially, since I cannot directly access the mind of another being, I cannot know that another being is conscious, I can only know that I am conscious. However, this problem also applies to humans in the form of the question "how do I know that other humans have minds, like I do, and are not automatons?"

There are two usual strategies to escape this problem. First, one can appeal to linguistic reports. Lack of language has been a historically deployed criteria to deny mentation in non-human animals, yet affirm it in humans. But, *prima facie*, only humans can possibly meet this criteria, which makes it question begging since we are trying to establish if non-humans can be conscious. Further, linguistic reports do not actually escape the problem, as even if we restrict the question to humans we still face the problem that only humans capable of speaking or writing

could possibly pass the test (so the illiterate and mute cannot be conscious) and, moreover, we can ask whether the being in question is really a sophisticated machine capable of mimicking a native speaker.²⁰⁷ Second, one can appeal to analogous capacities. Another human being is so much like me that it seems to license an ascription of consciousness to another human.

Presumably, this is how Kant answers the problem of other minds (though he never actually weighs in on the issue directly).²⁰⁸ However, recall (from §1.5) that Kant suggests that what capacities non-human animals have are analogues to what humans have. Thus, the analogy strategy to escape the problem of other minds raises one of Dennett's concerns: that what seems analogous between humans and non-human animals is actually vastly different. The question is whether or not they really are vastly different, or if resistance like Dennett's sets the bar for comparison far too high. Clearly, we do not expect gorillas to understand books, or dolphins to do logical proofs, or dogs to "give in to reason," but does the absence of these capacities and the command over language really suggest a vast disparity between the capacities of humans and non-human animals? I suggest this is only the case if we, again, ignore evolutionary theory.

As noted above, Rollin tracks the idea of phylogenetic continuity of conscious experience and, in so doing, he establishes that Darwin and Romanes both believed that the continuity of biological traits also suggested continuity in psychological traits. To build upon the arguments above, given continuity and small variation, convergent developments of biological traits in multiple strains of biological ancestry—for example, visual perception—, and the enhanced

²⁰⁷ For another version of a similar point, see Alan Turing's "Computing Machinery and Intelligence" (or the Turing Test) and John Searle's response in "Minds, Brains and Programs" (or the Chinese Room Argument).

²⁰⁸ For this point, see Steve Naragon, "Reason and Animals: Descartes, Kant, and Mead on the Place of Humans in Nature," PhD diss., University of Notre Dame, 1987, re-formatted 190 p draft, accessed October 20, 2013, <http://users.manchester.edu/facstaff/ssnaragon/Naragon/Research.html>. Hereafter, "Reason and Animals." Also, Naragon points to a passage in Kant's lectures, specifically, see "*Metaphysik Herder*," in *Lectures on Metaphysics*, ed. and trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997), 28:116.

survival value of the traits noted above, it would be surprising to *not* find some, or all, of these traits distributed among the phylogenetic community rather than solely appearing in humans as some sort of epiphenomenon. Further, it is likely that the advanced capacities of humans were further developments in complexity on top of the cognitive capacities of other beings. While it is true that evolution does exhibit explosions of complexity, the explosions are often widespread, such as the development of a central nervous system in vertebrates, or the development of the cerebral cortex in mammals, and leads to further complexity. But, as Rollin notes, this further complexity comes in small variations not quantum leaps. Thus, though we can say that the development of language or the advanced cognitive capacities of humans do represent a new explosion of complexity, the precursors for that explosion most likely exist in at least some non-human animals. Thus, to conclude that the disparity between human capacities and non-human capacities is “so great that most of the speculative translations of imagination from our case to theirs *make no sense*,” as Dennett holds, is to ignore biology and evolutionary theory.²⁰⁹

The same points about ignoring evolution and setting the bar of comparison too high applies to arguments such as those made by Bob Bermond. Bermond maintains that consciousness and experiences of suffering require a prefrontal cortex (PFC)—a “higher order association area” in the brain that acts as a “top down control” over “evolutionarily older neural structures,” but “limits tremendously the number of species in which consciousness may be expected” due to it being a relatively recent development in brain structure.²¹⁰ Bermond supports this view with claims that self concept, planned behavior, a distinction between pain experience

²⁰⁹ Dennett, “Animal Consciousness,” 116, emphasis in text.

²¹⁰ Bob Bermond, “A Neuropsychological and Evolutionary Approach to Animal Consciousness and Animal Suffering,” in *The Animal Ethics Reader*, 2nd ed., ed. Susan J. Armstrong and Richard G. Botzler (London: Routledge, 2008), 107.

and pain emotion, prolonged duration of pain-like sensory inputs, “reflection, imagination and understanding of ‘duration’, of past and future” are required for consciousness and suffering and that these are absent in most animals.²¹¹ The most interesting aspect of Bermond’s argument is the tie of higher-order informational organization to brain structures—a fairly common tactic for resisting ascriptions of mental capacities. However, it is also interesting that, strangely, at various points in his argument, Bermond relies not only on there being a qualitative aspect to the consciousness and suffering of humans, but that we can detect higher-order functioning in observable behavior.²¹²

My suggested model of degrees of self-consciousness takes qualitative aspects to experience (or subjective aspects, as I have called them) as basic, and, as discussed above, I believe Kant does as well. So the question for Bermond regarding qualitative aspects is if only humans have qualitative aspects to experience, because these qualities only arise with brain structures such as the PFC, or if the qualitative aspects of human experience are vastly different from those of non-human animals. If the former, then Bermond has far more to prove than he thinks, as why would the PFC be necessary for there to be a qualitative aspect to perception, such as the echolocation of bats discussed above? Moreover, one need only point to octopi for an example of a non-human animal that exhibits complex behavior suggesting a high degree of organization, intelligence, and conscious responsiveness to its environment—an animal which has a nervous system fundamentally different than our own. If the latter, however, then Bermond is positing something that may be impossible to prove in the absence of linguistic reports and the

²¹¹ Ibid., 103.

²¹² For instance, on page 109 of the article cited above, Bermond claims that “Consciousness may therefore only be expected in animals which show these higher cognitive capacities in their behavior.”

absence of being able to empirically measure qualitative aspects of experience, as how are we to know that there really is a vast difference in qualitative experience?

Further, the remainder of Bermond's considerations can be discredited easily by a large volume of scientific evidence of the presence of these characteristics in animals. For instance, against Bermond's suggestion that emotion is a necessary criterion, and that emotion requires a prefrontal cortex, there is a vast body of evidence that many non-human animals do experience emotions. Moreover, I must, again, point to phylogenetic continuity versus singling out human life and experience. It does not follow, logically or biologically, that features of human experience, emotion, and brain structure have no significant precursors in the phylogenetic community—precursors implied by a capacity for experience itself and a full range of complexities which suggests evolutionary development and, in turn, phylogenetic continuity. Similarly, Bermond's reference to reflection and a sense of future and past has been accounted for on my suggested model with reference to numerous behavioral studies.

Thus, as I have argued against this second objection, the typical resistances to ascriptions of higher cognitive functioning in non-human animals on grounds of scientific investigation and proof fails. Appeals to brain structure, to analogous yet vastly different capacities, or to empirical measurability set the bar too high and anthropocentrically begs the question when we decide that specifically human faculties are necessary criteria for consciousness or subjective aspects of experience, yet exempt the need to prove, by the same lights, the same capacities in humans. Of course, it is true that the lack of linguistic reports makes the differences between humans and non-human animals seem large, especially since language is at the forefront of human reason. But is this enough to deny the analogies and similarities evidenced by evolutionary theory and observation? I believe it is not. When we observe the behavior of non-human animals we are

faced with a choice: we can leave the explanatory gaps unfilled, or we can fill them with explanations, hopefully ones that meet our highest standards. In the absence of linguistic reports, and in the need for some sort of explanation, I have tried to show that linguistic reports are an insufficient criterion, as is pointing to human capacities as drastically different from non-human animal capacities. As such, we are left with empirical investigation (such as neuroscience) and observation (such as ethology) to establish the capacities of non-human animals. Both methods require explanatory theories, which, I have argued, we cannot burden with insufficient or arbitrary criterion to bar us from certain conclusions—namely, that non-human animals have rich, conscious, cognitive lives.

But, how are we to explain away the persistent resistance to ascribing conscious subjective states to non-human animals in empirical investigation and observation? To follow Rollin a bit further, the logical positivism and behaviorism that dominated much of early twentieth century science and philosophy has left a specter over considerations of consciousness and subjective experience in non-human animals in that it relegated “consciousness-talk” to mystical nonsense because it lacked empirical measurability.²¹³ Essentially, the positivist position is that if something cannot be reduced to physical, “natural,” empirically verifiable terms, then, depending on the exact position, it is nonsense, fiction, illusion, or, at least, not worth being taken as anything above and beyond whatever physical states they answer to. For instance, a positivist might explain pain experience as certain C-fibers firing, or happiness as the presence of certain neurochemicals in one’s system. Terms such as “consciousness” or

²¹³ In Chapter 4 of his, *The Unheeded Cry*, Rollin tracks this denial from Darwin and Romanes through the supposed fathers of behaviorism, suggesting that behaviorism ought to be considered a speed bump along a previously dominant historical view that not only ascribes consciousness to humans, but also, through logic, common sense, and scientific theory, to many animals on the phylogenetic scale as well.

“subjective quality” do not have ready explanations in terms of neurology, brain-states, physics, or chemistry in that we do not know how to explain the arising of these phenomenon (such as consciousness) or what explanations we do have do not capture the subjective quality, and, thus, the subjective quality is taken as something superfluous.

Any student of the history of science would recognize this ideology, though perhaps in different terms. The figures of the Age of Enlightenment fought long and hard to free scientific method from religious authority and dogma, as well as from superstition and supernatural explanations. As seen in §2.2, Kant joins in such an ideology when he warns against our tendency to use reason to engage in speculative metaphysics and claim things about the world beyond all possible experience. Similarly, the positivists sought to eject non-confirmable hypotheses from proper science. Since consciousness and subjective qualities to experience are not empirically confirmable, the position goes, they ought not be used in explanations. All that should be used are explanations that appeal to properties that we can test. Of course, science is supposed to be open-minded, so a positivist ideology could give ground, but only on the terms that such explanations meet a burden of proof cashed out in the rigorous criteria of verifiability and falsifiability—criterion whose bar has often been set too high, like in the considerations above.

There are two main problems with this view. First, who is to say this is the correct way to do science regarding consciousness and experience? On the one hand, perhaps the most readily confirmable aspect of anyone’s own experience, as Descartes showed, is that one exists and is consciously thinking. Yet, according to the positivist view, one cannot even say this, as it is as equally non-confirmable as any other being’s consciousness or thinking and, as seen above regarding the problem of other minds, extrapolating from one’s own experience to the experience

of others is problematic. On the other hand, the naturalism at the heart of the positivist view is evaluative regarding how science ought to be done. But this evaluation is not confirmable empirically. Thus, the evaluation would not belong in science in light of the very view that says it is how science ought to be done. Second, and more importantly, the positivist way of thinking denies the work of many ethologists, and denies the ability to rely on the observable behavior of non-human animals due to drastically limiting the explanatory power of common sense. Logic, common sense, and scientific theory may grant consciousness and high cognitive functionality to many animals on the phylogenetic scale based upon one well-grounded and specific reason: we ought to reject any explanation that does not actually explain what *prima facie* needs to be explained. This point brings us to my response to the third objection regarding parsimony.

The typical deployment of a principle of parsimony as an objection to a competing explanation claims that we ought not accept any explanation that utilizes superfluous terms or overly complex and roundabout hypotheses. For instance, once we have a theory of gravity we no longer need to explain planetary motion by appeal to God, and once we have a model of elliptical planetary orbits we no longer need to appeal to complex epicycles. Occam's Razor is, perhaps, the most famous principle of parsimony (and one which Kant would have been familiar with, as suggested in Chapter One) and suggests that we ought not unnecessarily multiply explanatory entities—a suggestion usually seen as appealing to simplicity and against redundancy. But, Morgan's Canon provides a more relevant formulation for what concerns us here. According to Morgan's Canon, we ought not interpret a behavior or action as being caused by a higher cognitive faculty when we can explain it in virtue of a lower cognitive faculty.²¹⁴

²¹⁴ For a historical discussion of Morgan's Canon and its application to non-human animals, see Rollin's *UC*, Chapter 4.

Under a positivist view, Morgan's Canon explains what looks like complex conscious behavior in non-human animals as unconscious or minimally conscious, or a product of conditioning or instinct.

As we saw in Chapter One, Kant makes a similar appeal regarding non-human animals. To quote the same passages, Kant says "animals are an analogue of humanity... the acts of animals arise out of the same *principium* from which human actions spring, and the animal actions are analogues of this"²¹⁵ and, "The principle that guides animals, as analogue of reason... is called instinct, the faculty for performing actions without consciousness, for which human beings require consciousness."²¹⁶ Apparent rational, social, or autonomous behavior is relegated to a rudimentary analogue of complex developed capacities in humans. This leads Kant to deny that non-human animals possess consciousness, understanding, apperception, or judgment and, instead, to explain what appears like these capacities by appeal to instinct.²¹⁷ As we saw above, more modern denials appeal to necessary brain structures, to the presence of language, emotion, reflection, or a sense of future and past. Further, those denials are predicated on scientific verifiability and investigation. In the absence of those criteria whatever remains unexplained in non-human animals is deemed explicable by unthinking instinct or unconscious response.

²¹⁵ Immanuel Kant, "Moral Philosophy: Collins's Lecture Notes," in *Lectures on Ethics*, ed. Peter Heath and J.B. Schneewind, trans. Peter Heath (Cambridge: Cambridge University Press, 1997), 27:459, emphasis in text. Hereafter, *LE (Collins)*.

²¹⁶ Immanuel Kant, "*Metaphysik Dohna*," in *Lectures on Metaphysics*, ed. and trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997), 28:690. Hereafter, *LM (Dohna)*.

²¹⁷ See, for instance, Kant, *LE (Collins)*, 27:267 and 27:459; *LM (Dohna)*, 28:690; "Kant on the Metaphysics of Morals: Vigilantius's Lecture Notes," in *Lectures on Ethics*, ed. Peter Heath and J.B. Schneewind, trans. Peter Heath (Cambridge: Cambridge University Press, 1997), 27:710; "*Metaphysik Mrongovius*," in *Lectures on Metaphysics*, ed. and trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997), 29:907. Also see Naragon, "Reason and Animals," 42.

Herein lies the problem and we are now in a position to see why, as well as how my model purports to answer this objection given the resources discussed. Essentially, a principle of parsimony does not require us to make no assumptions at all. Background assumptions are a fundamental cornerstone of science. Kant most certainly knew this, as part of the reason he posits the categories of the understanding is as a response to the arguments of David Hume—who suggested that we never actually see the causes of things, we only see two things customarily brought together in experience and infer that one causes the other, where our inference is based upon a presumed uniformity in nature. In other words, we assume causality and regularity in order to even do science and Hume’s arguments seemed to threaten that foundation. Kant’s solution was to suggest that uniformity and the notion of causation were necessary features of our minds and cognition—features necessary for the very possibility of having experiences and that we bring to experiences. In modern science, we still assume uniformity and causation—for instance, that the laws of gravity hold everywhere in the Universe and will continue to hold tomorrow—despite not yet being able to prove it, nor having yet isolated a brain feature that specifically accounts for either how we bring uniformity and causality to experience or how we detect those things about the world.

Of course, we must be careful what background assumptions we license, but what could possibly be more of a cornerstone, more of a necessity, more of a fundamental assumption than that consciousness exists? If it did not, we would not even be able to investigate the world, nor ask whether or not consciousness exists, nor make any assumptions at all. As noted above, it is a problem for more modern forms of resistance to take human consciousness as granted—only individual consciousness is provable, and that only to the individual. Further, it is problematic to deny consciousness to non-human animals when we consider that we must explain how features

in humans arose through evolutionary precursors and that so many features, both biological and psychological, are distributed on the phylogenetic scale. It defies evolutionary theory to make consciousness, mindedness, reason, autonomy, self-consciousness, emotions, and social thinking an epiphenomenon of humans beings, even if, and especially because, we do not yet know what caused them to develop.

However, this is an even more problematic assertion for Kant to make, as the theory of cognition which he provides—which also claims that rationality, autonomy, and self-consciousness may exist in angels and aliens rather than in creatures in our backyards—gives absolutely no account of how cognition could possibly work given what Kant denies to non-human animals. In other words, Kant claims that the faculty of understanding, the categories of the understanding, and apperception (among other things) are necessary features for the possibility of human experience. In the absence of those same things, which he denies to non-human animals, how are we to explain that they experience anything at all (as he grants that they do)? The point is that, for Kant, instinct has no explanatory value beyond a posited barrier between cognition in humans and unconscious perception in non-human animals. He provides no deeper explanation of what instinct is beyond an analogue of reason that looks like reason, but is not reason. As an argument from parsimony, then, Kant's denial resists granting more complex capacities that might explain observable features of non-human animal life, by positing a mysterious and unexplained faculty that does not really explain anything at all.

Much like modern resistances are often anthropocentric, question-begging ways of setting the comparative bar between humans and non-human animals too high, positing instinct is an empty promissory note—it suggests an explanation that it never actually gives. What looks like reason is not reason because it does not involve the understanding and judgments, just as

what looks like conscious responses, learning, and complex behavior is not actually those things because non-human animals lack language or brain structures. Further, like with Bermond's possible suggestion that human experience is qualitatively different than non-human animal experience, we may wonder if instinct can ever explain what it needs to given all the problematic criteria noted above that is asserted by those who resist granting conscious experiences to non-human animals. In essence, then, instinct is a theoretical conclusion that results from a predilection to resist other possible explanations—or an attempted application of the principle of parsimony to prevent undue assumptions.

This leads to a second point about background assumptions and a principle of parsimony, which I noted above: we ought to reject any explanation that does not actually explain what needs to be explained. Positivists, and their legacy, try to avoid this by denying that there is anything to be explained (i.e. consciousness does not exist) and Bermond, Dennett, and Kant posit empty distinctions that do not actually explain the observable complex behavior in non-human animals. In essence, and to keep with the comparisons to astronomy, following their thinking would be like suggesting that the forces and causes of planetary orbits around the Sun do not actually apply to a planet's satellites due to their not having their own orbits around the Sun. Or, if one prefers, it would be like saying that Galileo's original telescope did not function on the same principles as modern telescopes because his did not have as much light-gathering capacity. In the absence of explaining two instances with the same principle, we would actually be defying parsimony by suggesting that what holds in one case is not the same as what holds in the other case. This defiance goes further when we explain the other case by reference to a completely different theory, and even further when that different theory does not sufficiently explain the case when the other theory would. Or, in other words, if gravity explains the orbits of

both planets and satellites, then a separate theory for satellites is superfluous. Similarly, if reason and advanced cognitive faculties can explain observable behavior and empirical investigation of both human and non-human animals, then instinct is superfluous.

Perhaps I am being unfair to proponents of instinct. I do not believe I am and for one reason: if instinct is a viable explanation for the behaviors and capacities of beings on the same phylogenetic scale as humans, why would we not use it to explain human behavior and capacities? In other words, why is language not instinctual? Why is human learning, emotions, and culture not subject to the exact same explanations as that in non-human animals? Is it possible that we could explain all human behaviors in terms of instinct? If not, why not? Or, to borrow a page from Kant's suggestions of aliens and angels, it seems entirely possible that beings far more developed in their rational capacities than humans may exist and would describe all human characteristics by appeal to lower faculties, perhaps even to instinct. With the epistemic limits of Kant's philosophy and the problems of scientific investigation and confirmation noted in the objections above, we do not actually know ourselves well enough to establish that we really are all that different than non-human animals. We think we are, but we do not know that we are. Humans do not get a free pass, they do not get to escape the fundamental explanatory power of evolution, nor do they, as seen in the problem of other minds, get to help themselves by comparison to other humans. In other words, explaining the behavior of all known beings in terms of instinct would preserve parsimony and evolutionary theory, though it would not preserve how we wish to think about ourselves.

Alternatively, we can do as my model suggests and put humans and non-human animals on the same evaluative scale in terms of the same capacities for reason, autonomy, and self-consciousness. Doing so allows us to, I believe, affirm basic observation, common sense, logic,

evolutionary theory, and the principles set forth by Kant which form the foundation of his cognitive and ethical theories. We can, in other words, grant that what looks like consciousness and self-consciousness *actually is* consciousness or self-consciousness, but hold that these capacities admit of degrees—differences of the same sort of capacity rather than of different kinds. This allows us to both avoid multiplying explanatory theories and terms, while also licensing us to utilize theories that seem to work in explaining human behavior and, due to the notion of degrees, extend that explanatory fecundity to the rest of the phylogenetic community. As I put this in Chapter One, this allows us explain more—about ourselves, about non-human animals, and about the relationship between the two.

Need we worry that humans become less special on this model? No, we are still the most advanced beings we have found to date. Need we allow anthropomorphic ascriptions of advanced capacities to non-human animals, such as that a dog actually understands language? No. In fact, because my model suggest degrees we can avoid such ascriptions by utilizing parsimony in the form of Morgan’s Canon and suggest that while the dog does not “understand” or does not “know” it, nevertheless, is consciously responding to something and deploying some sort of understanding, though certainly not one that equivocates between basic understandings and full-blown human reason. In this sense, then, my model allows for the explanatory value of evolution and phenomenology, and the methods of psychology and ethology, while allowing us to avoid anthropomorphism and abandon anthropocentrism. As such, I believe it is the more parsimonious model, and the more explanatorily fecund model. Further, its use of Kant as a base allows us to move from discussions of cognition to discussions of moral considerability based upon the same theories and the same explanatory model. Might I be mistaken about what constitutes some of the levels of my model, or what creatures can be mapped to each level? Of

course. But, my model also provides its own grounds for investigating, criticizing, explaining, and elaborating on itself—as we can always, with the same criteria that generated the model, move capacities or beings up or down the model given sufficient argument and evidence. If such built-in explanatory fecundity and room for elaboration and criticism is not a desirable feature for a parsimonious scientific model, then I do not know what is.

REFERENCES

- Beck, Lewis White. *Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press, 1960.
- Bekoff, Marc and Jessica Pierce. *Wild Justice: The Moral Lives of Animals*. Chicago: University of Chicago Press, 2009.
- Bermond, Bob. "A Neuropsychological and Evolutionary Approach to Animal Consciousness and Animal Suffering." In *The Animal Ethics Reader*, 2nd ed., edited by Susan J. Armstrong and Richard G. Botzler, 99-112. London: Routledge, 2008.
- Brook, Andrew. "Kant's View of the Mind and Consciousness of Self." *Stanford Encyclopedia of Philosophy*. Last modified January 22, 2013. Accessed March 14, 2013. <http://plato.stanford.edu/entries/kant-mind>.
- Clayton, Nicky, Nathan Emery, and Anthony Dickinson. "The Rationality of Animal Memory: Complex Caching Strategies of Western Scrub Jays." In *Rational Animals*, edited by Susan Hurley and Matthew Nudds, 197-216. New York: Oxford University Press, 2006.
- Dawkins, Marian Stamp. "Animal Minds and Animal Emotions." In *The Animal Ethics Reader*, 2nd ed., edited by Susan J. Armstrong and Richard G. Botzler, 120-125. London: Routledge, 2008.
- Denis, Lara. "Kant's Conception of Duties Regarding Animals: Reconstruction and Reconsideration." *History of Philosophy Quarterly* 17, no. 4 (October 2000): 405-423.
- Dennett, Daniel. "Animal Consciousness: What Matters and Why." In *The Animal Ethics Reader*, 2nd ed. Edited by Susan J. Armstrong and Richard G. Botzler, 113-119. London: Routledge, 2008.
- Foot, Philippa. "Morality as a System of Hypothetical Imperatives." *The Philosophical Review* 81 (1972): 305-316.
- Ganeri, Jonardon. *The Self: Naturalism, Consciousness, and the First-Person Stance*. Oxford: Oxford University Press, 2012.
- Gennaro, Rocco J. "Animals, Consciousness, and I-thoughts." In *The Philosophy of Animal Minds*. Edited by Robert W. Lurz, 184-200. Cambridge: Cambridge University Press, 2009.
- Gómez, Juan Carlos. "Are Apes Persons? The Case for Primate Intersubjectivity." In *The Animal Ethics Reader*, 2nd ed., edited by Susan J. Armstrong and Richard G. Botzler, 169-174. London: Routledge, 2008.
- Gordon, James and Israel Abramov. "Color Vision." In *The Blackwell Handbook of Sensation and Perception*, edited by E. Bruce Goldstein, 92-127. Malden, MA: Blackwell, 2001.

- Grandin, Temple. "Behavior of Slaughter Plant and Auction Employees Toward the Animals." *Anthrozoos* 1, no. 4 (1988): 205-213.
- Hampton, Robert R. "Can Rhesus Monkeys Discriminate Between Remembering and Forgetting." In *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, edited by Herbert S. Terrace and Janet Metcalfe, 272-295. New York: Oxford University Press, 2005.
- Holland, Jennifer S. *Unlikely Friendships: 47 Remarkable Stories from the Animal Kingdom*. New York: Workman Publishing Company, 2011.
- Johnson, Robert N. "Duties to and regarding others." In *Kant's "Metaphysics of Morals": A Critical Guide*, edited by Lara Denis, 192-209. Cambridge: Cambridge University Press, 2010.
- Kain, Patrick. "Duties Regarding Animals." In *Kant's "Metaphysics of Morals": A Critical Guide*, edited by Lara Denis, 210-233. Cambridge: Cambridge University Press, 2010.
- Kant, Immanuel. *Critique of Practical Reason*. In *Practical Philosophy*. Edited and Translated by Mary J. Gregor, 134-271. Cambridge: Cambridge University Press, 1996.
- Critique of Pure Reason*. Edited and Translated by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press, 1998.
- Groundwork of the Metaphysics of Morals*. In *Practical Philosophy*. Edited and Translated by Mary J. Gregor, 38-108. Cambridge: Cambridge University Press, 1996.
- "Kant on the Metaphysics of Morals: Vigilantius's Lecture Notes." In *Lectures on Ethics*. Edited by Peter Heath and J.B. Schneewind. Translated by Peter Heath, 249-452. Cambridge: Cambridge University Press, 1997.
- Metaphysics of Morals*. In *Practical Philosophy*. Edited and Translated by Mary J. Gregor, 354-603. Cambridge: Cambridge University Press, 1996.
- "Metaphysik Dohna," In *Lectures on Metaphysics*. Edited and Translated by Karl Ameriks and Steve Naragon, 357-394. Cambridge: Cambridge University Press, 1997.
- "Metaphysik Herder," In *Lectures on Metaphysics*. Edited and Translated by Karl Ameriks and Steve Naragon, 3-18. Cambridge: Cambridge University Press, 1997.
- "Metaphysik Mrongovius." In *Lectures on Metaphysics*. Edited and Translated by Karl Ameriks and Steve Naragon, 109-288. Cambridge: Cambridge University Press, 1997.
- "Moral Philosophy: Collins's Lecture Notes." In *Lectures on Ethics*. Edited by Peter Heath and J.B. Schneewind. Translated by Peter Heath, 37-222. Cambridge: Cambridge University Press, 1997.
- Physical Geography*. In *Natural Science*. Edited by Eric Watkins. Translated by Olaf Reinhardt, 434-679. Cambridge: Cambridge University Press, 2012.

- “Universal Natural History and Theory of the Heavens.” In *Natural Science*. Edited by Eric Watkins. Translated by Olaf Reinhardt, 191-308. Cambridge: Cambridge University Press, 2012.
- Kellert, Stephen R. and Alan R. Felthous. “Childhood Cruelty toward Animals among Criminals and Noncriminals.” *Human Relations* 38, no. 12 (December 1985): 1113-1129.
- Kitcher, Patricia. *Kant’s Transcendental Psychology*. New York: Oxford University Press, 1990.
- Krueger, Joel W. “The Who and the How of Experience.” In *Self, No Self?: Perspectives from Analytical, Phenomenological, and Indian Traditions*, edited by Mark Siderits, Evan Thompson, and Dan Zahavi, 27-55. New York: Oxford University Press, 2011.
- Lee, Barry B. “Visual Pathways and Psychophysical Channels in the Primate.” *J Physiol* 589, no.1 (December 2010): 41-47. Accessed June 13, 2013. <http://onlinelibrary.wiley.com/doi/10.1113/jphysiol.2010.192658/pdf>.
- MacKenzie, Matthew D. “The Illumination of Consciousness: Approaches to Self-Awareness in the Indian and Western Traditions.” *Philosophy East and West* 57, no. 1 (January 2007): 40-62.
- Menzel, Charles. “Progress in the study of chimpanzee recall and episodic memory.” In *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, edited by Herbert S. Terrace and Janet Metcalfe, 188-224. New York: Oxford University Press, 2005.
- Metcalfe, Janet and Lisa Son. “Anoetic, Noetic, and Auto-noetic Metacognition.” In *The Foundations of Metacognition*, edited by Michael J. Beran, Johannes Brandl, Josef Perner, and Joëlle Proust, 289-301. Oxford: Oxford University Press, 2012.
- Nagel, Thomas. “What Is It Like to Be a Bat?” In *The Nature of Consciousness: Philosophical Debates*. Edited by Ned Block, Owen Flanagan, and Güven Güzeldere, 519-528. Cambridge: MIT Press, 1997.
- Naragon, Steve. “Kant on Descartes and the Brutes.” *Kant-Studien* 81, no. 1 (January 1990): 1-23.
- “Reason and Animals: Descartes, Kant, and Mead on the Place of Humans in Nature.” PhD diss., University of Notre Dame, 1987. Re-formatted 190 p draft. Accessed October 20, 2013. <http://users.manchester.edu/facstaff/ssnaragon/Naragon/Research.html>.
- O’Neill, Onora. “Consistency in Action.” In *Ethical Theory 2: Theories About How We Should Live*, edited by James Rachels, 256-281. Oxford: Oxford University Press, 1998.
- Orwell, George. *Animal Farm: A Fairy Story*. Boston: Houghton Mifflin Harcourt, 2009.
- Picard, Laurence, Isméry Reffuveille, Frances Eustache, and Pascale Piolino. “Development of Auto-noetic Autobiographical Memory in School-Age Children: Genuine Age Effect or Development of Basic Cognitive Abilities?” *Consciousness & Cognition* 18, no. 4 (December 2009): 864-76.

- Piolino, Pascale, Béatrice Desgranges, David Clarys, Bérengère Guillery-Girard, Laurence Taconnat, Michel Isingrini, and Francis Eustache. "Autobiographical Memory, Auto-noetic Consciousness, and Self-Perspective in Aging." *Psychology & Aging* 21, no. 3 (October 2006): 510-25.
- Regan, Tom. "The Case for Animal Rights." In *The Animal Ethics Reader*, 2nd ed., edited by Susan J. Armstrong and Richard G. Botzler, 19-25. London: Routledge, 2008.
- Rollin, Bernard E. *Animal Rights & Human Morality*, 3rd ed. Amherst, NY: Prometheus Books, 2006.
- "There is Only One Categorical Imperative." *Kant-Studien* 67, no. 1-4 (January 1976): 60-72.
- *The Unheeded Cry: Animal Consciousness, Animal Pain, and Science*. Expanded ed. Hoboken, NJ: John Wiley & Sons, 1998.
- Rosenthal, David. *Consciousness and Mind*. New York: Oxford University Press, 2005.
- Sellars, Wilfrid. "...this I or He or It (The Thing) Which Thinks..." *Proceedings and Addresses of the American Philosophical Association* 44, (1970 -1971): 5-31. Accessed November 30, 2013. <http://www.jstor.org/stable/3129681>.
- Singer, Peter. (1974). "All Animals are Equal." In *Ethics in Practice*, 3rd ed., edited by Hugh LaFollette, 171-180. Malden, MA: Blackwell Publishing, 2007.
- Steinbeck, John. *Of Mice and Men*. New York: Penguin Books, 1993.
- Stone, Valerie, Simon Baron-Cohen, and Robert T. Knight. "Frontal Lobe Contributions to the Theory of Mind." *Journal of Cognitive Neuroscience* 10, no. 5 (September 1998): 640-656.
- Sutton, John. "Memory." *Stanford Encyclopedia of Philosophy*. Last modified February 3, 2010. Accessed March 19, 2013. <http://plato.stanford.edu/entries/memory/>.
- Tulving, Endel. "Episodic Memory and Auto-noesis." In *The Missing Link in Cognition: Origins of Self-Reflective Consciousness*, edited by Herbert S. Terrace and Janet Metcalfe, 3-56. New York: Oxford University Press, 2005.
- "Memory and Consciousness." *Canadian Psychology* 26, no. 1 (January 1985): 1-12.
- Varner, Gary E. *Personhood, Ethics, and Animal Cognition: Situating Animals in Hare's Two Level Utilitarianism*. New York: Oxford University Press, 2012.
- "Personhood, Memory, and Elephant Management." In *Elephants and Ethics: The Morality of Coexistence*, edited by Christen Wemmer and Catherine Christen, 41-68. Johns Hopkins University Press, 2008.
- Walker, Ralph C.S. "Kant and Transcendental Arguments." In *The Cambridge Companion to Kant and Modern Philosophy*, edited by Paul Guyer, 238-268. Cambridge: Cambridge University Press, 2006.

Whitehead, Hal, Luke Rendell, Richard W. Osborne, and Bernd Würsig. "Culture and Conservation of Non-humans with Reference to Whales and Dolphins: Review and New Directions." In *The Animal Ethics Reader*, 2nd ed., edited by Susan J. Armstrong and Richard G. Botzler, 181-192. London: Routledge, 2008.

Zahavi, Dan. "The Experiential Self: Objections and Clarifications." In *Self, No Self?: Perspectives from Analytical, Phenomenological, and Indian Traditions*, edited by Mark Siderits, Evan Thompson, and Dan Zahavi, 56-78. New York: Oxford University Press, 2011.