

HUMAN BEINGS AND THE MORAL LAW: MORAL PRECARIOUSNESS IN
KANT'S ETHICAL PHILOSOPHY

Bradley M. Taylor

A DISSERTATION

in

Philosophy

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2014

Supervisor of Dissertation

Paul Guyer, Jonathan Nelson Professor of Humanities and Philosophy, Brown University

Graduate Group Chairperson

Michael Weisberg, Associate Professor of Philosophy, University of Pennsylvania

Dissertation Committee

Elisabeth Camp, Associate Professor of Philosophy, Rutgers University

Karen Detlefsen, Associate Professor of Philosophy and Education, University of

Pennsylvania

HUMAN BEINGS AND THE MORAL LAW: MORAL PRECARIOUSNESS IN
KANT'S ETHICAL PHILOSOPHY

COPYRIGHT

2014

Bradley Martin Taylor

This work is licensed under the
Creative Commons Attribution-
NonCommercial-ShareAlike 3.0
License

To view a copy of this license, visit

<http://creativecommons.org/licenses/by-nc-sa/2.0/>

This dissertation is dedicated to my parents,

Martin Joseph Taylor

and

Patricia Heeran Taylor

ACKNOWLEDGMENT

I would like to take this space to extend my thanks to all the professors and mentors that have contributed to my development as a student of philosophy. The excellent examples that were set for me by these teachers and scholars provided me with an ideal toward which I aspire in my own teaching and scholarship.

I would also like to thank my fellow graduate students and colleagues at the department of philosophy at the University of Pennsylvania. I have been both inspired and influenced by your own work. In particular, I would like to thank Reed Winegar for all the incredibly insightful conversations and discussions about Kant and Hegel we had over the years, especially during the early years of my time at Penn.

I would also like to thank the University of Pennsylvania Critical Writing Program for making me a teaching fellow for the Academic Year 2012-2013. This experience provided invaluable teaching experience while also allowing me to continue work on my dissertation. In particular, I would like to thank Valerie Ross, Patrick Werner, Doug Paletta, and Roger LeGrand. My time as a Critical Writing Teaching Fellow has profoundly impacted my approach to teaching. I will carry the influence of your wonderful program for all my teaching life.

I would also like to thank my committee members Paul Guyer, Karen Detlefsen, and Elisabeth Camp for their assistance, advice, and guidance through the dissertation writing process.

Finally, I would like to thank Sarah Jane Compton for her constant encouragement, steadfast support, and companionship.

ABSTRACT

HUMAN BEINGS AND THE MORAL LAW: MORAL PRECARIOUSNESS IN KANT'S ETHICAL PHILOSOPHY

Bradley M. Taylor

Dr. Paul Guyer

This dissertation is an examination of human moral precariousness in Kant's ethics. Human beings are in a state of moral precariousness insofar as they are ever-capable of transgressing the moral law and are often uncertain of the moral worth of their actions. Put another way, in this dissertation I argue that the basic relationship between human beings and the moral law, in Kant's moral philosophy, is, most fundamentally, one of tenuousness and vacillation. This relation is the fundamental characteristic of the human moral condition because such a relation is built into Kant's account of human moral agency. We have a tenuous relation to the moral law because we always have at least the possibility of conflict between our desire for happiness (i.e. the satisfaction of our inclinations) and the requirements of the moral law. We also may have a vacillating relation to the moral law insofar as we usually find ourselves acting in accordance with the moral law (that is, we often find ourselves committing lawful actions), while also (for many, if not all, human beings) finding ourselves occasionally deviating from the requirements of the moral law (this is a consequence of Kant's doctrine of radical evil). In my dissertation, I argue that this moral precariousness manifests itself as a set of ongoing, perpetual moral crises. That is, there are several crucial points in Kant's moral

thinking that provide the occasion for a crisis. In each of these crucial points, Kant's account of human beings as dual-natured (both natural and rational) generate a struggle (or at least the representation of a struggle) that has a variety of moral consequences. This dissertation approaches this moral precariousness through analyses of several key features of Kant's moral philosophy: his concept of humanity, the dignity of human beings, the moral feeling of respect, and experience of sublimity, and Kant's theory of radical evil. Through an examination of each of these topics, I argue that Kant's account of the moral condition of human beings is one of perpetual open-endedness, uncertainty about one's moral worth, and, above all, (at least) potential crisis.

TABLE OF CONTENTS

ACKNOWLEDGMENT	IV
ABSTRACT	V
PREFACE	VIII
CHAPTER 1	1
CHAPTER 2	46
CHAPTER 3	91
CHAPTER 4	135
CHAPTER 5	166
BIBLIOGRAPHY	234

Preface

Human Beings and the Moral Law: Crises in Kant's Moral Philosophy

One of the more striking aspects of Kant's moral philosophy is the moral precariousness of human beings.¹ The moral law presents itself as an imperative, as an imposition on us that may limit our striving for happiness. Oftentimes a person will encounter moments where they must choose between the morally correct thing to do and their self-interested happiness. While Kant does claim that we can cultivate certain moral feelings that create greater and greater harmony between happiness and morality, there is nevertheless always the possibility of a conflict between what our self-love would have us do and what the moral law requires of us. Kant's moral philosophy is committed to the idea that a perfect harmony between our natural and rational beings is unlikely to ever occur, and even if it does occur, is always subject to deterioration. For most human beings, the clash between self-love and morality is a recurring issue; for the rare few who have disciplined their inclinations to be broadly consistent with (and even conducive to)

¹ Some recent scholarship on Kant's moral thinking has focused on this precariousness, emphasizing the extent to which human beings find themselves on the wrong side of morality. While my approach to Kant's moral thinking is hugely indebted to recent work on Kant's theory of evil, I take as my starting point the moral precariousness that human beings face, even when acting in accordance with morality. That is, human beings neither as virtuous, nor as mired in immoral behavior, but rather as uncertain (and vaguely pessimistic) about their moral worth. For discussions of the pervasiveness of evil in Kant's moral thinking, see "Kantian Moral Pessimism" by Patrick Frierson and "Evil Everywhere: The Ordinariness of Kantian Radical Evil" by Robert B. Loudon both in *Kant's Anatomy of Evil*, edited by Sharon Anderson-Gold and Pablo Muchnik: Cambridge University Press, New York, NY (2010).

morality must stay vigilant in their discipline, lest they experience anew the unhappy clash between the two competing principles.

This dissertation takes this clash (whether it be actual and merely potential) as the starting point for its investigation of several aspects of Kant's moral philosophy. I argue that the clash between self-love and morality is a crucial and recurring issue in Kant's thinking about morality and its relation to human beings. That is, human beings, for Kant, have a fraught relation to the moral law, one that tends towards moral failure or (at least) moral struggle. The inextricable precariousness of human beings' moral situation is an important part of the backdrop for Kant's thinking about morality. Without a full appreciation for the clash between self-love and morality, Kant's moral philosophy may seem overly formalistic and dry. Kant's moral philosophy, though, does indeed take this striking, potentially dramatic struggle between competing principles (self-love and morality) as embedded into the very nature of moral choice itself, and into his account of the moral psychology that flows from moral choice. The clash between self-love and morality is a recurring and inescapable aspect of Kant's moral thinking. This clash manifests in Kant's moral philosophy as various crises. By "crisis" I mean it in its two ordinary senses: 1) a crisis is of a time of conflict and tension between competing interests; crisis is resolved when one of the competing interests takes precedence and affects the chain of events that follows the resolution; 2) a crisis is an experience of heightened tension that accompanies a significant struggle whose outcome is not readily known. A crisis in this sense is a psychological phenomenon as well as an objective one.

In Kant's moral philosophy, we see many instances of crises in both senses of the word. The presence of these crises is an unsurprising consequence of certain key features of Kant's moral thinking. These crises appear largely on account of our (as human beings) status as natural and rational beings. As we will see, these crises appear in Kant's thinking in crucial points of conflict between our dual natures. The appearance of these crises, then, is not accidental, but rather is integral to Kant's overall moral theory of how human beings experience the moral law. In Kant's moral philosophy, there are (on my reading) three kinds of crisis that occur in the moral life (and moral theorizing) of human beings. Each type of crisis is grounded in the ever-present potential (or actual) clash between self-love and morality, but each figures into Kant's moral thinking in a distinct way, and each operates in different ways on how human beings are affected by their knowledge of, and struggle with, the moral law. The crises in Kant's moral philosophy I describe as follows:

1. In the noumenal realm, there is a crisis that we must represent as taking place between the principle of self-love and the moral law. That is, the formation of our fundamental maxim involves a choice, according to which self-love is prioritized (by the will) over morality or vice-versa.² The kind of choice that the will makes determines the

² In his book *Kant on Religion within the Boundaries of Mere Reason*, Lawrence Pasternack refers to the formation of this fundamental maxim the *Gesinnung* choice. The choice of our fundamental maxim, though, must occur in the phenomenal world, since maxims are embedded in our volitional state (which is psychological). The *Gesinnung* choice must be determined by some activity in the noumenal world, which is genuinely imputable to us. Therefore, if we think about the noumenal choice, we can represent it as consisting of a struggle between self-love and morality, but the actual struggle between those two principles occurs in the phenomenal world. We can think of the noumenal choice as adopting either a positive or a negative stance toward morality because the immediate consequence of the noumenal choice will be our *Gesinnung* choice and because the more remote consequences of the noumenal choice are the more specific

morally relevant conduct of the moral agent in the phenomenal world. This crisis I call the “crisis from transcendental idealism” because it follows from Kant’s rigid distinction between the noumenal and the phenomenal realms. It is a crisis we cannot experience directly, but which we must represent as a crisis between two competing volitional principles.

2. From our perspective, in the phenomenal world, we have no access to the workings of the noumenal choice. As a result, we can never experience directly the exercise of our free choice in choosing our fundamental maxim. In our phenomenal choices, we can recall, and even think about in real time, the processes by which we render a decision. This is not the case for noumenal choices; we only can observe the consequences of our choice. As a result of this epistemological inaccessibility and uncertainty, we can never rest assured of our possession of a good will and therefore our moral worth. Even when we are engaged in seemingly appropriate behavior, we might simply be acting in accordance with morality, rather than out of pure respect for the moral law. Furthermore, we know that it is always possible for us to commit an immoral action, even if we have hitherto not done so. This means that we must represent the noumenal choice as perpetually open-ended. Our inability to observe the noumenal choice directly can result (especially perhaps in particularly morally earnest people) in an uneasy sense of tension in which our relation to the moral law is unknowable and the possibility for immoral actions is ever-present. I call this kind of crisis the “crisis from transcendental epistemology.”

maxims we adopt in our day-to-day lives. Pasternack seems to treat the Gesinnung choice and the noumenal choice as one and the same (pp. 115-117).

3. Finally, there are the crises that emerge from empirical introspection. Kant, in the *Groundwork*, states that our (psychological) motives are opaque to us. He writes early on in Section II of the *Groundwork*: “[I]t is absolutely impossible by means of experience to make out with complete certainty a single case in which the maxim of an action otherwise in conformity with duty rested simply on moral grounds and on the representation of one’s duty.”³ We cannot trust introspection to provide us with insight into our motives in part because we tend to overestimate the moral worth of ourselves: “...for we like to flatter ourselves by falsely attributing to ourselves a nobler motive, whereas in fact we can never, even by the most strenuous self-examination, get entirely behind our covert incentives, since, when moral worth is at issue, what counts is not actions, which one sees, but those inner principles of actions that one does not see.”⁴ In this passage, Kant seems to be making a slightly ambiguous statement when he discusses the “inner principles of actions that one does not see.” Does he mean that we do not have access to our the full depth of our psychological state when we act, or does he mean by that phrase the noumenal choice, which, of course, we do not see? While this particular phrase might be ambiguous, I would argue that if it is, then that is only because Kant affirms both. The inaccessibility of the noumenal gives rise to the second sort of crisis I discuss immediately above and the opacity of our psychological motives is this third sense of crisis. This is true when thinking about morally neutral actions as well as morally relevant ones. This opacity of motives means that there is no way of knowing whether or not an action was motivated by respect for the moral law, or by some self-

³ G 4:407.

⁴ G 4:407.

interested inclination. Moral earnestness, according to Kant, brings about uncertainty and doubt about one's moral worth, not self-satisfaction. Such doubt in human beings, who all recognize the categorical and imperative nature of the moral law, may bring about a psychological crisis.

This dissertation will examine five aspects of Kant's moral philosophy that all, in one way or another, manifest at least one of these three sorts of crises. I argue that much of Kant's moral philosophy has some sense of moral crisis underlying, framing, or informing many of its core concepts and ideas. I hope to show this by devoting each chapter to a central concept in Kant's moral thinking and to show how each one, in its own way, brings to light the potential for moral crisis in us human beings that dramatizes or otherwise manifests the precarious relationship between human beings and the moral law.

Chapter one, titled "Humanity in Kant's Moral Philosophy," discusses Kant's conception of humanity. Here, I argue that members of humanity, according to Kant, always exercise their capacity to set and pursue ends freely (i.e. exercise their humanity) according either to the principle of self-love or to the principle of morality. That is, the supreme maxim that a person forms in setting his or her ends either has the moral law or happiness as its foundation. If a human being, in exercising their capacity for free choice, pursues ends on the basis of respect for the moral law, then the expression of the humanity is also an expression of the moral personality. Alternatively, if a human being makes an immoral choice and opts to satisfy an inclination in opposition to his or her duty, then the expression of humanity is not also an expression of one's moral

personality. Humanity, then, occupies a vacillating place on the continuum between what Kant calls animality (our predisposition to pursue happiness) and personality.

Chapter two, titled “Human Dignity in Kant’s Moral Philosophy,” deals with the notion of human dignity in Kant’s moral philosophy. I argue that the basis for human dignity, according to Kant, rests upon our capacity for morality. Human dignity lies in our capacity to elevate ourselves above the dictates of nature (as expressed through our natural inclinations) and to organize our choices on the basis of morality. Human beings, therefore, are in a crisis position in which we either live up to our dignity, or fail to do so. Human dignity, then, is a dynamic concept that involves capacity to lift ourselves up out of the merely natural. The vacillating dynamism of human dignity is an integral part of the human moral condition because, while we can fail in achieving the fullness of our dignity, we can never lose entirely our predisposition to the good. There is always the possibility of moral conversion, so that no matter the kind of life a person has led up to a certain point, it is always possible to change course. Each moment can be a crucial point in the moral life of a human being. This possibility for moral conversion is the baseline, necessary and sufficient condition for human dignity.

Chapter three, titled “Respect in Kant’s Moral Philosophy,” analyzes Kant’s account of respect as a moral category. Here, I argue that Kant employs respect as a moral term in at least three distinct ways: 1) as the feeling human beings have for the moral law; 2) as the attitude we are required to harbor for other human beings; 3) the feeling of approbation we have towards human beings who demonstrate virtue. I claim that the second and third senses of respect are ultimately grounded in the first sense. That

is, the respect we owe all human beings and the praiseful respect we pay to moral exemplars are dependent on, and derive from, the respect we feel for the moral law itself. In this chapter, I also argue that each sense of respect, in different ways, arises out of (and helps clarify) our own crisis-ridden relation to the moral law. That is, it is only in beings like us (that is, beings who are prone to moral crisis) that the feeling of respect can take on the character that it does.

Chapter four, titled “Sublimity and Morality,” is about the sublime and its relation to Kant’s moral philosophy. Here I claim that Kant argues that the sublime can have an indirect contribution to a moral life, and that Kant’s account of the sublime manifests and dramatizes the potential crisis between our natural and rational natures. The sublime experience (particularly the dynamical sublime) makes us feel (or at least imagine) the rupture point between our animal and rational natures and enlivens in us an awareness of our potential for moral action. The sublime does not resolve the crisis between self-love and morality, but rather brings about a heightened awareness of it, which we experience as pleasurable because of the sense it allows us to feel of being elevated above nature. Sublimity, I argue, gives Kant a way of doing justice to those intense moments of moral self-awareness without giving way to enthusiasm and fanaticism.

Chapter five turns to the question of radical evil in Kant’s moral philosophy. Here, I argue that, given Kant’s doctrine of radical evil, human beings tend to find themselves in a perpetual moral crisis because we are not morally good, but, given our predisposition to personality, strive (or at least wish we were striving) for moral goodness. Kant’s account of radical evil is not a despairing doctrine, but rather one that

describes human beings as drawn in two directions by competing principles. There remains, even for immoral human beings, the possibility of moral conversion and the achievement of a good will. Correspondingly, no good human being is immune from the possibility of regress back into immorality. Human beings, on Kant's account, always find themselves in a potential crisis-point. Moral life is a perpetually open-ended endeavor with no assurances or guarantees of our moral goodness or of our moral impropriety.

Kant's moral philosophy, I argue, stakes out a middle ground between ungrounded moral optimism and self-defeating moral despair. Ungrounded moral optimism can instill a sense of laziness, complacency, and arrogance that inevitably wear away at our moral fortitude. In extreme cases, such simplistic optimism can give rise to moral fanaticism where one falsely assumes that one's own desires and morality form a one-to-one relationship. Desire and morality, in such a case, combine and, as a result, the potential for a moral crisis is nullified.⁵ So, human beings, when using their reason properly, relate to the moral law through its appearance to us as an imperative, as if from without. We do not identify with the moral law except insofar as we see ourselves as potentially acting out of respect for it.

The opposite extreme regarding the moral law and its relation to human beings that one might take is one of moral despair. If the moral law can never be totally absorbed

⁵ In her paper, "Kant's Account of Practical Fanaticism," Rachel Zuckert makes a similar point when she writes, "Practical reason provides us with legitimate concepts of the supersensible—of God and of the morally good—and these ideas are also legitimate objects of feeling (moral feeling). Nonetheless, Kant argues, such feeling does not identify the good or ground moral claims...Rather, moral feeling is an effect of one's recognition of the validity of the moral law" (308).

by us so as to prevent the possibility of a clash between happiness and morality, and the majority (if not the totality) of human beings are evil, then it seems that people may lapse into a kind of despair in which one loses hope that one can become good. Kant's moral philosophy, though, does provide space for moral hope. In the *Collins Lectures on Ethics*, Kant warns against fixating too strongly on our own (or humanity's in general) moral shortcomings: "However, let us not brood upon the weakness of human nature, and wonder whether it be incapable of moral purity; for the endeavour to find all one's actions impure is apt to cause a man to lose self-confidence in his ability to perform good and morally pure actions, and to believe his nature too weak and incapable of that; we have to believe, rather, that *rectitude moralis* might be a strong motivating ground for us."⁶ Human beings are always capable of moral conversion to the good, and are always subject to experiences (sublimity) or feelings (moral respect and benevolence) that can embolden them in their hope of attaining a good will and realizing the fullness of their moral dignity.

⁶ Collins Lectures 27:293.

CHAPTER 1

Humanity in Kant's Moral Philosophy

Introduction

This chapter examines the concept of humanity in Kant's moral philosophy. I argue that there are two distinct tendencies in Kant's discussions of humanity. One is Kant's tendency to treat humanity as consisting in possession of a rational nature. The other is Kant's regular discussion of humanity being in some relation to the moral law as well as being in possession of a moral predisposition. I also contend these two ways do not seem to be straightforwardly consistent with each other. In this chapter I examine both of these tendencies in Kant in some detail, independent of each other. Once I have examined these dual notions of humanity, I will discuss two possible ways in which, their prima facie distinctness notwithstanding, these two tendencies in Kant's discussions of humanity can be rendered consistent with each other. The first I label the 'moral capacity reading of Kantian humanity,' the second the 'good will reading.' After introducing these two ways of reconciling the two tendencies, I will discuss some problems facing each interpretation of Kantian humanity. I argue that the good will reading cannot be maintained, and that the moral capacity reading can be maintained, but only after taking into account, and resolving, the problems I raised. I conclude by defending a version of the moral capacity reading of Kantian humanity. I will argue that, properly understood, the moral capacity reading reconciles the two tendencies in Kant's moral philosophy and also that these tendencies are actually deeply connected to each other when Kant considers the humanity of human beings.

I. The Importance of Humanity: Humanity as an End-In-Itself

Determining the precise nature of humanity has wide-ranging consequences for any interpretation of Kant's moral philosophy. The second formulation of the categorical imperative reads: "*So 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.*"⁷

Earlier in the *Groundwork*, Kant seems to suggest that human beings are ends in themselves as a result of their being rational beings: "Now I say that the human being and in general every rational being *exists* as an end in itself, *not merely as a means* to be used by this or that will at its discretion; instead he must in all his actions, whether directed to himself or also to other rational beings, always be regarded *at the same time as an end.*"⁸

Humanity, according to Kant, has a value above all price and cannot be traded away at equal value for anything else. Kant writes, "What is related to general human inclinations and needs has a *market price*...but that which constitutes the condition under which alone something can be an end in itself has not merely a relative worth, that is, a price, but an inner worth, that is, *dignity.*"⁹ Humanity is a dignity that is to be treated as sacrosanct, never to be undermined or violated for another purpose, no matter how valuable that purpose may be.¹⁰

As a result of its status as an end in itself, humanity is also granted more particular moral protections within Kant's moral philosophy. Members of humanity are

⁷ G 4:429.

⁸ G 4:428.

⁹ G 4:435.

¹⁰ In his notes on moral philosophy, Kant describes humanity as "holy and invulnerable" (Notes and Fragments 19:165).

not to be treated with contempt or disrespect, but rather with respect: “Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other. Humanity itself is a dignity; for every human being cannot be used merely as a means by any human being (either by others or even by himself) but must always be used at the same time as an end.”¹¹ Kant specifically lays out “duties of respect” that are to be extended to all of humanity; we must not, according to our duties of respect, defame, ridicule, or act arrogantly toward members of humanity.¹²

Given the importance of humanity in Kant’s moral philosophy, it becomes especially important to have a clear understanding of what exactly Kant means by humanity. Kant, it seems, does not have a simple view of what humanity is, but there do seem to be two tendencies in Kant’s various descriptions of humanity that form the basis of Kantian humanity: 1) humanity as possessing a rational nature, capable of setting ends, and 2) humanity as being receptive to morality and capable of moral action.

II. Rational Nature and Setting Ends

Kant at times discusses humanity as consisting in possession of a rational nature.¹³ The basic distinction that Kant makes in discussing humanity as rational nature is that between humanity and animality. Whereas humanity involves possession of rational capacities, animality involves those features that belong to the human being considered solely as an animal. The basic principle that underlies animality is what Kant calls in the *Religion* “physical or merely *mechanical* self-love, i.e. a love for which

¹¹ MM 6:462.

¹² MM 6:465-469.

¹³ G 4:428-430.

reason is not required.”¹⁴ In the case of human beings, our animality is concerned chiefly with our instinctual drives toward self-preservation, propagation of the human species, and inclusion within a community. In the *Religion* Kant writes, “It [the predisposition to animality in the human being is threefold: *first*, for self-preservation; *second*, for the propagation of the species, through the sexual drive, and for the preservation of the offspring thereby begotten through breeding; *third*, for community with other human beings, i.e. the social drive.”¹⁵ Animality, and the various drives that it involves, are not the result of rational deliberation and evaluation; that is, they are instinctual.¹⁶

Humanity, in contrast to animality, involves rationality. Kant holds that humanity consists in possession of a rational nature; he contends that the crucial feature of humanity, or possession of a rational nature, that which distinguishes humanity from the rest of nature (especially animality), is the capacity to set ends for oneself. The paradigmatic expression of one’s rational nature is when agents’ (members of humanity) set ends for themselves. Kant asserts in the *Groundwork* that the capacity to set ends is the crucial distinguishing characteristic between rational nature and animal nature. Kant writes, “Rational nature is distinguished from the rest of nature by this, that it sets itself an end.”¹⁷ The ends that the agent pursues are not passively received through the

¹⁴ Rel. 6:26.

¹⁵ Rel. 6:27.

¹⁶ Allen Wood describes the predisposition to animality in similar terms: “They [the drives of animality] give us (as they do nonrational) an instinctive urge to do what is necessary for our own survival and that of the species. But they operate at an entirely prerational level in that they involve no conscious representation of these natural ends and no deliberation as to how they are to be achieved” (*Kant’s Ethical Thought* 118).

¹⁷ G 4:437.

impulses and inclinations nature implants in the agent, but rather the agent in possession of rational nature “sets *itself* an end.”¹⁸

In the *Metaphysics of Morals*, Kant reiterates the connection between humanity and the capacity to set ends when he writes, “The capacity to set oneself an end—any end whatsoever—is what characterizes humanity (as distinguished from animality).”¹⁹ As he did in the *Groundwork*, here Kant suggests that humanity’s essential characteristic is its capacity to set ends.²⁰ He emphasizes setting ends as humanity’s central characteristic earlier in the *Metaphysics of Morals* in writing that humanity is that “by which he alone is capable of setting ends.”²¹

The activity of humanity, unlike animality, is not determined solely by instinct. Animals may appear to act in deliberate and careful manners; such mannerisms may even resemble the ways in which human beings (humanity) behave. The crucial difference, Kant argues, between animality and humanity is that those who possess humanity set ends that need not be the result of a dictation from instinct.²² Animals uncritically pursue ends that are set by instinct. Humanity, by contrast, is capable of deliberation and

¹⁸ Ibid (emphasis added).

¹⁹ MM 6:392.

²⁰ Christine Korsgaard writes, “In the *Groundwork*, Kant interchanges the terms “humanity” and “rational nature.”” (*Creating the Kingdom of Ends* 110). She also writes, “Kant takes the characteristic feature of humanity, or rational nature, to be the capacity for setting an end (110). Allen Wood, in *Kantian Ethics*, equates humanity with “rational nature” (85). He also writes, “Humanity contains our rational capacity to set ends and devise means to them, and our rational self-love, giving us grounds for forming a conception of our happiness and pursuing it.” Wood also describes humanity as “the capacity to set ends according to reason” (94).

²¹ MM 6:387.

²² Paul Guyer emphasizes this feature of Kantian humanity when he writes that Kant holds, “The only natural capacity unique to humankind is its potential to free itself from mere nature by exercise of its freedom” (*Kant on Freedom, Law, and Happiness* 169).

reflection before choosing whether or not to pursue some end that instinct recommends. In the *Critique of the Power of Judgment*, Kant holds that human beings, in addition to acting as animals would, act “purposively and independently of nature.”²³ Kant also elaborates on humanity’s unique capacity to set ends when he describes the human being (in general) as “the sole being on earth who has reason, and thus a capacity to set voluntary ends for himself...”²⁴ We set our ends voluntarily, and thus independently of nature, insofar as human beings are not necessitated to pursue ends recommended by instinct.

Instinct does not provide the basis according to which members of humanity set and pursue ends. Rather the basis is provided through an evaluative and deliberative process whereby one reflects on ends (deliberative) and renders an evaluation about which ends are worth pursuing. Christine Koorsgard emphasizes the deliberative and evaluative components of humanity’s capacity to set ends when she writes, “When Kant says that the characteristic of humanity is the power to set an end...he is referring to a more general capacity for choosing, desiring, or valuing ends; ends different from the ones that instinct lays down for us, and to which our interest is directed by the operations of reason.”²⁵

²³ CPJ 5:434.

²⁴ CPJ 5:431.

²⁵ *Creating the Kingdom of Ends* 114. She goes on to write that “the distinctive feature of humanity, as such, is simply the capacity to take a rational interest in something: to decide, under the influence of reason, that something is desirable, that it is worthy of pursuit or realization, that it is to be deemed important or valuable, not because it contributes to survival or instinctual satisfaction, but as an end—for its own sake” (114).

The act of freely, or voluntarily, setting and pursuing an end involves rendering an evaluative stance toward the end, namely, the agent must deem the end as worth pursuing. When instinct does not simply dictate what one does, members of humanity must take it upon themselves to assign value to some ends. Kant imagines what might be involved with humanity's coming to reject instinct as the sole arbiter over what they will do when he writes, "He discovered in himself an ability to choose his own way of life without being tied to any single one like the other animals. But the momentary gratification which this realization of his superiority may have afforded him was inevitably followed at once by anxiety and fear as to how he should employ his newly discovered ability...He stood, as it were, on the edge of an abyss. For whereas instinct had hitherto directed him towards individual objects of his desire, an infinite range of objects now opened up, and he did not yet know how to choose between them. Yet now that he had tasted this state of freedom, it was impossible for him to return to a state of servitude under the rule of instinct."²⁶ Here again we see Kant emphasize humanity's freedom from instinct as being the central feature of humanity as well as the realization that freedom involves one's own choice in deciding about which objects of desire (or ends) one will pursue. Involved in such a choice-making process must be some way of determining which objects of desire are *worth pursuing*.

It is important to note that the capacity to set ends does not necessarily include the capacity to set *moral* ends. That is, one can have the capacity to set ends without having the separate, but related, capacity to set ends for the sake of morality. Humanity, insofar

²⁶ Conjectures on the Beginning of Human History 224.

as the concept is exhausted by the capacity to set ends, does not necessarily involve morality as entering into the process whereby members of humanity set ends.²⁷ Setting an end involves the use of one's rational capacities, but one can exercise a rational capacity without reference to morality. In the *Religion* Kant seems to make this point when writes, "The most rational being of this world might still need certain incentives, coming to him from the objects of inclination, to determine his power of choice. He might apply the most rational reflection to these objects—about what concerns their greatest sum as well as the means for attaining the goal determined through them—without thereby even suspecting the possibility of such a thing as the absolutely imperative moral law which announces to be itself an incentive, and indeed, the highest incentive."²⁸

Kant, then, seems to be suggesting that humanity is amoral in the sense that an individual may possess humanity without also possessing a receptivity to the moral law. Kant argues that we can readily imagine a perfectly rational being that sets and pursues ends who, at the same time, does not consult with the moral law at all in his or her deliberation. The agent in this case exhibits humanity (and therefore agency), but not *moral* agency.

It may seem natural, given Kant's apparent amoral description of humanity, to think of Kant as committed to a view a humanity that does not include moral agency. Humanity on this reading need not involve any reference to morality in the determination

²⁷ Koorsgaard emphasizes the distinction between setting ends and setting moral ends when she writes, "[i]t is the capacity for the rational determination of ends in general, not just the capacity for adopting morally obligatory ends, that the Formula of Humanity orders us to cherish unconditionally" (*Creating the Kingdom of Ends* 111).

²⁸ Rel. 6:26n.

of one's ends. An individual can be a member of humanity without having any relation whatsoever to the moral law. This interpretation of humanity can be labeled as the amoral reading of Kantian humanity. In short, according to the amoral reading, Kantian humanity consists solely in one's rational capacity to deliberate about, to evaluate, and to set voluntary ends. Moral considerations need not enter into one's deliberative procedure or factor into the content of one's chosen ends. One can exhibit humanity while also exhibiting complete indifference to the moral law.²⁹

II. Moral Predisposition

While the amoral tendency of Kantian humanity has considerable textual support, there is also another tendency in Kant's discussions of humanity that suggests Kantian humanity involves moral consideration as well. This morally thick view of humanity poses a potential problem for an interpretation of humanity that focuses solely on rational deliberation in setting voluntary ends. According to this moralistic interpretation of humanity, the amoral picture, while not incorrect as far as it goes, is incomplete.

The general thrust of the morally thick view of Kantian humanity consists in Kant's suggestion that human beings have an "original predisposition to the good."³⁰ In the *Religion* Kant holds that predispositions or properties are in human beings originally

²⁹ This, of course, is not the case in ordinary human beings. This sort of humanity is not commonly (perhaps never) realized in human beings. The amoral reading of Kantian humanity is simply an interpretation of the concept of humanity in Kant's moral philosophy and not the complex reality of humanity in the case of human beings, who also have a relation to the moral law, which figures into the decision-making process of human beings.

³⁰ MM 6:441.

when “they belong to the possibility of human nature.”³¹ Part of what it is to be a human being is to have the predisposition to the good. That is, all human beings considered simply as human beings must have tendencies that make them receptive to the moral law, its imperatives, and morally relevant affective states. In the *Metaphysics of Morals*, for instance, Kant writes that a human being “can never lose entirely his predisposition to the good.”³²

Furthermore, Kant sometimes suggests that if a human being were to lack, or lose, an original predisposition to the good, and its attendant morally relevant affective states, then that person would also lack humanity. This morally thick view of humanity often arises when Kant is discussing human beings’ receptivity to moral feeling. For instance in the *Metaphysics of Morals*, Kant writes, “No human being is entirely without moral feeling, for were he completely lacking in receptivity to it he would be morally dead; and if (to speak in medical terms) the moral vital force could no longer excite this feeling, then humanity would dissolve (by chemical laws, as it were) into mere animality and be mixed irretrievably with the mass of other natural beings.”³³ Without receptivity to moral feeling, our humanity would “dissolve” into “mere animality.” If this were to occur then the human being would simply be an inhabitant of the natural world, “mixed irretrievably with the mass of other natural being.” Kant does not suggest that our moral personality

³¹ Rel. 6:28.

³² MM 6:464. Kant also reiterates his view that there is a universally shared predisposition to the good in human beings in *Toward Perpetual Peace* when argues that since states at least pretend to care about the concept of right, it proves that “there is to be found in the human being a still greater, though at present dormant, moral predisposition to eventually become master of the evil principle...” (8:355).

³³ MM 6:400.

would dissipate if we lacked moral feeling, rather he suggests that our *humanity* would wither away and we would be back into a state of mere animality. If the amoral reading maintains that an agent could be indifferent to morality while still retaining their humanity, then this passage seems flatly to contradict the amoral interpretation. According to this passage, a characteristic of human beings that is necessary for us to retain our humanity is a decidedly moral one, not one that involves mere deliberative rationality in setting ends. If Kant is committed to the claim that a predisposition to the good is an essential component of humanity, then a straightforward amoral reading of Kantian humanity cannot be sustained.

There is further textual support for a more morally thick reading of Kantian humanity. In the *Critique of Practical Reason* Kant suggests that reason alone (that is deliberative, amoral reason) does not raise human beings above animality. He argues, on the one hand, it is undeniable that, insofar as human beings are members of the natural world, it is inevitable (and not necessarily problematic) that they will use reason to aid in the satisfaction of their animal desires: “The human being is a being with needs, insofar as he belongs to the sensible world, and to this extent his reason certainly has a commission from the side of his sensibility which it cannot refuse, to attend to its interest and to form practical maxims with a view to happiness in this life and, where possible, in a future life as well.”³⁴ On the other hand, Kant argues that amoral use of reason, on its own fails to raise the human being above mere animality: “But he is nevertheless not so completely an animal as to be indifferent to all that reason says on its own and to use

³⁴ CPrR 5:61.

reason merely as a tool for the satisfaction of his needs as a sensible being. For, that he has reason does not at all raise him in worth above mere animality if reason is to serve him only for the sake of what instinct accomplishes for animals; reason would in that case be only a particular mode nature had used to equip the human being for the same end to which it has destined animals, without destining him to a higher end.”³⁵ Reason used only in the pursuit of the needs of one’s sensible being would simply be a mechanism nature has equipped human beings with. Reason, then, would not be all that different from the fast-twitch musculature of a rabbit, the long nose of an anteater, or the visual system of an eagle. This would not give to human beings a worth above and beyond that of the rest of nature. But Kant is also clear that possession of humanity gives an agent a worth beyond accounting.³⁶ If humanity is that by which one attains absolute worth, apart from mere animality, and reason simply in the service of sensibility does not separate an agent from mere animality, then humanity must be something beyond the mere exercise of reason in the setting and choosing of ends.

III. Reconciling the Tendencies: Two Options

It seems clear, then, there exists two tendencies in Kant’s moral philosophy regarding humanity. On the one hand Kant at times describes humanity as consisting solely in a capacity to set voluntary ends, independent of instinct. On the other hand, Kant occasionally suggests that humanity involves some relation to morality. If we accept that both tendencies factor into Kantian humanity, then we need an account that captures

³⁵ CPrR 5:61-62.

³⁶ G 4:434-435.

both the capacity to set ends and the capacity for morality. The task, then, for a satisfactory interpretation of Kantian humanity would incorporate the capacity to set ends while also taking into account Kant's repeated claims that humanity must also involve the agent having some relation to morality.

In this section I will describe two ways in which some commentators have brought together these tendencies in Kant's treatment of humanity. The first approach is what I will call the 'moral capacity reading' of Kantian humanity; the second approach is the 'good will reading' of humanity put forward by Richard Dean. After briefly laying out the basics of both readings, I will consider some difficulties I take to be facing each approach.

The Moral Capacity Reading

There is a significant strand of Kant scholarship that interprets humanity along more moralistic lines than the amoral reading. This interpretation treats humanity as a rational nature, which, in addition to having the power to choose and set ends, includes the *capacity for morality*. We see this tendency in Thomas Hill,³⁷ Barbara Herman,³⁸ Roger Sullivan,³⁹ and also in John Rawls.⁴⁰ Rawls expresses the basic description of humanity on the moral capacity reading in his *Lectures on the History of Moral*

³⁷ Thomas E. Hill, *Respect, Pluralism, and Justice* (page 87). Hill's formulation here seems to be less purely in the humanity as capacity for morality camp, but he does describe humanity as possessing "certain normative capacities and dispositions" which include, among other things, the willingness "to reciprocate with others in endorsing principles that respect each person as a potential source of legitimate values" (87).

³⁸ Barbara Herman, *The Practice of Moral Judgment* (page 238).

³⁹ Roger Sullivan, *Introduction to Kant's Ethics* (page 70).

⁴⁰ John Rawls, *Lectures on the History of Moral Philosophy* (page 188).

Philosophy, “Kant means by humanity those of our powers and capacities that characterize us as reasonable and rational persons who belong to the natural world... These powers include, first, those of moral personality, which make it possible for us to have a good will and a good moral character; and second, those capacities and skills to be developed by culture: by the arts and sciences and so forth.”⁴¹ The existence of some sort of relation between the agent and the moral law is not irrelevant, or beside the point, to this interpretation of humanity. Rather humanity consists precisely in having capacities suited for morality and in being capable of possessing a good will. These capacities must include the ability to set ends voluntarily and apart from what instinct recommends. Also included in the capacity for morality must be receptivity to the moral law. A person must be aware of the moral law and that it applies to him or herself in order for moral action to be possible.

The Good Will Reading

In his book *The Value of Humanity in Kant's Moral Theory* Richard Dean presents a provocative challenge to the capacity interpretation of humanity. Instead of conceiving humanity as the rational capacity to set ends as well as the capacity for moral personality, Dean argues that humanity consists in the actual possession of a good will. The various proponents of the moral capacity reading, regardless of whatever disagreements they may have, are united in the claim that possession of a good will is not a necessary feature of humanity. That is, one could be a member of humanity while also

⁴¹ Rawls 188.

lacking a good will. Richard Dean's basic departure, then, from the capacity interpretation is his rejection of that claim.

Dean argues that the term 'humanity' in Kant's moral philosophy is reserved for a rational being "who actually accepts moral principles as providing sufficient reasons for action."⁴² He then labels such a being as possessing a "properly ordered will," which is a will in which moral considerations are given priority over self-love. Dean then claims that this sort of will is a good will.⁴³ A good will is found in any rational being "who actually accepts moral principles as providing sufficient reasons for action."⁴⁴ So, on Dean's reading, the humanity formulation can be interpreted as claiming that those who possess the property of having a good will are to be treated as ends-in-themselves. This implies that anyone or anything that lacks a good will is not required to be treated as ends-in-themselves.

IV. Problems with the Moral capacity reading and the Good Will Reading

I would argue that both the capacity and good will interpretations face problems in explaining Kant's notion of humanity. In this section I explore some potential problems with each approach. There are also problems facing the good will interpretation of Kantian humanity. While the good will reading is worth consideration, given the consequences it would have for Kant's moral philosophy, it leads to conclusions that cannot be withstood within Kant's moral philosophy. In what follows I first run through three problematic implications of the good will reading. After my discussion of the good

⁴² Dean 6.

⁴³ *Ibid.*

⁴⁴ *Ibid.*

will reading I attempt to bring out features of humanity that a capacity interpretation would allow but that bear emphasis.

The Moral Capacity Reading

The chief difficulty facing the moral capacity reading is the fact that Kant seems to locate the capacity for morality not in humanity, but in what he calls our “predisposition to personality.” In the *Religion*, for instance, Kant writes, “The predisposition to personality is the susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice.*”⁴⁵ Without such susceptibility, it is not clear how an agent could have the capacity for morality. In the *Groundwork*, Kant argues that respect for the moral law is the only determining ground for a morally good will: “Now, an action from duty is to set aside entirely the influence of inclination and with it every object of the will; hence there is left nothing that could determine it except objectively the law and subjectively *pure respect* for this practical law, and so the maxim of complying with such a law even if it infringes upon all my inclinations.”⁴⁶ If the susceptibility to respect for the moral law is part of our predisposition to personality, and not to humanity, then it seems that the capacity for morality must not be included in an interpretation of Kantian humanity.

Kant in fact explicitly states that the concept of personality must be separated by the concept of humanity: “We cannot consider this predisposition as already included in the concept of the preceding one [humanity], but must necessarily treat it as a special

⁴⁵ Rel. 6:28.

⁴⁶ G 4:400-1.

predisposition. For from the fact that a being has reason it does not at all follow that, simply by virtue of representing its maxims as suited to universal legislation, this reason contains a faculty of determining the power of choice unconditionally, and hence to be “practical” on its own; at least, not so far as we can see.”⁴⁷ One can readily imagine a being with reason, but at the same time lacking receptivity to the moral law. Such an agent would be able to decide on an end to pursue, calculate likely outcomes from his or her actions, determine the best course of action, and then skillfully go about pursuing the end. None of this deliberative process needs to consult with the moral law or ask whether or not the maxim upon which the agent is acting is suited for universal legislation. An additional property, beyond humanity, is required to bring moral considerations to the forefront in an agent’s deliberative process. The moral capacity interpretation of humanity thus seems to miss the mark in identifying the unique characteristic that distinctively marks humanity.

Problems with the Good Will Reading

The first problematic implication of the good will reading is that humanity would be incredibly scarce and therefore that the group whose members are to be treated as ends in themselves would be almost vanishingly tiny. Kant in the *Groundwork* writes, “One need not be an enemy of virtue but only a cool observer, who does not take the liveliest wish for the good straight-away as its reality, to become doubtful at certain moments (especially with increasing years, when experience has made one’s judgment partly more shrewd and partly more acute in observation) whether any true virtue is to be found in the

⁴⁷ Rel. 6:27.

world.”⁴⁸ If true virtue is taken to be possession of a good will, then this passage, coupled with Dean’s reading, would seem to suggest that it’s entirely possible that no one in the world is to be treated as an end in itself. We may hear all sorts of pious-sounding *wishes* that virtue be found in the world and in oneself, but we still would not be able to find *true* virtue anywhere.

Dean tries to preempt concerns about the first implication I discussed by arguing that good will is more widely dispersed than is commonly believed. Dean writes, “[A] good will, properly understood, is not such a rarity among humans...”⁴⁹ Dean makes this claim and then provides some textual support for it. Dean initially defined a good will as being possessed by a person “who actually accepts moral principles as providing sufficient reasons for action.”⁵⁰ A bit later in the book we get a fuller discussion of the good will. He defines the good will as follows, “A good will is the will of an agent who is committed to moral principles, and this commitment can be present even when one is not performing actions that display it.”⁵¹ So he has in mind here a person who is choosing between different morally permissible ends. Such a person, Dean plausibly contends, can still be in possession of a good will. Dean further specifies his definition of a good will when he writes, “A good will is the will of a being who is committed to acting morally, who gives priority to moral principles rather than acting simply to satisfy her own desires, inclinations, impulses, or sentiments.”⁵²

⁴⁸ G 4:407.

⁴⁹ Dean 7.

⁵⁰ Dean 6.

⁵¹ Dean 21.

⁵² Dean 24.

Interpreting the good will as giving priority to moral principles over the principle of self-love is indeed consistent with Kant's description of the good will in the opening section of the *Groundwork*. Giving priority to moral principles over selfish ones is another way of saying that one is prepared to do one's duty, even if doing so would require keeping one's selfish inclinations in check. Kant writes that the concept of "duty...contains that of a good will though under certain subjective limitations and hindrances."⁵³ The "subjective limitations and hindrances" can be interpreted as the inclinations we have that cause us to experience the requirements of the moral law as constricting and as setting limits to what we can do. In fact, Kant says exactly this in the *Mrongovius Lectures* when discussing why people so often fail to do what they know a reasonable person would do: "We are well aware of what a being with such reason would do; but we do it not, for we have inclinations that are hindrances."⁵⁴ Dean's interpretation of a good will, so far, is also consistent with the argument laid out by Kant in the *Religion* where the moral predicament of human beings is construed as a struggle to properly order the requirements of the moral law on the one hand and the inclinations to self-love on the other hand.⁵⁵

In my view, though, Dean problematically argues that possession of a good will is "compatible with significant degrees of self-deception, lack of attention to the moral dimensions of one's choices, and weakness of will."⁵⁶ Dean's interpretation of a good will is perhaps most easily understood in the context of the degrees or "grades" of evil in

⁵³ G 4:397.

⁵⁴ Mrongovius Lectures 29:605.

⁵⁵ See e.g. Rel. 6:36.

⁵⁶ Dean 7.

the human being that Kant discusses in the *Religion*.⁵⁷ Kant writes, “We can think of three different grades of this natural propensity to evil. *First*, it is the general weakness of the human heart in complying with the adopted maxims, or the *frailty* of human nature; *second*, the propensity to adulterate moral incentives with immoral ones... i.e. *impurity*; *third*, the propensity to adopt evil maxims, i.e. the *depravity* of human nature, or of the human heart.”⁵⁸ Dean argues, I think, that a good will is consistent with frailty and impurity, but not with depravity. Dean writes, “Frailty can keep one from choosing particular right actions even though one truly wills to act on the moral law as one’s supreme and overriding principle.”⁵⁹

I would argue that possession of a good will is not compatible with any of these grades of evil. Human beings possess good wills when they perform their duties out of respect for the moral law. They also can possess good wills when, while choosing from different morally permissible ends, the demands of the moral law are in fact given priority over self-love. In cases of frailty, a person has failed to perform his or her duty because they did not in fact properly order the relationship between self-love and morality. The maxim on which they act placed, in that instance, the inclinations of self-love over the requirements of the moral law. Even though, on some level, the person does indeed recognize that morality must be given priority over self-love, the maxim they acted upon had to privilege self-love over morality. The person might be making a pious *wish* that their will was properly ordered, but as Kant notes in the *Groundwork*, even the

⁵⁷ Rel. 6:29-30.

⁵⁸ Rel. 6:29.

⁵⁹ Dean 21.

liveliest of wishes for the good does not amount to its reality. I think this is the best way to interpret Kant when he writes, “An evil heart can coexist with a will which in the abstract is good.”⁶⁰ Dean takes this passage to support his view, but it is not clear how precisely it does. In *abstract* the will may be good, but this, it seems, does not amount to saying that the will is good *in fact*. The maxim on which the person acted cannot possibly, in my view, have properly ordered morality and self-love, otherwise the person would not have intentionally committed the misdeed.

Dean also suggests that there is textual support of his reading of the good will in the *Religion* where Kant is discussing moral progress. Dean writes, “To embark firmly on the path of continual moral improvement is to have the commitment to moral principles that marks a good will.”⁶¹ Kant writes, “If by a single and unalterable decision a human being reverses the supreme ground of his maxims by which he was an evil human being...he is to this extent...a subject receptive to the good; but he is a good human being only in incessant laboring and becoming...to find himself upon the good (though narrow) path of constant *progress* from bad to better...this is the same as actually being a good human being...”⁶² I don’t think this passage provides a sound basis for asserting, as Dean does, that a good will is consistent with occasional intentional immoral actions. Dean writes, “A good will is not present only when an agent acts on the motive of duty. It endures when performing merely permissible actions and can even coexist with some

⁶⁰ Rel. 6:37.

⁶¹ Dean 22.

⁶² Rel. 6:48.

intentional immoral actions.”⁶³ It seems right to say that *person* can be on the whole good even if there are occasional misdeeds along the way, so long as they are engaged in the struggle for moral improvement from bad to better that Kant discusses in the *Religion*, but a good *will*, does not seem to be consistent with intentional immoral deeds.

A background assumption in Dean’s discussion of the good will seems to be the claim that good will is not identical to perfect will. Dean describes a perfect will as follows: “We could imagine a will that is short of a holy will, because it is subject to the temptations of inclination, yet still chooses always to act according to the principles of reason. Kant thinks (and intuition seems to fall on his side here) that humans do not have this sort of perfect, finite will either.”⁶⁴ Since Dean thinks it is reasonable to suppose that good wills are fairly common, it seems to follow that Dean would not think of good and perfect wills as equivalent.

It is not clear, though, that there really is a distinction to be drawn between a perfect will, as defined above, and a good will. A good will, as Kant lays out in the *Groundwork*, is a will that is determined by respect for the moral law and that chooses to act in ways that fulfill one’s duty (even if there are inclinations of self-love that would, if acted upon, lead one to immoral action). It is not clear, then, what a perfect will would consist in over and above a good will. Kant even seems to define moral perfection, which we may suppose is at least in part achieved through possessing a perfect will, in terms rather similar to his description of good will: “[P]erfection consists subjectively in

⁶³ Dean 22.

⁶⁴ Dean 20.

the *purity* of one's disposition to duty, namely, in the law being by itself alone the incentive, even without the admixture of aims derived from sensibility, and in actions being done not only in conformity with duty but also *from duty*.”⁶⁵

Perhaps one can think of a perfect will as a will that remains good over an entire lifetime or a will can be perfect from the moment of a moral conversion on out. But then a perfect will would just be a good will that persists through time. The internal features of a good and perfect will would be identical. If a good will is equivalent to a perfect will, then it is clear that Kant thinks a good will would be rare. This reading of a good will also has the benefit of cohering with Kant's statement from the *Groundwork* suggesting that it's conceivable and perhaps even probable that true virtue is nowhere to be found among human beings.

Furthermore, Dean's contention that good wills are not all that scarce among human beings seems to rub against much of what Kant says about the attitude we should take toward our own will. Kant suggests that it is not conducive to morality to assume we have a good will. He writes, “[O]ne is never more easily deceived than in what promotes a good opinion of oneself. Moreover, it seems never advisable to be encouraged to such a state of confidence but much more beneficial (for morality) to “work out one's salvation with *fear* and *trembling*” (a hard saying which, if misunderstood, can drive one to the darkest enthusiasm).”⁶⁶ Kant suggests that the assumption that one indeed has a good will is actually often evidence of a lack of moral honesty:

⁶⁵ MM 6:446.

⁶⁶ Rel. 6:68.

This is how so many human beings (conscientious in their own estimation) derive their peace of mind when, in the course of actions in which the law was not consulted or at least did not count the most, they just luckily slipped by the evil consequences; and [how they derive] even the fancy that they deserve not to feel guilty of such transgressions as they see others burdened with, without however inquiring whether the credit goes perhaps to good luck, or whether, on the attitude of mind they could well discover within themselves if they just wanted, they would not have practiced similar vices themselves, had they not been kept away from them by impotence, temperament, upbringing, and tempting circumstances of time and place (things which, one and all, cannot be imputed to us).⁶⁷

It seems that Kant does not think we should assume that we have a good will; nor do we seem justified in supposing that good will is commonplace among human beings. So, I would argue that the first implication still holds from Dean's reading. That is, if Dean's reading is correct, then humanity would be rare and so too would be human beings who are required to be treated as ends in themselves.

A second implication of Dean's thesis is the following: If humanity is the good will, then we cannot ever know who we are required to treat as an end in itself. In Section II of the *Groundwork* Kant asserts that it is indeed impossible to know if an action was committed out of respect for the moral law, "In fact, it is absolutely impossible by means of experience to make out with complete certainty a single case in which the maxim of an action otherwise in conformity with duty rested simply on moral grounds and on the representation of one's duty."⁶⁸ Even after we have undergone "the keenest self-examination" and have still found in our maxim "nothing besides the moral ground of

⁶⁷ Rel. 6:38.

⁶⁸ G 4:407.

duty that could have been powerful enough to move us to this or that good action” we still cannot infer “with certainty that no covert impulse of self love, under the mere pretense of that idea, was not actually the real determining cause of the will.”⁶⁹ Kant expresses a similar sort of skepticism about moral self-knowledge in *The Metaphysics of the Metaphysics of Morals* when he writes, “The depths of the human heart are unfathomable. Who knows himself well enough to say, when he feels the incentive to fulfill his duty, whether it proceeds entirely from the representation of the law or whether there are not many other sensible impulses contributing to it that look to no one’s advantage...and that, in other circumstances, could just as well serve vice?”⁷⁰ This is true for our *own* wills, to which we have a degree of privileged access. It is more pressingly true in the case of others.

Furthermore, Kant argues that we have reason to *not* give to ourselves the benefit of the doubt with regard to having a good will: “[F]or we like to flatter ourselves by falsely attributing to ourselves a nobler motive, whereas in fact we can never, even by the most strenuous self-examination, get entirely behind our covert incentives, since when moral worth is at issue, what counts is not actions, which one sees, but those inner principles of actions that one does not see.”⁷¹

So, if humanity is to be treated as an end in itself and never merely as a means, and humanity is possession of a good will: a) we cannot know who we are required to

⁶⁹ *Ibid.*

⁷⁰ MM 6:447.

⁷¹ G 4:407.

treat as an end in itself, and b) we have reason to believe that in our own person we are likely to falsely attribute a good will to ourselves out of our desire to “flatter” ourselves.

Given this epistemological concern, it becomes possible that the humanity formulation would fail to provide us with any concrete requirements on how to conduct ourselves with regard to real individuals. We would know to treat humanity as an end in itself, but we would have no reliable way of determining who in particular we are to treat that way.

In the case of the second implication (That we cannot know who to treat as an end in it self), Dean claims 1) We are justified in believing that good will is common, so we know that most people, or at least many people should be treated as ends in themselves, and 2) We have good reason to treat others as ends in themselves even if we suspect that they do not have a good will. So, the epistemological impossibility of knowing whether or not someone has a good will does not lead us to abandon the idea that we are to treat most human beings as ends in themselves. As we have seen, I do not think Dean’s first claim is justified (that good will is common), so I do not think it can be used here to mitigate concerns about the second implication. This leaves the second claim he makes (That there are good reasons for treating people who lack good will with respect). I will wait in responding to this claim until I turn my attention to the third implication because the two discussions are related to each other. For now, we should simply take Dean as asserting that the epistemological problem is not a genuine worry because we do not need to know or even suspect that someone has a good will in order to be morally required to treat them as an end in itself.

A third potentially problematic implication for Dean is that his account of humanity and the good will seems to be incompatible with Kant's duties of respect. This third implication has its basis in the first implication that few people, if any, have humanity. That is, the first implication holds that the set of ends in themselves is extremely small, if perhaps empty (at least with regard to human beings). If that implication holds, then the number human beings who are owed respect as described in Kant's duties of respect is equally small. This would be a problem for Dean's reading, in my view, because of Kant's regular and consistent claims that all human beings are owed respect.

Dean's reading of humanity entails that some rational human beings are not owed the respect that Kant describes in his discussion of the "duties of respect." This, however, seems to flatly contradict Kant's more universal-sounding claim that opens his discussion of the duties of respect: "Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other."⁷² In this same passage Kant asserts that "the dignity of humanity" is "in every other human being."⁷³ Even according to Dean's more expansive reading of the good will, at least some vicious wrongdoers would not be members of humanity. If the duties of respect are grounded in one's humanity, then we would not, on Dean's reading, have duties of respect to scoundrels. Kant, though, explicitly states that vicious people too are owed respect. Kant

⁷² MM 6:462.

⁷³ MM 6:462.

writes, when dealing with such a person, “Nonetheless I cannot deny all respect to even a vicious man as a human being...”⁷⁴

Dean recognizes that Kant seems to suggest that respect is owed to all human beings and his reading might seem to conflict with this tendency in Kant. Dean, then, tries to make his reading more consistent with Kant’s more universal formulation of the duties of respect by arguing that we still have good reasons for treating human beings who lack good will with respect. So, the duties of respect would be extended to a wider group of people, including even those who lack humanity. The duties of respect would not, according to Dean, be based on humanity (or good will), but rather on something else. He writes, “[T]here are reasons to treat most humans with respect and concern, even if they do not fully earn this treatment by possessing a good will.”⁷⁵

There are two reasons that Dean outlines for why we should respect humans who lack a good will (and therefore lack humanity). The first is the following: Dean says, “First, to treat any human with disrespect lessens our respect for all humans.”⁷⁶ In defense of this claim, Dean refers to Kant’s discussion of defamation. While Dean is certainly right that Kant does argue that in cases of defamation (one of the vices of disrespect) we “cast a shadow of worthlessness over our race itself” and that defamation “diminishes respect for humanity as such,”⁷⁷ those concerns do not exhaust the wrongs and injuries that have been committed. In cases of disrespect, a person’s *humanity* has

⁷⁴ MM 6:463.

⁷⁵ Dean 7. Dean also writes, “Kant himself provides some good reasons to treat all humans with respect, even if some have not earned that respect by possessing a good will” (94).

⁷⁶ Dean 94.

⁷⁷ MM 6:466.

been undermined. Even if the person is indeed guilty of wrongdoing, *that person* was also wronged when subjected to contempt, which consists precisely in denying them their humanity. Kant writes, “Failure to fulfill mere duties of love is *lack of virtue (peccatum)*. But failure to fulfill the duty arising from the *respect* owed to every human being as such is a *vice (vitium)*. For no one is wronged if duties of love are neglected; but a failure in the duty of respect infringes upon one’s lawful claim.”⁷⁸ Dean’s reading precludes the possibility of a person lacking a good will being *wronged* as a result of being disrespected; the basis has for showing respect to scoundrels, according to Dean, is just the need to avoid the *effects* that disrespect may have on our conception for the species in general. The wrongdoer is not, himself, wronged in this case; Kant, however, is clear that in cases of disrespect an individual person’s lawful claim is infringed upon.

Kant, again, is fairly explicit that we are to avoid contempt because it would deprive a person of the respect they are owed on the basis of their being human beings: “To be contemptuous of others, that is, to deny them the respect owed to human beings in general, is in every case contrary to duty; *for they are human beings.*”⁷⁹ Kant also notes that we are rightly scandalized by ruthless treatment of criminals because we are appalled that someone belonging to our species can be treated in such a way.⁸⁰ So it seems that Dean cannot try to base the duties of respect on anything other than the respect owed to all human beings and since the duties of respect clearly extend to all human beings who

⁷⁸ MM 6:464.

⁷⁹ MM 6:463 emphasis added.

⁸⁰ MM 6:463.

lack a good will, it follows that humanity cannot be equivalent to as rare a property as good will.

The second reason Dean argues that we have good reason to treat with respect even vicious persons is that in failing to do so we would, “undermine his chances of reforming his character.”⁸¹ Dean argues that showing contempt for those who lack good will discourages the wrongdoer from moral improvement. I think Dean is correct in saying that the possibility of moral improvement in all human beings is one reason why we should not treat others with contempt. Kant, for instance, writes, “The same thing [the duty to avoid contempt] applies to the censure of vice, which must never break out into complete contempt and denial of any moral worth to a vicious human being; for on this supposition he could never be improved...”⁸² I think, though, that Dean is overlooking the fact that Kant seems to suggest that the very possibility for improvement is a consequence of an individual being a human being. In this same passage, Kant claims that the notion that a person cannot be improved is inconsistent with the very idea of a human being.⁸³

Dean goes on to claim that “This duty to show respect for others is not based on a supposition that all humans necessarily deserve respect. In fact it presupposes the opposite.”⁸⁴ This seems correct so far as it goes, but it is important to remember that Kant is hardly univocal in his usage of the term “respect” (*Achtung*). There certainly is a sense

⁸¹ Dean 94.

⁸² MM 6:463.

⁸³ *Ibid.*

⁸⁴ Dean 95.

of respect in Kant's moral philosophy that is owed only to those who have exhibited a steadfast commitment to the moral law, but there is another sense of respect that seems to be more inclusive and does not depend on meritorious action or virtue. This second sense of respect is not based on possessing a good will, but rather simply being a human being. Kant writes, "It is only through the noble predisposition to the good in us, which makes the human being worthy of respect, that one can find one who acts contrary to it contemptible (the human being himself, but not the humanity in him)."⁸⁵ Here, Kant refers to humanity as being "in" even a vicious person who has committed contemptible misdeeds. In this sense of respect, it is not a matter of "deserving" respect in the ordinary sense that someone might deserve something. A vicious person is owed respect because of an inextricable property of rational human beings in general (namely, their "noble predisposition to the good," which Kant suggests, here at least, to be their "humanity.") So here too the good will reading overlooks the fundamental role humanity places in anchoring the universalistic extension of the duties of respect.

IV. A More Refined Moral Capacity Reading?

Given these problems that arise for both the moral capacity reading and the good will reading, coupled with a recognition of the tendencies in Kant's moral philosophy to treat humanity both as an amoral capacity to set ends as well as a receptivity to morality, we may well ask: what sort of interpretation of humanity remains viable?

⁸⁵ MM 6:441.

Even though the moral capacity reading faces problems, there is much in the moral capacity reading that remains promising in providing a satisfactory interpretation of Kantian humanity. Unless we are to dismiss the morally thick tendency Kant displays in describing humanity, then we must account for humanity's necessary relation to, and ultimately capacity for, morality. Furthermore, the moral capacity reading does so in a way that does not posit that humanity must always display morally virtue. That is, a complete understanding of humanity and its role within Kant's moral philosophy must involve in some way humanity's receptivity to the moral law, without involving possession of a good will.

The problem that the capacity interpretation runs into is, I would argue, the result of not taking into account a crucial distinction in its reading of Kantian humanity. The distinction is between what I will call "humanity as such" and "humanity as it is."⁸⁶ Humanity as such requires a narrower, amoral interpretation, while humanity as it is does indeed involve the capacity for morality and receptivity to morally relevant feelings. On this reading, humanity as such is the capacity to set voluntary ends; in the case of human beings humanity also necessarily involves the capacity for morality and the accompanying receptivity to respect for the moral law, and other morally relevant feelings.

⁸⁶ My phrase "humanity as such" derives from Korsgaard's discussion of humanity in *Creating the Kingdom of Ends* (114).

Humanity As Such

Humanity as such, considered on its own, is simply possession of the rational capacity to set voluntary ends.⁸⁷ This amoral tendency is, according to Kant, the core faculty of humanity, that by which members of humanity do indeed separate themselves from the rest of nature. Humanity as such also possesses the foundational characteristic of moral action; moral action must be based on the capacity to choose ends free from the dictates of instinct. Humanity as such is captured by Kant's imaginative reconstruction of the beginnings of human history when human beings come to realize that instinct need to dictate to them their course of action: "He discovered in himself an ability to choose his own way of life without being tied to any single one like the other animals."⁸⁸ Humanity as such involves the rational deliberative process whereby individuals make choices about what sorts of ends are worth pursuing, but this sort of activity can, in principle, occur without reference to morality.

Humanity in the Moral World (Humanity as it is)

While humanity as such is conceptually distinct from the exercise of freedom considered from a moral perspective, in the case of human beings, the activity of humanity as such always and necessarily must be situated within a moral context. Furthermore, the choices that mark the capacity for setting ends must always, at the same

⁸⁷ Korsgaard writes, "[T]he distinctive feature of humanity, as such, is simply the capacity to take a rational interest in something: to decide, under the influence of reason, that something is desirable, that it is worthy of pursuit or realization, that it is to be deemed important or valuable, not because it contributes to survival or instinctual satisfaction, but as an end—for its own sake" (*Creating the Kingdom of Ends* 114).

⁸⁸ Conjectures on the Beginning of Human History 224.

time, be made in some conscious awareness of the requirements of the moral law.

Actually existing humanity (humanity as it is), involve humanity's core activity of setting ends, but it also involves, in the case of human beings, morally relevant actions.⁸⁹

Whenever we encounter an exercise of one's humanity in a particular instance, it must exist in some morally relevant state.

Any and all exercises of humanity as such (i.e. rational end-setting), are also instances of humanity as it is, here in this moral world (that is, an end-setting action that is subject to moral consideration).⁹⁰ While this fact brings together humanity as such and humanity as it is (exercised within a moral framework), the connection is empirically based, and therefore not conceptually necessary. It is possible, I would argue, to draw a necessary and conceptual link between humanity as such and humanity as it is. That is, humanity as such would not be possible without the capacity for morality. If we define humanity as the capacity to set ends *voluntarily*, then we must think of humanity as *freely* choosing its desired ends. For Kant choices are free only when they come as the result of the exercise of one's free will. That is, if nature is dictating to you the ends you adopt, then one is not freely choosing them; they are being adopted *involuntarily*. If one goes about setting one end over another, one must have some way of adjudicating between two possible ends. Without morality providing an alternative set of ends, then one's inclinations will sort themselves out according to the relative power they are able to

⁸⁹ I use the phrase "actually existing" because while Kant does claim that it is possible to think of a perfectly rational agent setting ends without any consideration or relation to the moral law (Rel. 6:28), Kant also suggests that there are actually no instances of this sort of humanity (Rel. 6:36).

⁹⁰ It is possible that there can be cases of humanity as such without humanity as it is in non-human creatures perhaps, but Kant thinks that human beings will always be in a position of expressing their humanity in morally relevant contexts.

exercise over a person's decision-making. Choosing one end over the other, therefore, would not really be the result of authentic volition, but rather the result of a natural process whereby inclinations compete with each other in a power struggle. The notion of voluntarily setting ends, on this picture, is weakened down to the idea that no other *person* is forcing you to set one end over another. It is still true, though, that nature (that is, the predispositions nature has assigned to you) is dictating the ends you adopt. A person may, un-coerced by another person, choose to forego that piece of cherry pie in favor of a healthy apple, and that decision may seem voluntary, but really it is simply the case that the person's inclination to vanity (let us say) is determining his actions (summer is coming and he wants to cut a certain figure on the beach in order to impress people). In this example, inclinations are determining the person's decision-making process.

While for a compatibilist like David Hume, this is a sufficient account of free or voluntary action, for someone like Kant, who has a robust theory of free agency, this is a limited, incomplete, and, ultimately, illusory case of voluntary action. Truly voluntary actions require the possibility of grounding one's power of choice on the moral law because only the moral law provides an authentic alternative to the principle of self-love. On Kant's picture, the possibility of morality requires the postulating of a noumenal, free self exercising choice in a way that determines the choices we make in the phenomenal world. Choosing to forego the piece of cherry pie on prudential grounds would be an exercise of one's free will only if there is some noumenal choice that ultimately grounds and explains the choice. Without the possibility of noumenal choice, every phenomenal choice would be determined according to natural laws (be they laws of physics,

psychology, or something else). Whether or not one actually chooses the moral law, without the moral alternative to nature, humanity would collapse back to mere animality (albeit of a more complicated sort).

This morally thick conception of humanity is consistent with Kant's regular comments that humanity is something that can be improved, cultivated, and developed.⁹¹ Kant at times describes humanity as a goal or aim to which we must aspire. Moral progress is understood as increasingly coming to realize one's own humanity. In the *Vigilantius Lectures*, Kant says, "Humanity itself, if we wished to personify it, actually lacks any inclination to evil, but the more a man compares himself therewith, the more he finds out how far away he is from this."⁹² Here humanity seems to be an ideal that we compare ourselves with in order to determine the extent of our own moral shortcomings. It is not, at least here, simply a property "in" us, but something that is outside of us, something to aim towards in our moral development. The capacity to set ends (humanity as such) does seem to simply reside "in" us in a constant and unchanging way, but this seems to be only the bare precondition for a full flowering of our humanity that undergoes, potentially, radical change and improvement.

In Kant's published works we also get instances in which he treats humanity as a goal. In the *Metaphysics of Morals* while discussing moral education he writes, "[I]t is not comparison with any other human being whatsoever (as he is), but with the *idea* (of

⁹¹ Whereas a weak, amoral, and reductive account of humanity in Kant's philosophy would have trouble making room for, and reconciling its account with, Kant's tendency to treat humanity as a kind of moral ideal.

⁹² *Vigilantius Lectures* 27:609.

humanity), as he ought to be, and so comparison with the law, that must serve as the constant standard of a teacher's instruction."⁹³ Moral education and culture generally function as the means by which humanity is cultivated and perfected in human beings. While humanity as such must underlie any and all actual exercises of humanity (humanity as is), the particular sorts of ends that human beings choose may change and improve radically over the course of both an individual life as well as of the history of the species. Humanity is not simply a static, unchanging property in human beings. It undergoes change, development, and improvement.

The exercise of free choice undergoes change; that is, our very usage of our capacity to set ends develops and improves. Individual human beings undergo a process of development whereby they learn to exercise their humanity in increasingly cultivated ways. In his *Lectures on pedagogy*, Kant writes, "The human being is the only creature that must be educated. By education we mean specifically care (maintenance, support), discipline (training) and instruction, together with formation. Accordingly, the human being is first infant, then pupil, and then apprentice."⁹⁴ The very concept of humanity (humanity as such) leads to humanity as it that is in some level or degree of completion. Only through education does one realize their humanity: "Discipline and training changes animal nature into human nature."⁹⁵ Training and instruction allows the human being to formulate their own course of action and to shape the sort of life he or she will lead. Kant again contrasts the human situation and the animal situation by claiming that the potential

⁹³ MM 6:480.

⁹⁴ Lectures on pedagogy 9: 441.

⁹⁵ *Ibid.*

of non-human animals is determined by instinct: “An animal is already all that it can be because of its instinct; a foreign intelligence has already taken care of everything for it. But the human being needs his own intelligence. He has no instinct and must work out the plan of his conduct for himself.”⁹⁶

This is true in both the lives of individual human beings as well as for the species as a whole. Once human beings have discovered within themselves the capacity to choose their own ends, questions immediately arise as to which ends should be pursued and which should not. In his imaginative reconstruction of the beginning of human history, Kant expresses the immediate need for human beings to decide how their newfound freedom from instinct is to be best expressed: “But the momentary gratification which this realization of his superiority may have afforded him was inevitably followed at once by anxiety and fear as to how he should employ his newly discovered ability...”⁹⁷ Early attempts in exercising freedom may well be crude, uncivilized, and even immoral, but through the gradual improvement of one’s humanity, one’s humanity can become more in line with morality and with the promotion of humanity as an end in itself. Kant expresses the idea that humanity (i.e. the free setting and pursuing of ends) can only begin to be perfected through preliminary cruder attempts: “I admit that I am not comfortable with this way of speaking, which even clever men are wont to use: “A certain people...is not ripe for freedom” ... For on this assumption freedom will never come, since we cannot *ripen* to it if we are not already established in it (we must be free in order to be able to make use of our powers purposively in freedom). To be sure, the

⁹⁶ Lectures on pedagogy 9:441.

⁹⁷ Conjecture on the Beginning of Human History 224.

first attempts will be crude...yet we do not ripen to freedom otherwise than through our *own* attempts (and we must be free to be allowed to make them).”⁹⁸

An analogy may clarify the point here. When people are first learning how to read, their capacity to read exists, but only at a basic or crude level. As they develop into better readers, their reading capacity undergoes improvement. While we can rightly claim that the elementary reader and the highly developed reader are doing the same thing (i.e. reading), we can also rightly claim that there are significant differences in the exercise of the highly developed reader’s capacity to read and the basic reader’s capacity to read. The same can be said about the exercise of one’s humanity. As individuals exercise their freedom, with the constant improving tendency our relation to the moral law has on our humanity and with the stimulation and impetus for growth that culture generally allows for, human beings can move from crude and immoral uses of their humanity, to cultivated and moral ones.

One may raise two related concerns against my layered, multi-faceted account of humanity. Given that Kant seems to be of two minds about humanity (at times as amoral end-setting, at other times as a more morally thick notion regarding our capacity for and failures in morality), does Kant have a unified account of humanity at all? The related concern is that perhaps my reading of humanity fails on the same grounds as the basic moral capacity reading does; namely, what I describe as humanity as is or the morally thick conception of humanity is really personality. Humanity would remain the amoral

⁹⁸ Rel. 6:188n.

capacity to set ends, while personality, and not humanity, would include the capacity for morality.

I maintain that both concerns can be answered adequately within my account. First, I would argue that Kant does indeed have a unified account of humanity. The core activity of humanity, on my reading, is the capacity to set voluntary ends. Free immoral action, on my account, is still an exercise in one's humanity, just as free moral action is. In the case of the free actions of human beings, though (what I call 'humanity as is'), free actions are inseparable from moral considerations. In the case of human beings, there are morally acceptable and laudable exercises of freedom as well as morally unacceptable or blameworthy ones. So the very activity of setting ends must necessarily have a moral dimension.

One may retain the merest sort of humanity by neglecting one's duty, but built into humanity is an inherent tendency toward its completion, promotion, and perfection: "Now there are in humanity predispositions to greater perfection, which belong to the end of nature with respect to humanity in our subject; to neglect these might admittedly be consistent with the *preservation* of humanity as an end in itself but not with the *furtherance* of this end."⁹⁹ Moral actions are only possible if they were freely chosen (i.e. they were an exercise in one's humanity). In any exercise of a free act (i.e. setting an end and pursuing it) by a human being, one must consider whether the action is consistent with morality. Furthermore, in every case of humanity as is, the moral law will present itself as something that should be incorporated into one's maxim. This is the case for all

⁹⁹ G 4:430.

human beings, even those who engage in contemptible behavior: “There is no one—not even the most hardened scoundrel, if only he is otherwise accustomed to use reason—who, when one sets before him examples of honesty of purpose, of steadfastness in following good maxims, of sympathy and general benevolence (even combined with great sacrifices of advantage and comfort), does not wish that he might also be so disposed.”¹⁰⁰

The second concern (that humanity as is and personality are the same thing) can also be allayed. First, humanity as is and personality are not strictly identical because humanity exists on a wide-ranging continuum whereas personality occupies only the extreme end of the continuum. That is, there can be instances of humanity as is that fall somewhere on the continuum between animality and personality that do not fall onto the personality end of the spectrum. Kant emphasizes that there can be free actions that depart radically from the requirements of morality, and yet the agent involved has nevertheless retained his or her humanity. In the *Religion*, Kant maintains that humanity involves “a self-love which is physical.”¹⁰¹ He then asserts that the exercise of humanity is consistent with vices such as “envy, ingratitude, joy in others’ misfortunes...”¹⁰² Such vices, though, are not possible expressions of one’s personality. Personality can only be expressed through one’s will being determined objectively by the moral law itself and subjectively by pure respect for the moral law. In cases of envy, for example, one’s level

¹⁰⁰ G 4:454.

¹⁰¹ Rel. 6:27.

¹⁰² Rel. 6:27.

of humanity is in an undeveloped, immoral state. One's humanity as is remains in a state well short of the achievement of personality.

Second, it is not a problem with my reading to say that humanity and personality can *in some cases* be identical to each other. Humanity can exist in various degrees or levels. At the very bottom, perhaps not even realized in the natural world, is the amoral level humanity. This consists merely in an agent's capacity to choose ends voluntarily and not on the sole basis of instinct. An individual's humanity, in its fully developed form, would be the attainment of a good will or perfect moral personality. Korsgaard makes a similar point when she writes, "Humanity, completed and perfected, becomes personality, so that in treating the first as an end in itself we will inevitably be led to realize the second."¹⁰³ Personality, on this reading, is a peculiar form of humanity that sets its ends in a way that the maxim upon which the person acts in pursuing those ends is capable of universal legislation.

Personality, understood as our capacity to respond to moral incentives, must retain, when expressed in human beings, the capacity to set and choose ends. Personality, in human beings, is achieved through a specific sort of exercise of one's freedom (i.e. of one's humanity); furthermore, personality consists precisely in our capacity to set ends being directed toward, not arbitrary and conditionally valuable ends, but rather necessary and unconditionally valuable ones. Personality is only exercised when respect for the moral law operates as the subjective determining factor of one's will. Personality is realized when one the moral law and pure respect for it are the determining grounds for

¹⁰³ *Creating the Kingdom of Ends* 114.

one's maxim: "The idea of the moral law alone, together with the respect that is inseparable from it, cannot be properly called a *predisposition to personality*; it is personality itself."¹⁰⁴

So while humanity is not in all cases identical with personality, it can be said that humanity, fully developed and perfected, is identical with personality.¹⁰⁵ In the *Religion*, for instance, Kant describes personality as "the idea of humanity considered wholly intellectually."¹⁰⁶

Humanity as is always sets ends from the perspective of an agent with a predisposition to the good or to personality. In the *Religion* Kant describes the predisposition to personality: "The predisposition to personality is the susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice*."¹⁰⁷ The disposition to personality is present in all human beings that are capable of setting ends, even wicked ones. In the *Religion*, for example, Kant writes, "The human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it). The law rather imposes itself irresistibly, because of his moral predisposition; and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e. he would be morally good."¹⁰⁸ The law imposes itself irresistibly on all of human

¹⁰⁴ Rel. 6:28.

¹⁰⁵ Koorsgaard makes this point when she writes, "Humanity, completed and perfected, becomes personality, so that in treating the first as an end in itself we will inevitably be led to realize the second" (*Creating the Kingdom of Ends* 114).

¹⁰⁶ Rel. 6:28.

¹⁰⁷ *Ibid.*

¹⁰⁸ Rel 6:36.

beings and, as a result, even if a person has totally subverted his or her will and has privileged self-love over morality, morality still presents itself to that person as something that should be acted upon. Our very capacity to set ends is inextricably related to our moral predisposition precisely because the moral law introduces constraints and demands about what sorts of ends we can legitimately willfully pursue.

As a consequence of our predisposition to personality human beings are receptive to morally relevant affective states. Humanity is always exercised with the morally relevant feelings impinging on our affective states, either strongly as in cases where one is tempted to commit an immoral act or weakly in cases where one seems to be choosing between morally neutral alternatives. In instances where the determination of our will was formed with reference to respect for the moral law (and only in cases like these), personality and humanity become identical. An expression of one's humanity becomes also an expression of one's personality, and vice versa.

Humanity, on this multi-faceted account, has as its core feature the capacity to set and choose ends. Humanity is also inseparable, in the case of human beings, from the improvement and cultivation of that very same capacity to set and choose ends. Humanity exists in a cultivated state when a person exhibits fine taste and culture, as well as sound moral virtue. Humanity, when perfected, also promotes the humanity of others so that individuals are able to pursue ends that in turn promote one's own humanity, as well as the humanity of others.¹⁰⁹ Humanity, then, is both a static, constant feature of (rational)

¹⁰⁹ In his paper "Setting and Pursuing Ends: Internal and External Freedom," Paul Guyer develops an account for how exactly Kant's conception of humanity coheres with the various formulations

human life and a dynamic process that is an ongoing project in human life. Humanity is a constant in human life because rational human beings always have the capacity to set and pursue ends of their own choosing; it is dynamic because it is always expressed in some determinate way, which can be cultivated and improved throughout one's life.

Humanity, then, is one way in which human beings exist in a state of moral precariousness. Humanity can be exercised in a way that is consistent with morality, and an expression of our commitment to the moral law, but it can just as easily be exercised in a manner that works only to promote our happiness (at the expense of morality). Humanity is a necessary feature of a morally good life (for human beings), but humanity also figures crucially in a morally problematic life. Without the freedom to choose our own ends, we cannot have an authentic moral status (either for good or for ill). Humanity, in its capacity for *freely* setting ends, must be protected and promoted because freedom is the only source of value in the world.¹¹⁰ But, by that same token, humanity is a necessary component of an immoral human life in which a person sets and pursues ends that conflict with the requirements of morality. Humanity, as we will see in the next chapter, occupies a tenuous place on a continuum with animality and personality as its polar ends.

of the categorical imperative and his system of duties (developed in *The Metaphysics of Morals*). On Guyer's reading, Kant's account of humanity, combined with the second formulation of the categorical imperative, leads to the following moral obligation: "[O]ne must always use one's own capacity to set ends in a way compatible with the greatest possibility of other exercises of this capacity by oneself and others, that is to say, in a way compatible with the greatest possible use of freedom."

¹¹⁰ In his Lectures on Ethics from the 1784, Kant is quoted by his student Georg Ludwig Collins as saying, "Freedom...is the capacity which confers unlimited usefulness on all others. It is the highest degree of life. It is the property that is a necessary condition underlying all perfections." He goes on, "[T]he inner worth of the world, the summum bonum, is freedom...Freedom is thus the inner worth of the world."

CHAPTER 2

Dignity in Kant's Moral Philosophy

Introduction

The notion that human beings have dignity, and that morality requires us to treat others and ourselves with respect seems to be a broadly held and deeply felt moral intuition. The Preamble to the Universal Declaration of Human Rights says, “the inherent dignity and the equal and inalienable rights of all members of the human family is the foundation of freedom, justice and peace in the world.” According to the Universal Declaration, then, all human beings possess dignity; dignity is not limited to certain persons, but rather extends to “all members of the human family.” The importance of human dignity is also affirmed in the Third Geneva Conventions, which prohibits, “outrages on personal dignity” against prisoners of war. Social and political movements from all over the world often appeal to the conscience of other people in terms of human dignity and mutual respect. One of the striking features of these appeals to human dignity is their democratic, egalitarian nature. Dignity is not something earned through meritorious action, nor is it claimed on the basis of noble birth or high social standing.

When considering such deployments of the concept of dignity, it is reasonable to look back to the history of moral thought to find antecedents and forerunners. Such historically-oriented individuals would seem to find in Kant's moral philosophy an earlier formulation of the same concept of dignity. With Kant's frequent and regular invocation of the concept of human dignity, it might seem clear that Kant, in part, laid the

foundation for modern conceptions of human dignity. The truth, though, is a bit more complicated than it might initially appear. Kant's account of human dignity bears some clear resemblances to our shared, modern conception of dignity, but there are interesting differences that we find as well. This chapter is an analysis Kant's concept of dignity (especially human dignity).

There are features of Kant's moral philosophy that seem to resonate with moral sentiments of this sort. That is, Kant's moral philosophy seems to make similar claims about the importance of human dignity. As we will see, Kant repeatedly makes reference to the dignity of humanity, of human nature, and of human beings. This chapter examines Kant's views on human dignity. It argues that Kant affirms human dignity largely in perfectionist terms regarding our capacity for freedom. It also argues that Kant's perfectionist understanding of human dignity nevertheless is ultimately broadly inclusive and universal. That is, even though Kant claims that dignity rests upon the capacity for, and proper use of, freedom, dignity is widely possessed, universally so, among rational human beings.

In Section I, I will begin my examination of Kantian dignity by noting some of the objects to which the term applies. I review passages where Kant claims that an object or a thing has dignity. I emphasize the relationship between dignity and morally relevant concepts, showing that Kant repeatedly claims that morality, the moral law, human nature, human beings, and humanity all possess dignity.

In Section II, I lay out a general definition for Kantian dignity, emphasizing its relational component. That is, something has dignity when it is *morally elevated* over and distinguished from something else. I also discuss the connection Kant draws between sublimity and dignity. I claim that sublimity is a kind of dignity and that our moral being is both dignified and sublime on the basis of its superiority over nature.

In Section III, I examine in further detail what it is about morality and humanity that has dignity. That is, I examine the intrinsic properties of morality and humanity by virtue of which they are elevated. I conclude by arguing that they both possess a superiority over mere nature. I then consider the ways in which humanity can be understood as being superior to nature and its relation to the moral law.

Section IV defends the claim that human dignity, for Kant, is based on our capacity for freedom. And finally, Section V considers the relationship between respect and Kantian dignity. Here, I try to reconcile the perfectionist tendencies in Kant's discussions of human dignity with his inclusive understanding of our duty to respect all people. I conclude that Kant's account of dignity shares with our modern notion of human dignity the characteristics of universality and inalienability, but, unlike the modern concept of human dignity, Kant grounds his notion of dignity in the ability of human beings to attain virtue.¹¹¹

¹¹¹ In his book *Kant on Human Dignity*, Oliver Sensen makes a similar claim in which he argues that Kant's views on dignity has both modern and pre-modern features. Pre-modern accounts of dignity claim that dignity is less widely distributed and is limited to special human beings who are elevated above ordinary human beings. Modern notions of dignity tend to be more egalitarian and not merit-based (see Oliver Sensen, *Kant on Human Dignity*. De Gruyter: Boston, MA, 2011). My discussion of dignity departs with Sensen's approach slightly in emphasis. In this

I. What Has Dignity?

Kant repeatedly makes a connection between dignity and morality. That is, dignity, for Kant, is a moral concept. He makes this connection between dignity and morality in several ways: He claims that morality, the moral law, and duty have dignity. In the *Groundwork*, Kant claims that morality has dignity¹¹² and that it is only through the lens of morality that anything whatsoever can possess dignity. In the second *Critique* he describes the task of reason, in part, as the need to “work itself up as to gather strength to resist inclinations by a lively representation of the dignity of the law...”¹¹³ Also in *The Metaphysics of Morals* he warns against raising mere custom to “the dignity of the law.”¹¹⁴ In addition to the moral law’s dignity, Kant also claims that humanity or human nature has dignity. In the *Groundwork*, Kant writes the following: “[M]orality, and humanity insofar as it is capable of morality, is that which alone has dignity.”¹¹⁵

The idea that human nature itself has dignity seems to have been present in Kant’s moral thinking from a rather early stage. In *Observations on the feeling of the beautiful and sublime*, Kant claims that “true virtue” must be based on principles, which “are not speculative rules, but the consciousness of a feeling that lives in every human breast and that extends much further than to the special grounds of sympathy and complaisance. I believe that I can bring all this together if I say that it is the **feeling of the beauty and the**

chapter I claim that Kant’s notion of dignity does indeed have a merit component, but I argue that, on the whole, Kant’s account of human dignity is more modern than pre-modern. All human beings, as long as they are rational (with their predisposition to the good in place), never lose their dignity.

¹¹² G 4:435.

¹¹³ CPrR 5:147.

¹¹⁴ MM 6:464.

¹¹⁵ G 4:435.

dignity of human nature.¹¹⁶ In the *Groundwork*, Kant also writes of the “dignity of human nature and of every rational nature.”¹¹⁷ In the *Religion*, Kant again refers to the “dignity of human nature” and praises the Stoics for correctly, in his view, basing their moral philosophy on it, claiming, “they [the Stoics] could not have laid down a better or nobler principle for foundation.”¹¹⁸

In the second *Critique*, Kant references the dignity of the *human being*. He claims that what he calls the “pure moral motive...teaches the human being to feel his own dignity.”¹¹⁹ Kant also repeatedly refers to the dignity of *humanity*. In *the Metaphysics of Morals*, for instance, Kant refers to “the dignity of humanity within us” when discussing our duty to avoid servility.¹²⁰ These examples¹²¹ make plain, in my view, that Kant was committed to the view that the moral law, morality, humanity, human nature, and human beings have dignity. We now can begin to think about what Kant means by the dignity of the moral law or the dignity of humanity.

II. Dignity, Sublimity, and Elevation

Dignity and sublimity are closely related concepts in Kant’s moral philosophy. Human dignity is expressed, we will see, most pointedly when we act out of respect for the moral law, and thereby check the influence of merely natural inclinations in the determination of the ends we set. An action that is expressive of human dignity, for Kant,

¹¹⁶ OBS 2:217. See also OBS 2:219.

¹¹⁷ G 4:440.

¹¹⁸ Rel. 6:57-58n.

¹¹⁹ CPrR 5:152.

¹²⁰ MM 6:436.

¹²¹ Some other instances of Kant mentioning the dignity of humanity are: MM 6:423, MM 6:427, MM 6:444, MM 6:459.

can rightly be called sublime. Kant writes, for instance, “Among moral qualities, true virtue alone is sublime.”¹²² A brief examination of some of Kant’s remarks about the sublime provides a way to understand Kantian dignity.

The sublime, for Kant, is, “that which is **absolutely great**,” and also “that **which is great beyond all comparison**.”¹²³ The truly sublime, according to Kant, is incapable of being measured: “If...we call something not only great, but simply, absolutely great, great in every respect (beyond all comparison), i.e., sublime, then one immediately sees that we do not allow a suitable standard for it to be sought outside of it, but merely within it.”¹²⁴ It is impossible to bring to bear any “outside” standard regarding the sublime. The sublime cannot be broken down into terms of something else. The magnitude of the sublime is somehow radically untranslatable into sensory, empirical categories of quantity or magnitude.

In his moral writings, Kant describes the moral law and humanity (particularly its capacity for morality) in terms of sublimity. In *Observations on the feeling of the beautiful and sublime*, Kant writes, “Among moral qualities, true virtue alone is sublime.”¹²⁵ A bit later, Kant claims that virtue becomes more sublime as it becomes nobler.¹²⁶ In *The Metaphysics of Morals*, Kant describes the moral predisposition in humanity as sublime.¹²⁷ In the second *Critique* Kant claims that the moral law possesses

¹²² OBS 2:215.

¹²³ CPJ 5:248.

¹²⁴ CPJ 5:520.

¹²⁵ OBS 2:215.

¹²⁶ “Thus virtue can be only grafted upon principles, and it will become the more sublime and noble the more general they are” (OBS 2:217).

¹²⁷ MM 6:435.

“sublimity” and that it is a rather serious error to explain moral or practical concepts solely in terms of “experiential consequences (so-called happiness).”¹²⁸ A bit later in the second *Critique*, Kant refers to self-sacrificing actions done “for the sake of duty alone” as “*noble and sublime*.”¹²⁹ Also, in the *Religion* Kant writes that the respect we feel for the law “rouses a *feeling of the sublimity* of our own vocation...”¹³⁰ So Kant seems to apply both sublimity and dignity to similar sorts of things.

The reason for this convergence of terms might be explained by Kant’s tendency to think of both the sublime and the dignified in terms of *elevation*. Elevation is a metaphor that Kant consistently applies both to sublimity and to dignity.¹³¹ The sublime is a “disposition of the mind” that is always “striving to advance to the infinite.”¹³² Kant describes the sublime as follows: “**That is sublime which even to be able to think of demonstrates a faculty of the mind that surpasses every measure of the senses.**”¹³³ The sublime, then, provides us with proof, or a demonstration, that there must be a faculty of the mind that “surpasses” our senses’ capacity to measure.

This is particularly acutely dramatized in the case of the dynamical sublime, which makes us experience the receding power of our natural selves in the face of the irresistible power of nature and, at the same time, the retaining of our capacity for morality: “[I]t [the dynamical sublime] reveals a capacity for judging ourselves as

¹²⁸ CPrR 5:70.

¹²⁹ CPrR 5:85.

¹³⁰ Rel. 6:22n.

¹³¹ For a similar discussion about the centrality of the language of elevation in Kant’s discussions of sublimity and dignity see Oliver Sensen “Kant’s Conception of Human Dignity.”

¹³² CPJ 5:250.

¹³³ CPJ 5:250.

independent of it [nature] and a superiority over nature on which is grounded a self-preservation of quite another kind than that which can be threatened and endangered by nature outside us, whereby the humanity in our person remains undemeaned even though the human being must submit to that domination.”¹³⁴

As the sublime awakens in us recognition of a faculty of the mind that surpasses every measure of the senses, human dignity rests upon the existence of the sublime self. In the *Groundwork*, Kant claims that if one is presented with a truly moral act, then one’s soul “elevates” and strives to act in a similar fashion.¹³⁵ Human beings, when engaged in a good moral act, achieve sublimity. Kant also claims that a will with pure respect for the moral law is a “cast of mind” that is to be “cognized as dignity” and is “infinitely above all price, with which it cannot be brought into comparison or competition at all without, as it were, assaulting its holiness.”¹³⁶ Similarly in the second *Critique*, Kant writes that moral action, results in an “*elevating*” feeling. In *The Metaphysics of Morals*, Kant writes that humanity, through “our capacity for internal lawgiving” and our reverence for the “(moral) human being within” our own person, we feel an “*exaltation* of the highest self-esteem...”¹³⁷ A bit later in the same work, Kant writes that awareness of our moral and “sublime vocation” brings about an “*elation of spirit*.”¹³⁸ In the *Religion*, Kant writes, “Now it is our universal human duty to *elevate* ourselves to this ideal of moral

¹³⁴ CPJ 5:261.

¹³⁵ G 4:410n.

¹³⁶ G 4:435.

¹³⁷ MM 6:436. In his notes on moral philosophy, Kant also mentions the “Elevation of humanity” when people are properly motivated in accordance with authentic inner worth (*Notes on moral philosophy* 19:185-186).

¹³⁸ MM 6:437.

perfection” he also claims that the ideal of holiness is something to which we must strive to “raise” ourselves.¹³⁹

In a way similar to how the sublime is beyond all comparison, dignified things, like the moral law and humanity, stand above any equivalence, or even comparison, with the satisfaction of inclinations. The value of morality cannot be translated in terms of inclination-satisfaction; the dignity of humanity, also, is not measurable in terms of inclination-satisfaction. In fact, no amount of happiness (i.e. inclination satisfaction) can be exchanged, at equal value, for one’s humanity. In the *Groundwork*, Kant writes, “In the kingdom of ends everything has either a *price* or a *dignity*. What has a price can be replaced by something else as its *equivalent*; what on the other hand is raised above all price and therefore admits of no equivalent has a dignity.”¹⁴⁰ If something has a dignity, then, that thing cannot be replaced by something else (by a different sort of thing) of equal value. Its value is radically incommensurable with other sorts of things. Kant makes a similar point in his notes on moral philosophy when he writes, “A value: for which something else can be given as an equivalence. Virtue has no price. Dignity is the inner value, which therefore has no price.”¹⁴¹ In the *Groundwork*, Kant applies this basic point to the distinction between the demands of morality and the desire for satisfaction of inclinations; he claims that those things that can satisfy human inclinations have a price, but not a dignity: “What is related to general human inclinations and needs has a *market*

¹³⁹ Rel. 6:61.

¹⁴⁰ G 4:434.

¹⁴¹ *Notes on moral philosophy* 18:159.

price.”¹⁴² By contrast, if something has a dignity, then it admits of no equivalent in terms of pleasure: “[T]hat which constitutes the condition under which alone something can be an end in itself has not merely a relative worth, that is, a price, but an inner worth, that is, *dignity*.”¹⁴³ Here, dignity is extended to “the condition under which alone something can be an end in itself.” In other words, that in virtue of which it is possible for something to be an end in itself has dignity.¹⁴⁴ Kant then extends the concept of dignity to humanity explicitly: “Now, morality is the condition under which alone a rational being can be an end in itself, since only through this is it possible to be a lawgiving member in the kingdom of ends. Hence morality, and humanity insofar as it is capable of morality, is that which alone has dignity.”¹⁴⁵ So, both the moral law and humanity have dignity and are both above and beyond price.

Dignity, then, is a peculiar sort of property. It is not a straightforward intrinsic property things have, but rather is a property something has only considered in relation with something else. Oliver Sensen also makes this point when he writes, “Ontologically ‘dignity’ refers to a relational property of being elevated, not to a non-relational value property. ‘X has dignity’ is another expression for ‘X is elevated over Y.’”¹⁴⁶ Something has dignity only in comparison with things that merely have a price. A dignified thing,

¹⁴² G 4:434-5.

¹⁴³ G 4:435.

¹⁴⁴ This point will return below when we consider the connection between dignity and a human being’s status as an end in itself.

¹⁴⁵ G 4:435.

¹⁴⁶ Oliver Sensen, “Kant’s Conception of Human Dignity,” *Kant-Studien* 100 (2009), pp. 309-331.

then, would be something that is elevated over anything that merely has a price, but not an inner worth.

III. Dignity, Morality, and Humanity

A natural question to ask here would be the following: What is it about the moral law and humanity that gives them dignity? The answer, hinted at just now, is that both morality and humanity possess properties that transcend or rise above things that are amenable to price. Put another way: They possess dignity on account of their elevation over and their being superior to any satisfaction of one's inclinations.

Morality requires us to set aside completely the force of our natural inclinations and instead to act out of a sense of duty. In the *Groundwork*, Kant describes the paradigmatic moral action as follows: “[A]n action of integrity done with steadfast soul, apart from every view to advantage of any kind in this world or another and even under the greatest temptations of need or allurements...”¹⁴⁷ No amount of advantage or of satisfaction of inclinations can, in any way, ground or motivate one's will in a truly virtuous action. That is, a moral act is done from duty, not from inclination. Earlier in the *Groundwork*, Kant provides his famous definition of duty: “[D]uty is the necessity of an action from respect for law.”¹⁴⁸ An action from duty, furthermore, “is to put aside entirely the influence of inclination and with it every object of the will...”¹⁴⁹ A moral agent, when committing a moral action, possesses a will that is determined by something over and above our natural concern for happiness (i.e. the satisfaction of our inclinations),

¹⁴⁷ G 4:410n.

¹⁴⁸ G 4:400.

¹⁴⁹ *Ibid.*

namely, the moral law itself and rational respect for it: “[H]ence there is left for the will nothing that could determine it except objectively the *law* and subjectively *pure respect* for this practical law, and so the maxim of complying with such a law even if it infringes upon all my inclinations.”¹⁵⁰ This rather stark formulation is an expression of Kant’s view that morality has dignity because the moral law requires a total repudiation of the inclinations in forming and motivating our actions, in favor of a higher moral motivation.

In fact, Kant claims that the sublimity of moral action is augmented, or at least made clearer, in cases where there are powerful inclinations weighing against doing one’s duty: “[T]he sublimity and inner dignity of the command is all the more manifest the fewer are the subjective causes in favor of it and the more there are against it, without thereby weakening in the least the necessitation by the law or taking anything away from its validity.”¹⁵¹ The command of the moral law is sublime and possesses inner dignity because we recognize its validity in all cases, even when there are powerful inclinations standing in opposition to it. While harmonious inclinations can accompany a moral action (for instance, one can, indeed should, feel benevolence for others while engaged in charitable giving, so long as one is motivated only by respect for the law), inclinations can nevertheless play no role in forming the maxim upon which one acts. Furthermore, if acting on a moral imperative would run afoul of every single inclination to happiness we may have, we must still act on the moral imperative, and thereby frustrate *all* our inclinations. Morality, then, possesses dignity in part through its opposition to and superiority over the dictates of the inclinations.

¹⁵⁰ G 4:400-401.

¹⁵¹ G 4:425.

We have seen that the ground for morality's dignity consists in its superiority over the merely natural (i.e. inclinations). Its necessity and validity obtain even in cases where all inclinations are opposed to it. We now turn to the ground of humanity's dignity. In order for the claim "humanity has dignity" to be comprehensible, one has to understand what Kant means by "humanity." A helpful way to introduce Kant's concept of humanity is by reflecting on its relationship with what Kant calls "animality." In the *Religion*, Kant distinguishes between three "original predispositions."¹⁵² An original predisposition is a trait that something has simply by being the sort of thing it is. In the *Religion*, Kant writes, "They [original predispositions] are *original*, for they belong to the possibility of human nature." He goes on to write, "By the predispositions of a being we understand the constituent parts required for it as well as the forms of their combination that make for such a being."¹⁵³ The three original predispositions Kant lays out in the *Religion* are the following: animality, humanity, and personality. For our purposes here we need only focus our attention on the distinction between animality and humanity (we will consider in some detail the distinction between the predisposition to humanity and to personality a bit later). The predisposition to animality consists in those drives and tendencies we have in virtue of our belonging to the natural world and of having what Kant calls a "sensuous nature"¹⁵⁴ (*Sinnlichkeit*): "The predisposition to animality in the human being may be brought under the general title of physical or merely mechanical self-love, i.e. a love for which reason is not required."¹⁵⁵ Animality is non-rational. That is, one can have a

¹⁵² Rel. 6:26-28.

¹⁵³ Rel. 6:28.

¹⁵⁴ Rel. 6:35.

¹⁵⁵ Rel. 6:26.

predisposition to animality without possessing reason. Kant introduces the capacity for reason with the predisposition to humanity. The predisposition to humanity concerns human beings' "*rational being*" in addition to its animal being.¹⁵⁶ Rationality, then, is the characteristic that distinguishes humanity from animality. Kant at times uses the terms 'humanity' and 'rational nature' interchangeably.¹⁵⁷ This is especially evident in the *Groundwork*, where, for instance, Kant refers to "the mere dignity of humanity as rational nature..."¹⁵⁸ Rational nature is expressed in choosing and setting ends. In *The Metaphysics of Morals*, Kant writes, "The capacity to set oneself an end—any end whatsoever—is what characterizes humanity (as distinguished from animality)."¹⁵⁹ Kant makes a similar point in the *Groundwork*, when he writes, "Rational nature is distinguished from the rest of nature by this, that it sets itself an end."¹⁶⁰

The capacity to set and choose ends means, for Kant, independence from our natural instincts and inclinations. Kant, as we saw earlier, understands dignity as involving an elevation of one thing over something else. The dignity of humanity consists in its capacity to set aside completely the influence of inclinations in one's decision making. We can now refine this idea by saying that members of humanity are free to set aside any particular inclination in setting and choosing ends. That is, members of humanity can deliberate on their inclinations and then set ends according to their deliberation. The capacity to set ends freely, then, must involve the capacity to curb one's

¹⁵⁶ *Ibid.* For a fuller discussion of these original predispositions and their relation to each other, see Allen Wood, *Kant's Ethical Thought* (pp. 118-120).

¹⁵⁷ G 4:425-431.

¹⁵⁸ G 4:439.

¹⁵⁹ MM 6:391.

¹⁶⁰ G 4:437.

inclinations. Non-rational animals (i.e. creatures with a predisposition to animality, but not to humanity) are capable of intentional activity; they may even have the capacity to make choices and express preferences. In the *Collins Lectures*, Kant says, “All animals have the capacity to use their powers according to choice. Yet this choice is not free, but necessitated by incentives and *stimuli*.”¹⁶¹ If the choice one makes is determined solely by “incentives and *stimuli*,” then one is not exercising humanity, but rather mere animality. Kant also discusses incentives in the *Groundwork*; there he describes incentives as a subjective state that motivates one to act. He writes, “The subjective ground of desire is an *incentive*.”¹⁶² Kant reiterates this definition of an incentive in the second *Critique*. An incentive is “the subjective determining ground of the will.”¹⁶³ A moral action cannot be motivated by incentives, understood as the subjective ground of desire or of the will.

An incentive is a kind of enticement to act on the basis of an interest in satisfying an inclination. A good will acts solely from duty and not on the basis of incentives. In the *Groundwork*, Kant distinguishes between an incentive, which is grounded in one’s particular inclinations and a motive, which is not based on narrow self-interested inclinations, but which hold for all rational beings: “[T]he objective ground of volition is a *motive*; hence the distinction between subjective ends, which rest on incentives, and objective ends, which depend on motives, which hold for every rational being.”¹⁶⁴ What separates humanity from animality, then, is its capacity to set ends on the basis of

¹⁶¹ Collins Lectures 27:344.

¹⁶² G 4:427.

¹⁶³ CPrR 5:72.

¹⁶⁴ G 4:427.

something other than merely subjectively determined incentives (i.e. the capacity to set ends independently of what instinct and inclination may recommend).

Humanity, then, consists in the capacity to set and choose ends for oneself, independent of the dictates of inclination.¹⁶⁵ It seems natural, then, to suppose that humanity has dignity insofar as it is elevated above mere nature (i.e. inclinations) in its capacity to choose ends freely.¹⁶⁶ In the *Groundwork*, though, Kant is a bit more specific in the particular reason why humanity has dignity. There, Kant claims that humanity has dignity, “insofar as it is capable of morality.”¹⁶⁷ The phrase “capable of morality” merits some consideration, especially since Kant oftentimes describes the dignity of humanity specifically with regard to humanity’s relation to the moral law. Is humanity, or rational nature, necessary and sufficient for being capable of morality, or is there another faculty or property that moral action requires? It seems that Kant is committed to the view that the capacity for morality, in created rational beings, certainly involves humanity, *but* that it also requires what he variously calls a “predisposition to personality” or a “predisposition to the good.”¹⁶⁸

¹⁶⁵ Christine Korsgaard, in *Creating the Kingdom of Ends*, writes, “Kant takes the characteristic feature of humanity, or rational nature, to be the capacity for setting an end (110). Allen Wood, in *Kantian Ethics*, also describes humanity as “the capacity to set ends according to reason” (94).

¹⁶⁶ Paul Guyer emphasizes humanity’s freedom from nature in its capacity to set ends when he writes, “The only natural capacity unique to humankind is its potential to free itself from mere nature by exercise of its freedom” (*Kant on Freedom, Law, and Happiness* 169).

¹⁶⁷ G 4:435.

¹⁶⁸ Rel. 6:26, Rel. 6:27, Rel. 6:41, MM 6:464.

Kant sometimes describes humanity solely in terms of instrumental rationality, without reference to the moral law.¹⁶⁹ That is, possession of a rational nature is consistent with indifference to the moral law. Humanity therefore would be capable of existing without receptivity to morality. Revisiting Kant's definition of humanity in *The Metaphysics of Morals*, he writes, "The capacity to set oneself an end—any end whatsoever—is what characterizes humanity (as distinguished from animality)."¹⁷⁰ Humanity is characterized by the capacity to set "any end whatsoever," presumably both morally permissible and morally impermissible ends. In the *Religion*, Kant carefully distinguishes between the predisposition to humanity and to personality. The predisposition to humanity, as we saw before, consists in possession of a rational nature. The predisposition to personality, by contrast, concerns a human being as both "a rational and at the same time as *responsible* being."¹⁷¹ He also writes, "The predisposition to personality is the susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice*."¹⁷² He goes on to describe the predisposition to personality as "susceptibility to simple respect for the moral law within us..."¹⁷³ Personality, not humanity, is responsive to the moral law.

In a footnote in the *Religion*, Kant explains that personality cannot simply be included in the concept of humanity:

¹⁶⁹ For a similar discussion on the distinction between humanity and personality, see Allen Wood *Kant's Ethical Thought*.

¹⁷⁰ MM 6:392.

¹⁷¹ Rel. 6:26.

¹⁷² Rel. 6:27.

¹⁷³ Rel. 6:27.

We cannot consider this predisposition as already included in the concept of the preceding one, but must necessarily treat it as a special predisposition... The most rational being of this world might still need certain incentives, coming to him from the objects of inclination, to determine his power of choice... Were this law not given to us from within, no amount of subtle reasoning on our part would produce it or win our power of choice from determination by all other incentives (of our freedom) and thereby also of the accountability of all our actions.¹⁷⁴

Without susceptibility to the moral law, it is not clear how a person could engage in morally worthy action. Humanity, without personality, is capable of being exercised in the service of incentives other than the moral law: “The *second* [predisposition to humanity] is rooted in a reason which is indeed practical, but only as subservient to other incentives...”¹⁷⁵ Morality, though, presents itself to us as something that is necessary and universally valid for all rational beings. Through our predisposition to personality, we experience the moral law as the necessary determining ground of our will. That is, we take on an interest in the moral law because of our disposition to personality. Kant writes, that the moral law “announces itself to be itself an incentive, and, indeed, the highest incentive.”¹⁷⁶

Without our taking a rational interest in the moral law, we (as human beings, who must act on the basis of interest) would not be capable of moral action. Toward the end of the *Groundwork*, Kant claims that the “idea of a pure world of understanding as a whole of all intelligences, to which we ourselves belong as rational beings” is “useful

¹⁷⁴ Rel. 6:26n.

¹⁷⁵ Rel. 6:28.

¹⁷⁶ Rel. 6:26n.

and permitted for producing in us a lively interest in the moral law...”¹⁷⁷ In the second *Critique*, Kant claims that pure respect for the moral law “produces an *interest* in compliance with the law which we call *moral* interest...”¹⁷⁸ Our capacity for moral interest resides, it seems, in our predisposition to personality, not in our predisposition to humanity. Since human dignity, according to Kant, seems to be grounded in our capacity for morality, humanity alone, it would seem, is an insufficient ground for the basis of human dignity. Humanity, as Kant claims in the *Religion*, is “subservient to other incentives.”¹⁷⁹ Human dignity must then be based on possession of humanity *and* of personality. Without personality, our end-setting capacity can only be put in the service of satisfying inclinations and not in the service of morality.

Why, then, does Kant repeatedly claim that humanity has dignity? If dignity consists in capacity for morality and personality is required for that capacity, why does Kant identify humanity as possessing dignity, absolute worth, and as existing as an end in itself?¹⁸⁰ The answer, it seems, has to be that there is a degree of flexibility in the way Kant deploys terms like ‘humanity’ and ‘personality.’ While there is, to be sure, an authentic distinction between the two in the *Religion*, perhaps, the distinction is not as stark in other places. In other words, the distinction between humanity and personality may not be as clear as it appears.

¹⁷⁷ G 4:462.

¹⁷⁸ CPrR 5:80.

¹⁷⁹ Rel. 6:28.

¹⁸⁰ Kant describes humanity in these terms in the *Groundwork* (see 4:428-429).

There are at least two ways in which humanity and personality can coincide. First, in the case of human beings, humanity and personality are both ever-present in our nature and in that sense can never be separated. In the *Religion*, Kant claims that the predispositions to humanity and to personality are “original,” which, as we saw, means that they are included in the very concept of human nature itself: “They [predispositions] are *original* if they belong with necessity to the possibility of this being, but *contingent* if the being in question is possible in itself also without them.”¹⁸¹ In the *Religion*, Kant writes, “The human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it). The law rather imposes itself irresistibly, because of his moral predisposition; and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e. he would be morally good.”¹⁸² The moral law always inspires in human beings an attitude of respect (insofar as human beings are incapable of repudiating or renouncing it). Whenever a human being is going about deliberating about which ends to pursue, the moral law will always figure into one’s deliberations as something that should be accounted for and incorporated into one’s maxim.

There is another, perhaps stronger, sense in which one could attempt to draw a connection between humanity and personality. One could attempt to show that humanity and personality are related in such a way that the one leads to the other. That is, personality would be derivable from the humanity of human beings. Though I will

¹⁸¹ Rel. 6:28.

¹⁸² Rel 6:36.

ultimately argue that such an attempt would fail, there are nevertheless two interesting ways in which humanity and personality relate to each other, without actually forging a bona fide connection. In other words, while the distinction between humanity and personality is unbridgeable, a careful look at them brings them closer together; the distinction is perhaps not as stark as Kant seems to make in the *Religion*.

Humanity, as we saw earlier, consists in the capacity to choose among alternatives in setting ends. Kant writes, “The predispositions to humanity can be brought under the general title of a self-love which is physical yet involves comparison (for which reason is required)...”¹⁸³ The capacity to make comparisons, between the happiness of oneself and that of others, as well as between the efficacy of means in achieving one’s happiness, involves use of one’s reason without reference to morality and its separate requirements.

The exercise of one’s humanity always involves making value judgments about the particular ends that one sets. Christine Korsgaard makes this observation when she writes, “the distinctive feature of humanity, *as such*, is simply the capacity to take a rational interest in something: to decide, under the influence of reason, that something is desirable, that it is worthy of pursuit or realization, that it is to be deemed important or valuable, not because it contributes to survival or instinctual satisfaction, but as an end—for its own sake.”¹⁸⁴ Instinct is the sole basis on which animals act. So, while animals may have preferences, their actions are determined entirely by instinct. Human beings, insofar as they have reason, are not determined by instinct in the setting and choosing of

¹⁸³ Rel. 6:27.

¹⁸⁴ Christine Korsgaard. *Creating the Kingdom of Ends* (p. 114).

ends. In *Conjectural beginning of human history*, Kant illustrates this point when he writes, “Instinct, that *voice of God* which all animals obey must have guided the novice.” He goes on to write, “As long as the inexperienced human being obeyed this call of nature, he did well for himself. Yet *reason* soon began to stir and sought through comparison of that which gratified with that which was represented to him by another sense than the one to which instinct was bound...”¹⁸⁵ Reason leads to reflection and comparison between different possible ends. At this early stage in human history, Kant does not suppose that the moral law presents itself clearly as an alternative to instincts, but, crucially, reason does make it possible to *begin* to make value judgments and comparisons between possible ends.

Humanity possesses reason, but it is used solely in the service of self-love. Reason may suggest particular means and ways of achieving the satisfaction of one’s inclinations; it may even produce a hierarchy of inclinations and deem some more worthy of satisfaction than others. Humanity, without personality, lacks a moral incentive. The moral law does not present itself as required. The most emphatic inclination we have is our desire for freedom (understood as not being under the control of another). Kant’s discussion of this particular inclination, and his description of some psychological states that accompany it, begin to shift the exercise of our humanity as solely concerned with our own freedom towards a recognition that the freedom of others may need to be taken into account.

¹⁸⁵ *Conjectural Beginning* 8:111.

A Natural Connection Between Humanity and Personality?

In *Anthropology from a pragmatic point of view* (hereafter *Anthropology*) Kant claims that human beings have a passion for freedom. The discussion on our passion for freedom takes place in a section entitled “On the inclination to freedom as a passion.”¹⁸⁶ Kant had earlier defined passion as follows: “Inclination that can be conquered only with difficulty or not at all by the subject’s reason is *passion*.”¹⁸⁷ Additionally, Kant defines inclination as “[h]abitual sensible desire.”¹⁸⁸ Kant, then, seems committed to the claim that human beings have a natural and habitual sensible desire for freedom that is difficult, or perhaps even impossible, to eradicate. Kant writes, “For the natural human being this [passion for freedom] is the most violent inclination of all...”¹⁸⁹ Our inclination for freedom is not simply one inclination among many. We desire freedom with great a vigor that is unequalled by our desires for all other things.¹⁹⁰

Kant, in this section, defines happiness largely in negative terms. In other words, freedom is defined in terms of *not* existing in a particular state, under a particular set of circumstances. Freedom, here, seems to consist in not being in a position where your happiness depends on the choices that someone else makes. Such a situation, even if one

¹⁸⁶ Anth. 7:268-7:270.

¹⁸⁷ Anth. 7:251.

¹⁸⁸ *Ibid.*

¹⁸⁹ Anth. 7:268.

¹⁹⁰ Jean-Jacques Rousseau, in his *Second Discourse on Inequality*, discusses the vigor with which free people (and animals) will defend their liberty when he writes, “But when I see free peoples sacrificing pleasures, tranquility, wealth, power, and life itself for the preservation of this sole good which is regarded so disdainfully by those who have lost it; when I see animals born free and abhorring captivity break their heads against the bars of their prison; when I see multitudes of utterly naked savages scorn European pleasures and brave hunger, fire, sword, and death, simply to preserve their independence, I sense it is inappropriate for slaves to reason about liberty” (*The Basic Political Writings* 72-3).

is under the thumb of a benevolent person, cannot possibly be a happy state: “Whoever is able to be happy only according to *another* person’s choice (no matter how benevolent this other person may be) rightly feels that he is unhappy.”¹⁹¹ Kant argues that the un-free person rightly judges him or herself as unhappy because living under the control of another is not a sound guarantee for securing one’s happiness: “For what guarantee has he that this powerful fellow human being’s judgment about his well-being will agree with his own?”¹⁹² Even a relatively benevolent person is not an especially reliable provider of one’s happiness and is likely to produce cases in which a person’s views about his or her happiness differs from the more powerful person. Furthermore, a person who stands to lose his or her freedom to another has good anthropological reasons for suspecting the supposed benevolence of the would-be master, namely the fact that people have a natural inclination to dominate others in the service of their happiness (and often with indifference to the happiness of others).¹⁹³

Interestingly, Kant does not stop at the claim that we have a violent inclination for freedom; he additionally claims that, in the interest of acquiring or protecting our freedom, we make a demand on others not to impede on our freedom. Kant writes that our passion for freedom places the human being “in a condition where he cannot avoid making reciprocal claims on others.”¹⁹⁴ One could imagine humans as being possessive and jealous of their freedom without making a claim that others are bound to respect their

¹⁹¹ Anth. 7:268.

¹⁹² *Ibid.*

¹⁹³ See Kant’s discussion of what he calls “The mania for domination” in *Anthropology from a pragmatic point of view* (7:273).

¹⁹⁴ Anth. 7:268.

freedom. The inclination to freedom, in such a case, would express itself solely in the readiness to defend one's freedom against anyone who might impede it. Kant describes this readiness as a "state of continuous warfare, by which he intends to keep others as far away from him as possible..."¹⁹⁵ In such a scenario, those who defend their freedom make no demand to others that they are bound to respect their freedom. Rather, they simply make it evident that they will not surrender their freedom without a robust and violent defense of it. Kant, though, does not claim that the human defense of freedom does not solely consist in violent defense of it, rather people do indeed claim that others have an obligation to honor their freedom. Kant, as we saw, calls demands of this sort "reciprocal claims on others."¹⁹⁶

We experience coercion or lack of outer freedom as an injustice. While we may not have a clear understanding of why it is unjust, or the authentic rational ground for our complaint, we nevertheless have an inchoate notion of moral outrage. Kant writes, "Even the child who has just wrenched itself from the mother's womb seems to enter the world with loud cries, unlike all other animals, simply because it regards the inability to make use of its limbs as *constraint*, and thus it immediately announces its claim to freedom (a representation that no other animal has)."¹⁹⁷ In a note to this passage, Kant develops this idea further, "[T]he fact that his feeling of uncomfortableness is not due to bodily pain but to an obscure idea (or a representation analogous to it) of freedom and its hindrance, *injustice*, is disclosed a few months later after the birth by the *tears* which accompany his

¹⁹⁵ Anth. 7:268.

¹⁹⁶ *Ibid.*

¹⁹⁷ *Ibid.*

screaming; they indicate a kind of exasperation when he strives to approach certain objects or in general merely strives to change his position and feels himself hindered in it.”¹⁹⁸ Human beings are distinct from other animals, in part, because the frustration of a wholly natural desire (freedom of movement) is tinged with an inchoate sense of beings wrongly constrained. In our relation to other people, also, we come to expect them to honor our outer freedom through our making “reciprocal claims” on them.

Kant makes a similar point in a note he wrote in his copy of *Observations*. There, he claims that we have a natural abhorrence of servitude, which he describes in terms that are similar to the deprivation of outer freedom. He writes:

The human being has his own inclinations, and by means of his capacity of choice has a clue from nature to conduct his actions in accordance with these. Nothing can be more appalling than that the action of one human stand under the will of another. Hence no abhorrence can be more natural than that which a person has against servitude. On this account a child cries and becomes bitter if it has to do what another wants without one having made an effort to make that pleasing to him. And it wishes only to become a man quickly and to operate in accordance with its own will.¹⁹⁹

Here, might seem reasonable to suppose that Kant envisioned the particular human inclination to freedom as providing a foundation for the moral law. Nature, through inclinations, has provided human beings with a “clue” about how to conduct their actions by means of our “capacity of choice.” If, in the *Religion*, Kant describes humanity as operating solely with ends that are set by one’s inclinations (and by one’s capacity to

¹⁹⁸ *Ibid.*

¹⁹⁹ Notes and Fragments pp. 10-11.

choose one end among many), then perhaps the strongest and most basic inclination (desire for freedom) can serve as a bridge between humanity and personality. Since we cannot help but make demands on others to respect our freedom. We protect our freedom, on this reading, by making a reciprocal demand that if the others will not impede my freedom, I will not impede theirs. The relationship would not be an analytic one, but would only establish itself with the discovery of a particular inclination (through anthropological and psychological inquiry) that is so highly valued by human beings, that they naturally take on the requirements of morality (namely, protecting and promoting the freedom of all, not just of oneself) in order to preserve and protect its satisfaction.

However, natural it may seem to make this connection, our strong natural inclination for freedom cannot serve as the basis of a connection between humanity and personality, for basic Kantian reasons: a merely natural inclination cannot serve as the foundation for morality. Even though the result of this movement from humanity to personality (via a consideration of the human desire for freedom) does end up in basing one's choices on the freedom of all, the fundamental starting point on which the whole edifice of morality rests is a psychological fact about human beings. In the beginning of the *Groundwork*, Kant describes his project as an attempt to “work out for once a pure moral philosophy, completely cleansed of everything that may be only empirical and that belongs to anthropology...”²⁰⁰ While empirical and anthropological claims may have a role at some level in moral reasoning, they cannot serve any function whatsoever in establishing why we possess personality in addition to humanity. So, while our natural

²⁰⁰ G 4:389.

inclination to freedom does seem to resonate with our ultimate moral vocation, and perhaps even the existence of such a strong desire for freedom provides a clue that we have a moral being that is not possessed by other animals and that thereby raises us above nature, it cannot provide a substantial link between our desire for our own freedom and our duty to preserve and promote the freedom of others.

A Metaphysical Connection?

If a connection between humanity and personality cannot be based upon an empirical and psychological insight, then perhaps a connection is located in rational nature itself. A connection based on rational nature would constitute a link not based on natural inclination, but rather on the basis of rationality itself and its necessary relationship to morality.

Humanity's value resides in its capacity for freedom. Protecting and promoting humanity involves protecting and promoting the exercise of human freedom (and the freedom of all rational beings in general). The second formulation of the categorical imperative claims: "*So 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.*"²⁰¹ One way in which a person can reduce others to mere means to an end is to tell them a false promise in order to cause a person to behave in a way that will satisfy an inclination one has. In such a case one person (the liar) would be putting another's humanity (their capacity to choose how to act) in the service only of the liar's ends, without regard for the

²⁰¹ G 4:429

ends of the other person. Such a scenario is rather similar to the case that Kant laid out in the *Anthropology* concerning the happiness of someone depending on the choice of another. Kant claims that being in such a state of dependency is (rightly) abhorrent to us. Kant's justification for describing a false promise in terms of treating others as mere means goes as follows: "For, he whom I want to use for my purposes by such a promise cannot possibly agree to my way of behaving toward him, and so himself contain the end of this action."²⁰² This may be read *prima facie* as providing a strong empirical hypothesis, according to which no human being actually would agree to be lied to and to be put in the service of another's ends on the basis of disinformation. If Kant were making an empirical claim here, his mode of argument would be similar to the discussion in the *Anthropology* about our natural inclination to freedom and our disinclination to servitude. It seems, though, that Kant is not making a naturalistic claim here at all, but is rather making a claim about to what it is possible for a rational being to consent. Another rational being, for Kant, "cannot possibly agree" to be used as a mere means, not because of an inclination that they may have, but rather because of rational nature itself.

A bit earlier in the *Groundwork*, Kant makes a claim that also potentially can be read as a naturalistic claim: "The human being necessarily represents his own existence in this way; so far it is thus a *subjective* principle of human actions."²⁰³ As Kant notes, this sort of description of human beings is based merely on inclinations (which are subjective, as opposed to objective) and is therefore not adding anything new to the discussion he had in the *Anthropology* about the human desire for freedom. He continues,

²⁰² G 4:429

²⁰³ G 4:428-9

though: “But every rational being also represents his existence in this way consequent on just the same rational ground that also holds for me; thus it is at the same time an *objective* principle from which, as a supreme practical ground, it must be possible to derive all laws of the will.”²⁰⁴ The basis for viewing oneself as an end in itself is no longer merely natural (i.e. a desire or inclination) but rather “rational” and “objective.” The wrongness of exploiting others for one’s own private ends now is justified through its basis in a rational ground, common to all rational beings.

My self-image as an end in itself and my rational demand to be treated as such is based on the same ground as other rational beings similar self-images and rational demands. The rational ground seems to be the freedom that all rational beings possess, and the dignity that each has on the basis of that freedom. The human being as an end in itself provides an *objective* end for guiding one’s moral choices, one for which all rational beings must, insofar as they are rational, strive. In describing human beings as ends in themselves, Kant writes, “These, therefore, not merely subjective ends, the existence of which as an effect of our action has a worth *for us*, but rather *objective ends*, that is, beings the existence of which is in itself an end, and indeed one such that no other end, to which they would serve *merely* as means, can be put in place...”²⁰⁵ An objective end has a status (a worth) that is exalted above any other merely subjective end. Kant goes on to write, “...since without it [an objective end] nothing of *absolute worth* would be found anywhere; but if all worth were conditional and therefore contingent, then no

²⁰⁴ G 4:429

²⁰⁵ G 4:428

supreme practical principle for reason could be found anywhere.”²⁰⁶ Only rational beings, viewed as ends in themselves, possess absolute worth: “But suppose there were something the *existence of which in itself* has an absolute worth, something which as an *end in itself* could be a ground of determinate laws; then in it, and in it alone, would lie the ground of a possible categorical imperative, that is, of a practical law.”²⁰⁷ Kant follows this discussion immediately with his claim that “the human being and in general every rational being *exists* as an end in itself...”²⁰⁸ The rational basis for viewing oneself as an end in itself, then, is not simply a strong inclination to outer freedom, but rather a rational recognition of the absolute worth one has on the basis of one’s humanity. If one is to remain rational in his or her demand that humanity (in his or her own person) be respected, then one must also recognize, preserve, and promote the humanity in others. Personality, then, is possession of a rational nature, coupled with this recognition.

Kant’s discussion, though, seems simply to assume that human beings view themselves as ends in themselves on the basis of a rational ground and that such a rational ground has the power to motivate us in our capacity of choice. That is, he seems to suppose, without argument, that we are receptive to the demands of morality and aware of the absolute worth of our freedom. Kant does indeed seem to hold, ultimately, that we are aware of our moral obligations and are affected by the moral law (i.e. possess a

²⁰⁶ *Ibid.*

²⁰⁷ *Ibid.*

²⁰⁸ *Ibid.*

predisposition to personality) on the basis of an indemonstrable and immediate “fact of pure reason.”²⁰⁹ Kant, in the *Critique of Practical Reason*, writes:

Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical, although it would be analytic if freedom of the will were presupposed; but for this, as a positive concept, an intellectual intuition would be required, which certainly cannot be assumed here.²¹⁰

The categorical imperative applies to us, but we only know this through a fact of reason, not through a demonstration of its validity. We again are a breaking point between humanity and personality. We recognize our own legitimate claim to freedom on the basis of our humanity (i.e. on our capacity for freedom), but we lack an account, without assumption, of the rational interest we take in the moral law (and in freedom for all rational beings).

If one takes our freedom as a given (as a fact of reason), then ultimately, there are two basic principles that one can apply in setting and choosing ends: satisfying one’s desires or abiding by the moral law.²¹¹ In the *Groundwork*, Kant simply assumes that if inclinations are not going to provide the motivational grounds for the will, then the only thing left to do so would be the moral law. In introducing the universal law formulation of the categorical imperative at the end of Section I of the *Groundwork*, Kant writes,

²⁰⁹ CPrR 5:31

²¹⁰ *Ibid.*

²¹¹ Rel. 6:35-36

“Since I have deprived the will of every impulse that could arise for it from obeying some law, nothing is left but the conformity of actions as such with universal law...”²¹² In the *Religion*, Kant describes the moral situation of human beings as the struggle to bring two volitional principles into a proper ordering: that of self-love and that of morality.²¹³ He claims that the moral law and the “law of self-love...cannot stand on equal footing, but one must be subordinated to the other...”²¹⁴ For any end that one sets, the reason one sets it, ultimately, is either based on considerations of self-interest (i.e. out of self-love) or it is based on considerations of the moral law and one’s duty. In the latter case, choosing the end and pursuing it is either consistent with the moral law or it is actually required by the moral law. In either case, the moral law must be the basic ground in determining the content of one’s will. In Kant’s terms, it must be “incorporated” into one’s maxim.²¹⁵ If one sets an end or makes a choice to neglect the demands of the moral law and to violate it, then one has willingly debased one’s humanity by putting it to use in a way that is no different than the way non-rational animals make choices on the basis of inclinations. In the *Collins Lectures*, Kant says, “All animals have the capacity to use their powers according to choice. Yet this choice is not free, but necessitated by incentives and *stimuli*. Their actions contain *bruta necessitas* [animal necessity]. If all creatures had such a choice, tied to sensory drives, the world would have no value.” If one sets ends solely in accordance with one’s instincts, then one is forfeiting one’s humanity, and also one’s dignity. Instead of being determined by the moral law within oneself and by autonomy,

²¹² G 4:402

²¹³ Rel. 6:36.

²¹⁴ *Ibid.*

²¹⁵ Rel. 6:32.

one is allowing nature, in the form of instincts and inclinations, to push oneself around.

If one were truly to exercise one's humanity, then one would choose, set, and pursue ends that are consistent with, and determined by, the moral law and its requirements.

Furthermore, any and all failures to conform one's will to the moral law are expressions of one's mere animality and not one's personality. Humanity, when perfected, would be identical with expressing one's personality.

In the case of human beings, humanity could not really exist without personality. Without a rational interest for the moral law, then, humanity would lapse into mere animality. In *The Metaphysics of Morals*, Kant describes what he calls "certain moral endowments" that are subjective consequences of our receptivity to the moral law: moral feeling, conscience, love of human beings, and self-esteem. In his discussion of these four moral endowments, Kant seems to be operating under the assumption that all human beings have a predisposition to personality (and therefore that the exercise of one's humanity will always be accompanied by a concern for the moral law in one's deliberations). He writes, "To have these predispositions cannot be considered a duty; rather every human being has them, and it is by virtue of them that he can be put under an obligation.—Consciousness of them is not of empirical origin; it can, instead, only follow from consciousness of a moral law, as the effect this has on the mind."²¹⁶ The existence of these endowments in all people cannot serve as *proof* that the moral law applies to us, but they are in concert with the assumption that the moral law does indeed apply to us. They are the effects that would befall human beings who existed as moral beings.

²¹⁶ MM 6:399.

In discussing moral feeling, Kant writes, “No human being is entirely without moral feeling, for were he completely lacking in receptivity to it he would be morally dead; and if (to speak in medical terms) the moral vital force could no longer excite this feeling, then humanity would dissolve (by chemical laws, as it were) into mere animality and be mixed irretrievably with the mass of other natural beings.”²¹⁷ Of course since the predisposition to personality is “original,” it is not really possible for one to lose it (without also perhaps losing one’s capacity to set and choose ends freely). In other words, no person is completely lacking moral feeling. Kant writes, “[E]very human being (as a moral being) has it [moral feeling] in him originally.”²¹⁸ We can, and do, and an obligation to cultivate and perfect our moral feeling, which is accomplished by meditating on its source: “Obligation with regard to moral feeling can be only to *cultivate* it and to strengthen it through wonder at its inscrutable source.”²¹⁹ This “inscrutable source” seems to be the part of us that is receptive to moral incentives, and not mere natural ones: “This [cultivation and strengthening of one’s moral feeling] comes about by its being shown how it is set apart from any pathological stimulus and is induced most intensely in its purity by merely rational representation.”²²⁰

IV. Dignity and Freedom

Human dignity, then, is grounded in the capacity to set ends apart from instinct; put another way: dignity is grounded in the capacity to set ends *freely*. The dignity of

²¹⁷ MM 6:400.

²¹⁸ MM 6:399.

²¹⁹ *Ibid.*

²²⁰ MM 6:400.

humanity therefore consists in freedom. Kant expresses this point in a note when he writes, “The dignity of humanity in one’s own person is personality itself, that is, freedom; for one is only an end in oneself insofar as one is a being that can set ends oneself.”²²¹ One can elevate oneself above nature only by setting aside completely the influence of inclinations and instead acting on the basis of pure respect for the moral law. In a note, Kant writes, “Freedom has a dignity on account of its independence; it also has a high price, for through it we are able to become authors of the good in accordance with our own concepts, which we can extend and multiply far beyond the natural instincts of animals.”²²²

Human dignity, though, is grounded in the possibility of *fulfilling* one’s capacity for freedom.²²³ That is, a person’s dignity, in some sense, depends on actually rising above mere animality and one’s natural instincts. In another note on moral philosophy, Kant writes, “The dignity of human nature lies solely in freedom; through it we alone can become worthy of any good. But the dignity of a human being (worthiness) rests on the use of freedom, whereby he makes himself worthy of everything good.”²²⁴ Humans become worthy of happiness only through the proper use of one’s reason.

Not only does freedom make people worthy of happiness, freedom is that which makes value possible in the world. There can be no value in the world without freedom: “Freedom...is the capacity which confers unlimited usefulness on all the others. It is the

²²¹ *Notes on moral philosophy* 19:307.

²²² *Notes on moral philosophy* 19:294.

²²³ For a similar discussion, see Paul Guyer *Kant and the Experience of Freedom* (347).

²²⁴ *Notes on moral philosophy* 19:181.

highest degree of life.” If no creatures in the world were free, then “the world would have no value.” With the possibility of freedom, though, it also becomes possible for there to be value in the world: “But the inner worth of the world, the *summum bonum*, is freedom according to choice that is not necessitated to act. Freedom is thus the inner worth of the world.”²²⁵

Freedom, then, is a kind of precondition for anything else to carry value (as opposed to carrying a mere price). That is, one can never rightly forfeit one’s freedom, even if one would thereby gain immense satisfaction of one’s inclinations (such as gaining massive amounts of wealth and power, for instance, in order to achieve happiness). In the case of human beings, only a free human being can have dignity. As Kant writes in the note quoted above, “[T]he dignity of a human being (worthiness) rests on the use of freedom...”²²⁶ Kant, in a perhaps rather extreme formulation contained in the *Collins Lectures*, claims that life itself has no value aside from freedom (that is, using one’s freedom in a morally proper fashion): “It is better to sacrifice life than to forfeit morality. It is not necessary to live, but it is necessary that, so long as we live, we do so honourably...”²²⁷ Kant, then, seems committed to a generally perfectionist conception of dignity.²²⁸ Free rational beings have dignity insofar as they use their freedom properly, which means they do not allow inclinations (no matter how numerous and powerful) to dictate one’s ends. On this reading, dignity is not a widely possessed property among human beings. Kant, at various times, expresses considerable skepticism about their

²²⁵ Collins Lectures 27:344.

²²⁶ *Notes on moral philosophy* 19:181.

²²⁷ Collins Lectures 27:373.

²²⁸ Oliver Sensen also describes Kant’s views on human dignity in perfectionist terms.

being any authentic instances of virtue at all. So he is skeptical that human beings have ever truly overcome the inclinations of self-love and have acted with a good will. In the next section I will try to show that Kant nevertheless holds that dignity is in fact widely possessed by human beings through their status as ends-in-themselves.

V. Dignity, Respect, and Contempt

Kant claims that beings with dignity are also ends-in-themselves and are owed respect by all other rational beings. In *The Metaphysics of Morals*, Kant writes, “[A] human being considered as a *person*, that is, as the subject of a morally practical reason, is exalted above any price; for as a person he is not to be valued merely as a means to the ends of others or even to his own ends, but as an end in itself, that is, he possesses a *dignity* (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world.”²²⁹ People are entitled to respect from others on account of their dignity. Kant writes, “The *respect* that I have for others or that another can require from me is therefore a recognition of a *dignity* in other human beings, that is, of a worth that has no price, no equivalent for which the object could be exchanged.”²³⁰ Human dignity, then, presents all rational beings with a requirement to treat all fellow rational

²²⁹ MM 6:434-435. This passage seems to echo the second formulation of the categorical imperative, which states, “So 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” (G 4:429). One might question why Kant decides to use “person” in the passage from *The Metaphysics of Morals* and “humanity” in the second formulation. On my reading of personality and humanity, though, this does not need to be troubling or surprising. On my reading personality and humanity are closely related and, ultimately, similar nearly to the point of being identical with each other (when humanity is perfected). For another discussion on why Kant uses the term “humanity” in the second formulation instead of “person,” see Allen Wood’s *Kant’s Ethical Thought*. Wood retains a more rigid distinction between humanity and personality than I maintain in this paper.

²³⁰ MM 6:462.

beings with respect. In *the Metaphysics of Morals*, Kant writes, “Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other. Humanity itself is a dignity; for a human being cannot be used merely as a means by any human being (either by others or even by himself) but must always be used at the same time as an end. It is just in this that his dignity (personality) consists, by which he raises himself above all other beings in the world that are not human beings...”²³¹ When one possesses dignity, one is also to be valued not merely as a means to some end, but rather also as an end in itself. Human beings are to be treated with respect on account of their status as dignified beings.

If we have a perfectionist understanding of dignity, though, it seems that we are to respect only those who possess true freedom (i.e. moral worth or virtue). This would be a troubling conclusion. It seems to come into tension with Kant’s more universal-sounding declarations of the duty to respect all persons, even vicious persons who clearly are not free within the perfectionist framework. In the *Collins Lectures*, for instance, Kant says, “Humanity, however, is worthy of respect, and even though somebody may be a bad man, the humanity in his person is entitled to respect.”²³² In fact, Kant clearly states that we have a *duty* to respect others: “To be *contemptuous* of others, that is, to deny them the respect owed to human beings in general, is in every case contrary to duty; for they are human beings.”²³³

²³¹ MM 6:462.

²³² Collins Lectures 27:373.

²³³ MM 6:463.

The basic way to reconcile a perfectionist interpretation of Kantian dignity with his inclusive remarks about the duty to respect all human beings, including vicious ones, is in Kant's claim that a person can command respect *as a moral being* (that is, as a rational creature that can set ends freely and takes on an interest in the moral law) while at the same time, through his actions, make himself unworthy of respect. Kant makes this claim when he writes, "I cannot deny all respect to even a vicious man as a human being; I cannot withdraw at least the respect that belongs to him in his quality as a human being, even though by his deeds he makes himself unworthy of it."²³⁴

Human dignity obtains when a person is capable of setting ends freely and when they take an interest in the moral law. These two dispositions, while only fully realized in a good will, can manifest themselves imperfectly in other ways. The predisposition to personality and an interest in the moral law is present in those who have engaged in immoral behavior. In the *Groundwork*, Kant writes, "There is no one—not even the most hardened scoundrel, if only he is otherwise accustomed to use reason—who, when one sets before him examples of honesty of purpose, of steadfastness in following good maxims, of sympathy and general benevolence (even combined with great sacrifices of advantage and comfort), does not wish that he might also be so disposed."²³⁵ The "hardened scoundrel" retains his dignity because he still has his humanity and personality, even though he, through his inability to prioritize properly the relationship between self-love and the moral law, is wasting or squandering his capacity for morality. Wrongdoers nevertheless retain *something* (predisposition to personality and humanity)

²³⁴ MM 6:463.

²³⁵ G 4:454.

that elevates them over the rest of nature and over mere animality. In the *Religion*, Kant writes, “[T]he command that we *ought* to become better human beings still resounds unabated in our souls; consequently, we must also be capable of it, even if what we can do is itself insufficient and by virtue of it, we only make ourselves receptive to a higher assistance inscrutable to us.—Surely we must presuppose in all this that there is still a germ of goodness left in its entire purity, a germ that cannot be extirpated or corrupted.”²³⁶ This is an interesting passage for a number of reasons, but for our purposes we need only focus on Kant’s insistence that there is a “germ of goodness” that makes us receptive to the command to become better human beings. We are responsive to the demands of the moral law, even if we fail to abide by them, and even if we need “higher assistance” to actually become good.

This elevation, however imperfect, manifests itself in ways other than pious wishes for virtue. Another way in which the moral predisposition in human beings expresses itself is through the care and sophistication that all people are able to bring to bear regarding questions of morality. In the second *Critique*, Kant makes the following observation:

Now of all arguments there are none that more excite the participation of persons who are otherwise soon bored with subtle reasoning and that bring a certain liveliness into the company than arguments about the *moral worth* of this or that action by which the character of some person is to be made out. Those for whom anything subtle and refined in theoretical questions is dry and irksome soon join in when it is a question of how to make out the moral import of a good or evil action that has been related, and to an extent one does

²³⁶ Rel. 6:45.

not otherwise expect of them on any object of speculation they are precise, refined, and subtle in thinking out everything that could lessen or even just make the suspect of purpose and consequently the degree of virtue in it.²³⁷

The moral law presents itself to all human beings (excepting those who may have severe mental ailments or are otherwise badly incapacitated) as necessary and universally valid.

As a result, human beings know what it would mean to act in a morally laudable way, and thus are capable of bringing a high degree of subtlety to moral discussions.

Furthermore, since all human beings take an interest in the moral law, those who may otherwise become bored by abstract, theoretical discussions tend to maintain a high degree of interest in discussions that center on moral worth. These sorts of propensities can be taken as expressive of human dignity and our separation from (or elevation over) non-rational or amoral animals.

Kant also affirms human dignity in the constant possibility for moral improvement. So long as one remains capable of moral improvement, one retains one's humanity and predisposition to personality (i.e. one's dignity) and so long as one remains a human being, one remains capable of moral improvement. We saw earlier how Kant claims that even "hardened scoundrels" maintain an interest in the moral law and a wish for moral virtue. Kant also claims that we are to treat others with dignity and respect because failure to do so would deny them even the possibility of moral improvement. In *The Metaphysics of Morals*, Kant writes that the censure of vice "must never break out into complete contempt and denial of any moral worth to a vicious human being; for on

²³⁷ CPrR 5:153.

this supposition he could never be improved, and this is not consistent with the idea of a *human being*, who as such (as a moral being) can never lose entirely his predisposition to the good.”²³⁸ So long as one has a predisposition to the good, one is thereby elevated over mere animality (which lacks responsiveness to the moral law), and therefore one has dignity.

If human dignity consists in freedom (and I have claimed that ultimately it does), then Kant even demonstrates considerable patience and tolerance for a certain degree of *misuse* of one’s freedom, at least in the early stages of learning to exercise freedom. For instance, as one begins to set ends free of coercion, both from internal coercion on the basis of inclinations and from external coercion on the basis of social and political forces (which in turn largely operate on the basis of manipulating a person’s inclinations). He writes,

I will admit that I am not comfortable with this way of speaking, which even clever men are wont to use: “A certain people (intent on establishing civil freedom) is not ripe for freedom”; “The bondsmen of a landed proprietor are not yet ripe for freedom”; and so too, “People are in general not ripe for freedom of belief.” For on this assumption freedom will never come, since we cannot *ripen* to it if we are not already established in it (we must be free in order to be able to make use of our powers purposively in freedom). To be sure, the first attempts will be crude, and in general also bound to greater hardships and dangers than when still under the command but also the care of others; yet we do not ripen to freedom otherwise than through our *own* attempts (and we must be free to be allowed to make them)...[T]o make it a principle that those who are once subjected to them [those in power] are essentially not suited for freedom, and that one is justified in keeping them from it for all time, this is an

²³⁸ MM 6:464.

intrusion into the prerogatives of Divinity itself, which created human beings for freedom.”²³⁹

Freedom, as we saw earlier, is for Kant the precondition for anything having true value at all in the world. If one could gain immense happiness by forfeiting one’s freedom, one should nevertheless renounce happiness and retain freedom. Similarly, if one can live a life of ease under the control of another or one can assert one’s freedom and lead a life of tumult and strife, one should select the latter, even though it may lead, at first, to great danger and merely “crude” attempts in exercising one’s freedom. For Kant, only a free life can be a fully dignified life.

Kant’s view on human dignity brings together a kind of moral perfectionism with an inclusive universalism. His conception of dignity does not lend itself to complacent indifference to the moral law, as though since one has dignity one can, in a sense, simply rest on that and not struggle for moral improvement. His views on dignity also, though, are inclusive and egalitarian. Everyone has dignity and everyone is owed mutual respect on account of it. Dignity is, perhaps, similar to a state of natural talent in something. A person can use their talent in a proper way, or one can misuse it and waste that initial talent one has.²⁴⁰ In the case of wasting talent, it would make no difference, in terms of the end result (remaining in an uncultivated state), whether or not one ever had the talent

²³⁹ Rel. 6:188n.

²⁴⁰ In his paper “Kant’s Conception of Human Dignity,” Oliver Sensen claims that there are “two stages” of human “elevation or dignity.” And only through the proper use of one’s freedom does one fully realize his or her dignity: “The initial elevation of humanity is only realized if one makes proper use of one’s moral capacity” (325). On my reading, it would be a mistake to say that the improper use of one’s freedom annuls one’s dignity. Human dignity lies in the capacity for moral elevation, not in its realization.

in the first place. Similarly, while a human being has the capacity for freedom, if one fails to prioritize properly the law of self-love and the moral law, then one is squandering that freedom. Dignity consists in all human beings possessing the means to elevate themselves above nature, but it can also be wasted if one descends back into mere animality.

CHAPTER 3

Respect in Kant's Moral Philosophy

Introduction

This chapter examines the role and meaning of respect in Kant's moral philosophy. I contend that Kant uses the term "respect" in a variety of contexts that, on the surface, seem inconsistent with each other. I argue that there are several senses of respect that Kant employs in his moral philosophy and that properly distinguishing between the various senses of respect helps dispel the apparent inconsistency. Specifically, I distinguish between three senses of respect that Kant uses: First, respect as the affective response human beings have toward the moral law. This is what Kant refers to as "respect for the moral law." Second, respect as the attitude we are required to hold towards fellow human beings. This sense of respect is what Kant describes in his discussion of the duties of respect. Third, respect as an expression of praise for another person's commitment to the moral law (this sense of respect I label "merit-based" respect).

In Section I, I discuss Kant's notion of respect for the moral law. In Section II, I examine the duties of respect that Kant lays out in the "Doctrine of Virtue" in *The Metaphysics of Morals*. In Section III, I discuss the respect that Kant thinks we reserve for agents whose virtue we admire. I conclude by maintaining that all the senses are ultimately grounded in the first sense (respect for the moral law). I also argue that each

sense of respect gives rise to moral humility in us by providing us with firsthand experience of our own tenuous and complex relation to the moral law.

I. Respect for the Moral Law

*Respect for the Moral Law in the **Groundwork to the Metaphysics of Morals***

In the *Groundwork of the Metaphysics of Morals* Kant claims that only acts committed from duty (and not merely in conformity with duty) are worthy of moral approbation. Actions from duty require a person to nullify the influence of particular sensible desires and inclinations in her decision-making. Inclinations, even benevolent ones, are to have no role in determining the will behind a morally relevant action. “Now, an action from duty is to set aside entirely the influence of inclination and with it every object of the will; hence there is left nothing that could determine it except objectively the *law* and subjectively *pure respect* for this practical law, and so the maxim of complying with such a law even if it infringes upon all my inclinations.”²⁴¹ Benevolent inclinations, animated by an authentic delight in helping others, are “amiable” and “honorable” and for that reason “deserves praise and encouragement but not esteem.”²⁴² Esteem is reserved for actions that are done from duty alone.

Since Kant is discussing the actions of human beings, though, he does not deny that *some sort* of feeling must always be present in a human being, even when she is acting from duty. There will also be a “subjective” side to our choices. Kant holds that subjectively the will must be determined solely by “pure respect” for the (moral) law.

²⁴¹ G 4:400-1

²⁴² G 4:398

While Kant admits that pure respect for the moral law is a feeling, it is of a singularly peculiar sort. He writes, “But though respect is a feeling, it is not *received* by means of influence; it is, instead, a feeling *self-wrought* by means of a rational concept...”²⁴³ The “rational concept” that is the ground of the “self-wrought” feeling is the moral law itself, the objective determination of the good will. The feeling, then, is occasioned not by anything empirical, but rather by a rational concept. In particular, a rational concept we impose upon ourselves.²⁴⁴ The moral law nevertheless has an *effect* on human subjectivity, namely respect. Kant continues:

What I cognize immediately as a law for me I cognize with respect, which signifies merely consciousness of the *subordination* of my will to a law without the mediation of other influences on my sense. Immediate determination of the will by means of the law and consciousness of this is called *respect*, so that this is regarded as the *effect* of the law on the subject, and not as the *cause* of the law.²⁴⁵

The moral law’s necessity²⁴⁶ with regard to the will brings about in subjects a distinctive *feeling* or *affect*, which Kant labels as “respect for law”²⁴⁷ or “pure respect for this practical law.”²⁴⁸ The character of the feeling is of a negative sort in this section of the *Groundwork*. Kant argues that the moral law serves as a check or limitation on our inclinations. He writes that the moral law “does not serve my inclination but outweighs it

²⁴³ G 4:401n

²⁴⁴ For a similar discussion see Sally Sedgwick (2008) *Kant’s Groundwork of the Metaphysics of Morals: An Introduction* pp. 73-76.

²⁴⁵ *Ibid.*

²⁴⁶ In the same note, Kant writes, “The object of respect is therefore simply the law, and indeed the law that we impose upon ourselves and yet as necessary in itself” (G 4:401n)

²⁴⁷ G 4:400

²⁴⁸ G 4:401

or at least excludes it altogether from calculations in making a choice...²⁴⁹ A genuine moral action, an action from duty, requires the subject “to put aside entirely the influence of inclination.”²⁵⁰ We recognize the necessity of *subordinating* our inclinations to the requirements of morality.

Besides the claim that the subjective determination of the good will is respect for the moral law and the observation that doing so makes us aware of the need to ground our choices on something other than our inclinations, Kant does not further elaborate on the feeling of respect in the *Groundwork*. For a fuller account of respect for the moral law, we need to examine Kant’s discussion of it in his *Critique of Practical Reason*.

Respect for the Moral Law in the Critique of Practical Reason

In the chapter entitled “On the incentives of pure practical reason” from the *Critique of Practical Reason*, Kant discusses at some length the effect that the moral law has on rational, non-divine (or created) beings. As we saw from our discussion of the *Groundwork*, Kant calls the effect of the moral law on a rational created creature “respect.” Kant defines “incentive” as “the subjective determining ground of the will of a being whose reason does not by its nature necessarily conform with the objective law.”²⁵¹ He then claims that the moral law itself is the only possible candidate for the incentive of the good will.²⁵² In other words, the moral law must always be that *for the sake of which*

²⁴⁹ G 4:400

²⁵⁰ *Ibid.*

²⁵¹ CPrR 5:72

²⁵² Allen Wood (1999) notes that in the *Groundwork* there is a distinction between what Kant calls “motives” and “incentives.” A motive is an “objective ground, which is an end given by reason alone and valid for all rational beings,” while an incentive is a “subjective ground, which

an action is performed, if the action is moral. So in the case of the moral law, the subjective and objective determining ground of the will are one and the same. Kant writes, “[T]he incentive of the human will (and of the will of any created rational being) can never be anything other than the moral law; and thus that the objectively determining ground must always and quite alone be also the subjectively sufficient determining ground of action if this is not merely to fulfill the *letter* of the law without containing its spirit.”²⁵³ If an action that conforms to our duty is subjectively determined by anything other than the moral law, then the action possesses “legality” but not “morality.”²⁵⁴

The objective character²⁵⁵ of the law is such that simply by being aware of it we are, at least in part, *motivated* to act for the sake of it. We become motivated to act for the sake of it insofar as reason recognizes the necessity of the moral law in the determination of our will (i.e. its universal validity).²⁵⁶ Part of this claim is the idea that there is something about human beings that makes us receptive to the moral law. In *Religion Within the Boundaries of Mere Reason*, Kant writes, “The human being (even the worst)

is an end based on empirical desires” (*Kant’s Ethical Thought* 111). In Kant this distinction is perhaps clearest in the Groundwork when Kant writes, “The subjective ground of desire is an incentive; the other ground of volition is a motive; hence the distinction between subjective ends, which rests on incentives, and objective ends, which depend on motives, which hold for every rational being” (G 4:427-8). Since in the Critique of Practical Reason Kant describes the moral law as the only possible incentive of a good will, it follows that it must be a unique sort of incentive, not merely based on empirical desires, but rather on something like a rational interest in the moral law.

²⁵³ *Ibid.*

²⁵⁴ CPrR 5:71. This view is reiterated in *The Metaphysics of Morals*: “The mere conformity or nonconformity of an action with law, irrespective of the incentive to it, is called its legality (lawfulness); but that conformity in which the idea of duty arising from the law is also the incentive to action is called its morality” (MM 6:219).

²⁵⁵ The objective character of the moral law is constituted by the principles it contains (see CPrR 5:66).

²⁵⁶ Note about “The analytic of practical reason” in CPrR 5:57-71.

does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it). The law rather imposes itself irresistibly, because of his moral predisposition; and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e. he would be morally good.”²⁵⁷ So, a (human) will that has successfully restricted the influence of inclination from its decision-making, by virtue of the human being’s “moral predisposition,” will be morally good.

Instead of further examining the claim that the moral law must be the incentive of a good will, Kant refocuses the discussion on the *effect* that the moral law has on subjects insofar as it is an incentive: “What we shall have to show a priori is, therefore, not the ground from which the moral law in itself supplies an incentive but rather what it effects (or, to put it better, must effect) in the mind insofar as it is an incentive.”²⁵⁸ Kant, in this section seems to believe he has already demonstrated that the moral law must in fact supply itself as incentive. Now, the concern is to show what we can know about the effect taking up the moral law as incentive involves for human beings.

A will is determined by the moral law only when inclination has been subordinated to the requirements of the moral law. That is, if I find in myself an inclination that, if pursued, would lead me to violate the moral law, then (insofar as my will is determined by the moral law) I will not act upon the inclination. Inclinations can figure into our decision-making, but whatever inclinations we act upon must be consistent with the dictates of the

²⁵⁷ Rel 6:36

²⁵⁸ CPrR 5:72

moral law.²⁵⁹ Kant locates evil in the human being not in the mere presence of inclinations (even vicious ones) or in the efficaciousness of inclinations in our decision-making, but rather in cases where a person has subordinated the moral law to them, not the other way around:

[T]he human being (even the best) is evil only because he reverses the moral order of his incentives in incorporating them into his maxims. He indeed incorporates the moral law into those maxims, together with the law of self-love; since, however, he realizes that the two cannot stand on equal footing, but one must be subordinated to the other as its supreme condition, he makes the incentives of self-love and their inclinations the condition of compliance with the moral law—whereas it is this latter that, as *the supreme condition* of the satisfaction of the former, should have been incorporated into the universal maxim of the power of choice as the sole incentive.²⁶⁰

Since the moral law appears even to the least moral person as something that is required to adhere to in one's choices, the law will never completely disappear from our deliberation of choice. That is, all human beings feel *compelled or required* to take up the moral law and incorporate it into their choices. Every human being has the same *kind* of relation to the moral law and finds him or herself in the same kind of predicament with regard to the power of choice. Namely, we have the moral law on the one hand and our inclinations on

²⁵⁹ Kant does not hold that inclinations are to be suppressed to such a degree as to play no role in our decision-making. Nor does he think moral worth consists in eradicating our inclinations. Rather he argues that they must play a secondary role in our decision-making. One must, as he says in his Lectures on Ethics, “bring them [inclinations] under a rule” (Collins Lectures, 27:345). See also his comment in Religion Within the Bounds of Mere Reason, when, while contrasting his views from those of the Stoics, he writes, “Considered in themselves natural inclinations are good, i.e. not reprehensible, and the want to extirpate them would not only be futile but harmful and blameworthy as well; we must rather only curb them, so that they will not wear each other out but will instead be harmonized into a whole called happiness” (6:58).

²⁶⁰ Rel 6:36

the other; and at times they will come into conflict with each other. The effect of the moral law on people, then, will be uniform in kind (though the particular dynamics between inclinations and duty will vary from person to person).

The notion that respect for the moral law operates in the same way in all human beings is another way of saying that no human being is perfectly *innocent*. Innocence can be understood as a state in which one's inclinations naturally lead one to do good and to act in accordance with the duties that arise from the moral law.²⁶¹ A human being's relation to the moral law is always complicated by the presence of inclinations that, if acted upon, would cause one to violate the moral law. No human being, therefore, is perfectly innocent. There are always tendencies that may gnaw at our moral resolve. Respect is the result of the *clashing* of those tendencies and our duty. Kant writes:

The moral level on which a human being (and, as far as we can see, every rational creature as well) stands is respect for the moral law. The disposition incumbent upon him to have in observing it is to do so from duty, not from voluntary liking nor even from an endeavor he undertakes unbidden, gladly and of his own accord; and his proper moral condition, in which he can always be, is *virtue*, that is moral disposition *in conflict*, and not *holiness* in the supposed *possession* of a complete *purity* of dispositions of the will.²⁶²

Moral purity is not a possible moral disposition for human beings, but rather moral discipline. Moral purity, like innocence, would involve no countervailing tendencies in our inclinations to our predisposition to the moral law. Moral purity would mean that all

²⁶¹ In this definition of innocence, I am following Allen Wood (1999) who writes that innocence is “a spontaneous and wholly pre-rational impulse to do good—that is, to do what proper rational reflection would tell us should be done” (*Kant's Ethical Thought*, p. 45).

²⁶² CPrR 5:84.

our inclinations would lead us to do our duty *gladly*. Moral purity is not possible for creatures like us.²⁶³ Kant even suggests that holding oneself up as an instance of moral purity and innocence is *contrary* to morality because it suggests a degree of arrogance in one's own righteousness without an honest recognition of having any inclinations that, if acted upon, would lead one to immoral action.²⁶⁴ Kant suggests that our reflection of one's duty does not lead us to "rove among *fancied* moral perfections" but rather to "set limits of humility (i.e. *self-knowledge*) to self conceit as well as to self-love, both of which are ready to mistake their boundaries."²⁶⁵

The feeling of respect that Kant describes in the *Critique of Practical Reason* must be widely (universally, in fact) shared feeling among human beings. The moral law appears to human beings as containing requirements that, at times, will clash with our inclinations. In other words, the moral law seems to human beings to set a limit on which inclinations one can satisfy. A good will, then, is one in which the relation between inclinations and duty is properly ordered; that is, a good will is determined by the moral law. Kant writes, "What is essential in every determination of the will by the moral law is that, as a free will—and so not only without the cooperation of sensible impulses but even

²⁶³ Kant puts this factual assertion rather bluntly when he writes, "[I]f a rational creature could ever reach the stage of thoroughly liking to fulfill all moral laws, this would mean that there would not be in him even the possibility of a desire that would provoke him to deviate from them... But no creature can ever reach this stage of moral disposition" (CPrR 5:84).

²⁶⁴ Allen Wood (1999) also makes this point when he writes, "He [Kant] thinks that when we who have lost innocence pretend to find in our own spontaneous feeling something either more reliable or more estimable than the self-constraint of reason, this enthusiasm involves a dangerous self-deception closely allied to moral corruption. What the sentimentalist's reflective endorsement of our spontaneous sentiments expresses is not the spontaneous love of goodness he pretends to find in himself but instead the mendacious arrogance of representing himself as too good to need rational constraint in order to do his duty" (*Kant's Ethical Thought*, p. 45).

²⁶⁵ CPrR 5:86 (emphasis added).

with rejection of all them and with infringement upon all inclinations insofar as they could be opposed to that law—it is determined solely by the law.”²⁶⁶

As in the *Groundwork*, Kant begins his discussion effect of the moral law on a subjects in distinctly negative terms: “So far, then, the effect of the moral law as incentive is only negative...”²⁶⁷ In particular, the moral law gives rise to a feeling that, in part, consists in a negative or painful *feeling*: “For, all inclination and every sensible impulse is based on feeling, and the negative effect on feeling (by the infringement upon the inclinations that takes place) is itself feeling.”²⁶⁸ An inclination often presents itself as a desire that we naturally seek to satisfy. The satisfaction of our desires is pleasurable, while the dissatisfaction or frustration of our desires results in a painful or unpleasant feeling. Therefore, whenever we have to infringe upon an inclination that would be immoral to act upon, we will experience some measure of unpleasantness or pain.²⁶⁹

Hence we can see a priori that the moral law, as the determining ground of the will, must by thwarting all our inclinations produce a feeling that can be called pain; and here we have the first and perhaps only case in which we can determine a priori from concepts the relation of a cognition (here the cognition of a pure practical reason) to the feeling of pleasure or displeasure.²⁷⁰

²⁶⁶ CPrR 5:72.

²⁶⁷ *Ibid.*

²⁶⁸ CPrR 5:72-3.

²⁶⁹ Paul Guyer provides a similar interpretation. He writes, “The mechanism of this pain might be readily understood as compatible with obvious laws of psychology. The decision not to satisfy one or more of one’s inclinations, where such satisfaction would require actions excluded by duty, causes frustration of those inclinations, which in turn naturally leads to pain” (*Kant and the Experience of Freedom* 357).

²⁷⁰ CPrR 7:73.

The negative or painful effect of the moral law on subjects with inclinations that do not always conform to the moral law has two components or aspects. Kant writes, “All the inclinations together (which can be brought into a tolerable system and the satisfaction of which is then called one’s own happiness) constitute regard for oneself (*solipsismus*).”²⁷¹ This system of self-regard divides into two sorts: 1) What Kant calls a “*love for oneself*, a predominant *benevolence* toward oneself (*Philautia*); and 2) a “*satisfaction with oneself* (*Arrogantia*).”²⁷² Kant also refers to benevolence toward oneself as “self-love” and satisfaction with oneself as “self-conceit.”²⁷³

Self-love is the ordinary tendency in human beings to satisfy their desires; self-conceit, by contrast, consists of inclinations to have a positive evaluation of one’s own worth or value. The moral law’s effect on each set of inclinations differs. Self-love is only limited by the moral law. Kant writes, “Pure practical reason merely *infringes upon* self-love, inasmuch as it only restricts it...”²⁷⁴ Self-love is consistent with having a good will (required even); when one’s inclinations are brought under the moral law, and satisfied only when they are consistent with it, self-love becomes *rational*.²⁷⁵

Self-conceit is a different case. Whereas self-love can be preserved in a will determined by the moral law, self-conceit must, necessarily, be completely done away with. Kant writes, “But it [pure practical reason] *strikes down* self-conceit altogether, since all claims to esteem for oneself that precede accord with the moral law are null and quite

²⁷¹ CPrR 5:73.

²⁷² *Ibid.*

²⁷³ *Ibid.*

²⁷⁴ *Ibid.*

²⁷⁵ *Ibid.*

unwarranted because certainty of a disposition in accord with this law is the first condition of any worth of a person...and any presumption prior to this is false and opposed to the law.”²⁷⁶ Self-conceit is a “claim to esteem for oneself” for reasons other than one’s will being determined by the moral law.²⁷⁷ Kant suggests that self-conceit can be understood in similar terms as any other sort of inclination; he holds that “the propensity to self-esteem, so long as it rests only on sensibility, belongs with the inclinations which the moral law infringes upon. So the moral law strikes down self-conceit.”²⁷⁸ It seems plausible to think of inclinations toward self-conceit as resting on sensibility if you extend to sensibility psychological pleasure. Simply put, it feels good to regard oneself as worthy of esteem. Inclinations toward self-conceit, though, will be completely struck down by the moral law when their basis is merely on the desire to feel good about oneself, regardless of the content of one’s will.

The moral law does not only generate a negative or painful affect in rational beings. Thus far, Kant’s discussion is little more than a reiteration and refinement of his treatment of the moral law’s effect on human beings from the *Groundwork*. Here, though, Kant adds a crucial aspect to the notion of respect for the moral law. Kant also holds that there is a *positive* affect that the moral law brings about in rational created beings. In fact, the positive aspect of the effect of the moral law on human beings is what makes the feeling

²⁷⁶ *Ibid.*

²⁷⁷ Kant cannot mean that no person has any claim to esteem or value unless he or she has in fact a good will. This would seem to rub against the second formulation of the categorical imperative where humanity is taken to be an unconditional value, regardless of the agent’s morals. Rather it seems that Kant here means “esteem” in the sense of achieving moral excellence. This seems like the third sense of respect, merit-based respect, that I discuss in this chapter.

²⁷⁸ CPrR 5:73.

respect and not mere frustration or dissatisfaction.²⁷⁹ An action that contains morality (and not mere legality) is a *free* one—the action of a will that is objectively determined by the moral law. Kant writes:

But since this law is still something in itself positive—namely the form of an intellectual causality, that is, of freedom—it is at the same time an object of *respect* inasmuch as, in opposition to its subjective antagonist, namely the inclinations in us, it *weakens* self-conceit; and inasmuch as it even *strikes down* self-conceit, that is, humiliates it, it is an object of the greatest *respect* and so too the ground of a positive feeling that is not of empirical origin and is cognized a priori.²⁸⁰

The very fact that a human being is capable of casting aside the influence of inclination brings about the positive feeling of freedom, or our capacity to act on the basis of a universal law.²⁸¹ The result is a complex affective state consisting of both positive and negative aspects.²⁸² The pleasure we derive from the moral law, coupled with the

²⁷⁹ It is interesting to note that in this passage Kant does not use the word “respect” (*Achtung*) until he begins to discuss the positive aspect of the effect of the moral law. “Respect,” though, designates the unified affective state consisting of both positive and negative aspects. They can be analyzed separately from each other, but in actual instances of it both aspects must always be present. The two aspects are mutually reinforcing; one grounds the other. The very source of the positive feeling of freedom consists precisely in our capacity to infringe upon inclination, while the source of the negative feeling derives from our rational nature and its adherence to the moral law.

²⁸⁰ CPrR 5:73.

²⁸¹ Furthermore, successfully limiting the influence of inclination and conducting ourselves morally is the only authentic basis for our inclinations of self-conceit to be satisfied rationally. We rightly esteem ourselves when we act morally. This is not to say that we act morally for the sake of satisfying our need for self-esteem, but rather as a consequence of moral action, we gain an authentic basis for self-esteem (see CPrR 5:73).

²⁸² The complexity I describe here is different from the complexity described by several other commentators (see Richard McCarthy “Motivation and Moral Choice in Kant’s Theory of Rational Agency,” *Kant-Studien*, 1994, vol. 85 and Philip Stratton-Lake *Kant, Duty and Moral Worth* pp. 35-39), who analyze respect as consisting of affective and cognitive aspects. My

dissatisfaction we derive from its requirement that we curb the influence of inclination in our decision making constitutes the complex affective state of respect for the moral law. Respect for the moral law, then, is, at root, a feeling of freedom (understood, though, as also conferring limitations on our inclinations) and an encounter with the authentic basis for self-esteem (namely, our capacity to adhere to the moral law).

As *submission* to a law, that is, as a command (indicating constraint for the sensibly affected subject), it therefore contains in it no pleasure but instead, so far, displeasure in the action. On the other hand, however, since this constraint is exercised only by the lawgiving of his *own* reason, it also contains something *elevating*, and the subjective effect on feeling, inasmuch as pure practical reason is the sole cause of it, can thus be called *self-approbation* with reference to pure practical, inasmuch as he cognized himself as determined solely by the law and without any interest, and now becomes conscious of an altogether different interest subjectively produced by the law, which is purely practical and *free*; and his taking this interest in dutiful action is not advised by any inclination; instead reason through the practical law absolutely commands it and also actually produces it, because of which it has quite a special name, that of respect.²⁸³

The negative aspect of the affective state never completely drops out. So long as we are human beings with empirical inclinations we will always be tugged in directions other than towards the moral law.²⁸⁴ Nevertheless respect for the moral law does have a

interpretation is consistent with the claim that respect has both affective and cognitive aspects. I simply am focusing on the affective aspect of it.

²⁸³ CPrR 5:80-81.

²⁸⁴ Failure to recognize this point would conflate virtue and holiness. Virtue consists in moral discipline, while holiness does not require moral discipline since there are no countervailing inclinations that would break our commitment to the moral law. In the *Metaphysics of Morals*

powerful positive aspect that, if not overwhelming the negative aspect, as least serves as a powerful counterweight to it.²⁸⁵ Through adherence to the moral law (and the feeling of respect that subjectively determines the will) we feel elevated and are rationally justified in engaging in self-approration. This self-approration is not based on morally irrelevant features about who we are, but rather on our will being in accordance with the command that the moral law lays out for it.

II. Respect Owed to All Human Beings: Duties of Respect

The Basic Duty of Respect: Respect as Avoiding Contempt

In the *Groundwork* as well as in “On the incentives of pure practical reason” from the *Critique of Practical Reason*, Kant focuses on the effect of the moral law on a single subject. That is, his discussion centers around the effect that the moral law has on the affective state of individual human beings. The subjective feeling of respect is exclusively directed to the moral law. A human being’s will, insofar as the agent adheres to the moral law, is subjectively determined by “pure respect” or simply “respect” for the moral law.²⁸⁶ The notion of respect, though, is not exhausted by Kant’s discussion of respect as the affective state one has when the moral law is the determining ground of one’s will. In *The Metaphysics of Morals*, Kant claims that human beings are morally required to respect

Kant writes, “Virtue signifies a moral strength of the will. But this does not exhaust the concept; for such strength could also belong to a holy (superhuman) being in whom no hindering impulses would impede the law of its will and who would thus gladly do everything in conformity with the law” (6:405). A holy will is a “superhuman” will, not a human one.

²⁸⁵ Paul Guyer (1993) also emphasizes the positive aspect of respect: “The feeling of respect is thus a complex but ultimately pleasurable state of feeling produced by our decision to adhere to the moral law, grounded, like other feelings of pleasure, in recognition of the possibility of the realization of our own objectives but reflecting in its very complexity the fact that not all of our objectives are conjointly satisfiable” (*Kant and the Experience of Freedom* 360).

²⁸⁶ See e.g. G 4:400, G 4:403, CPrR 5:74, CPrR 5:75, CPrR 5:76.

fellow human beings. He writes, “Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other.”²⁸⁷ The claim to respect is “legitimate” because of the dignity of humanity: “Humanity itself is a dignity; for a human being cannot be used merely as a means by human being (either by others or even by himself) but must always be used at the same time as an end.”²⁸⁸ The respect we owe to human beings is a second sense of respect that plays a crucial role in Kant’s moral philosophy.

Respect for other human beings involves recognition of their *dignity*. On account of that dignity, one cannot harbor *contempt* for human beings,²⁸⁹ nor can one *behave* in ways that *express* contempt for them.²⁹⁰ Something has dignity when it has something that is “of a worth that has no price, no equivalent for which the object evaluated (*aestimii*)

²⁸⁷ MM 6:462. Kant makes a similar factual claim in the Collins Lectures when he writes, “By nature we have two drives, whereby we demand to be respected by others, and to be loved” (Collins Lectures 27:407).

²⁸⁸ *Ibid.* Also, this echoes the second formulation of the categorical imperative formulated in the Groundwork: “The practical imperative will therefore be the following: So 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” (G 4:429).

²⁸⁹ Kant discusses contempt in his Lectures on Ethics as well, though the discussion there is along a more naturalistic line. He begins with a factual claim (quoted above) about how human beings by nature have two drives: 1) the demand for respect, and 2) the desire to be loved. He then claims that respect is “directed to our inner worth... One is respected because one has inner worth” (Collins Lectures 27:407). Kant then claims that our natural demand for respect is only legitimate when one has respect for other people: “If we wish to be respected, we must also have respect for other people, and for mankind in general” (Collins Lectures 27:407). Although the form of this imperative seems to be hypothetical, Kant will separately argue that not only is our demand for respect a natural drive, but it is also required by morality. He calls it the “love of honor,” and claims it is the “highest duty of humanity to oneself” (Vigilantius Lectures §110 27:664).

²⁹⁰ My discussion of humanity in this chapter is skeletal. I discuss Kant’s account of humanity in greater detail in chapter one of this dissertation. For my purposes here it is sufficient simply to assert that part of recognizing the humanity of fellow human beings consists in treating others with respect, rather than contempt. Another way to put this idea is as follows: human beings have dignity and are therefore owed respect by their fellow human beings.

could be exchanged.”²⁹¹ Given human beings have something which is of absolute value (i.e. something that cannot (legitimately) be traded away for any price), judging any human being as having no value is to treat someone with contempt.²⁹² Contempt is never consistent with the moral law, even contempt for people who have committed gross misdeeds. Kant writes, “To be *contemptuous* of others (*contemnere*), that is, to deny the respect owed to human beings in general, is in every case contrary to duty; for they are human beings.”²⁹³

This general sense of respect (that owed to all human beings) is owed to all: those who have a pattern of good moral action as well as those who have exhibited a pattern of immoral action. Kant apparently sensed a kind of resistance such a claim might be met with. Why, one may ask, do I owe respect to, say, scoundrels—to those who have apparently forsaken altogether any commitment to the moral law? It may seem almost grotesque to suggest that they are owed respect. Kant gives indirect voice to that concern when he writes, “At times one cannot, it is true, help inwardly *looking down* on some in comparison with others (*despicatui habere*); but the outward manifestation of this is, nevertheless, an offense.”²⁹⁴ This passage problematizes Kant’s account a bit. One could interpret it as asserting that, while in some cases it is justified to withdraw respect, one must never show *outward manifestations* of contempt. One may harbor contempt, but never act in a way that *expresses* it. Though here Kant seems to be suggesting such an interpretation, it cannot be what Kant’s position actually is. Respect, in the sense that

²⁹¹ MM 6:462.

²⁹² *Ibid.*

²⁹³ MM 6:463.

²⁹⁴ *Ibid.*

animates the duties of respect, must be held toward all people. Kant goes on to write, “Nonetheless I cannot deny all respect to even a vicious man as a human being; I cannot withdraw at least the respect that belongs to him in his quality as a human being, even though by his deeds he makes himself unworthy of it.”²⁹⁵ The sometimes inevitable “looking down on comparison with others” is not the same thing as withdrawing respect for others insofar as they are human beings.²⁹⁶

This is not to deny that we can never condemn immoral acts. One can condemn the immoral action while also respecting the person who committed it. What *is* disqualified by our duty to respect others are actions that would undermine the guilty person’s dignity. Kant argues that, as a result of a person’s dignity, there are limits for what can be done in the punishment of criminals; namely, anything that would degrade their humanity: “So there can be no disgraceful punishments that dishonor humanity itself (such as quartering a man, having him torn up by dogs, cutting off his nose and ears). Not only are such punishments more painful than loss of possessions and life to one who loves honor (who claims respect of others, as everyone must); they also make a spectator blush with shame at belonging to a species that can be treated that way.”²⁹⁷ The criterion Kant is laying out here is that no punishment is permitted that would degrade humanity (even in the person of vicious scoundrels).

²⁹⁵ *Ibid.*

²⁹⁶ This “looking down” does cause us to withhold respect in another sense, though. This kind of respect is discussed below under the label “merit-based” respect.

²⁹⁷ MM 6:463.

The same sort of reasoning applies to censure. We are right to censure vices in others, but the censure cannot turn into vicious contempt, which degrades the offender's humanity. Kant writes, "The same thing applies to the censure of vice, which must never break out into complete contempt and denial of any moral worth to a vicious human being; for on this supposition he could never be improved, and this is not consistent with the idea of a *human being*, who as such (as a moral being) can never lose entirely his predisposition to the good."²⁹⁸ Viewing another human being with contempt ultimately amounts to a denial of his or her *potential* for moral action. We harbor contempt for another when we deny the fact that they are capable of feeling respect for the moral law at all. If one has no predisposition to the good, then one is really not a member of humanity. A person, when viewed with contempt, is reduced to a mere thing.²⁹⁹ Respect, in this sense, in part consists in refusal to look down with contempt on others, as well as unyielding recognition of the other's humanity and the other's predisposition to the good.

In addition to the more general requirement of avoiding contempt, Kant discusses particular vices that violate duties of respect owed to other human beings. They are: 1) arrogance, 2) defamation, and 3) ridicule. The discussion of these vices brings out another

²⁹⁸ MM 4:463-4. Note here the reason why we should not have contempt for wrongdoers is because human beings never completely lose their predisposition to the good. This is not exactly the same ground for respect Kant laid out earlier when he grounded it in humanity.

²⁹⁹ Viewing an immoral person with contempt, then, would amount to seeing them in roughly the same terms as a natural disaster or a vicious, dangerous animal. In which case, one might say, harboring contempt for someone becomes irrational. We don't have contempt for natural disasters or to animals because we do not expect them to know any better. The ground for moral disapprobation must be some sense of respect (i.e. a recognition of their predisposition to the good). Kant seems to be making this point when he writes, "It is only through the noble predisposition to the good in us, which makes the human being worthy of respect, that one can find one who acts contrary to it contemptible (the human being himself, but not the humanity in him)" (MM 6:441).

element of the duties of respect that was implicit in the preceding discussion of respect and contempt in general. Namely, the duties of respect forbid us from thinking of ourselves as superior to others. Each vice with regard to duties of respect arise from inclinations we have for being deemed superior in comparison with others. Respect, by contrast, involves an *egalitarian* element that forbids such vicious competitiveness. The common feature, then, that each of the three vices share is their source in inclinations that drive us to make ourselves superior to others by undermining the bases for respect of others.

Arrogance

The first vice that violates duties of respect for other human beings (hereafter referred to as “vices of disrespect”) is arrogance. Kant offers a definition of arrogance when he writes: “*Arrogance (superbia* and, as this word expresses it, the inclination to be always *on top*) is a kind of *ambition (ambitio)* in which we demand that others think little of themselves in comparison with us.”³⁰⁰ An arrogant person declares it his or her *right* to view others with contempt: “Arrogance is, as it were, a solicitation on the part of one seeking honor for followers, whom he thinks he is entitled to treat with contempt.”³⁰¹ Arrogance, therefore, declares a right to hold others in contempt, which, obviously, is expressly forbidden by the very idea of a duty of respect. Kant provides a similar definition of arrogance in the *Collins Lectures*: “Arrogance...is founded on the demand a person makes, that others should hold themselves in low esteem, in comparison with

³⁰⁰ MM 6:465.

³⁰¹ *Ibid.*

himself; this, therefore, is associated with a demeaning of the other's personhood...³⁰²

Arrogance violates the egalitarianism that underlies the duties of respect in its demand for others to be reduced to mere "followers" who are not rightfully (in the mind of the arrogant person) to seek their own aims and purposes in life, but only to defer obsequiously to their superior.

Kant is clearly committed to the view that humans have a legitimate claim to respect from other humans, so arrogance cannot simply be vicious in virtue of its demand for respect. Rather, it consists in the inclination to demand esteem from others only at the expense of their own self-esteem. Arrogance, therefore, "differs from **pride proper** (*animus elatus*), which is *love of honor*, that is, a concern to yield nothing of one's dignity in comparison with others...for arrogance demands from others a respect it denies them."³⁰³ Pride proper is nothing more than the demand for equal recognition of our shared humanity. Arrogance, unlike pride proper, involves a violation of the basic egalitarianism that is at the heart of Kant's duties of respect. Everyone is worthy of the same kind of respect, in this sense, and in the same degree.

Defamation

³⁰² Collins Lectures 27:708.

³⁰³ MM 6:465.

The second vice that Kant describes as a violation of our duty of respect towards other human beings is defamation. Kant defines defamation as “the immediate inclination, with no particular aim in view, to bring into the open something prejudicial to respect for others.”³⁰⁴ This might strike one as an odd formulation because the ground for respect for humanity is absolute and cannot be forfeited by any action whatsoever. So long as one is a human being, nothing can undermine the legitimacy of one’s demand for respect from others. So, nothing that a defamer could possibly bring out about another person could possibly serve as grounds for others to withdraw respect.

Yet, Kant does seem to suggest here that defamation can in fact weaken the basis for respect for others, “[F]or every scandal given weakens that respect, on which the impulse to the morally good rests, and so far as possible makes people skeptical about it.”³⁰⁵ Given Kant’s account of humanity as possessing absolute value as an end in itself and his thesis that, on the basis of that absolute value, all human beings are owed respect, it seems peculiar that he would justify his prohibition on defamation by writing, “The intentional *spreading (propalatio)* of something that detracts from another’s honor—even if it is not a matter of public justice, and even if what is said is true—diminishes respect for humanity as such, so as finally to cast a shadow of worthlessness over our race itself, making misanthropy (shying away from human beings) or contempt the prevalent cast of mind, or to dull one’s moral feeling by repeatedly exposing one to the sight of such things and accustoming one to it.”³⁰⁶ The line of reasoning here seems to be based on an

³⁰⁴ MM 6:466.

³⁰⁵ *Ibid.*

³⁰⁶ *Ibid.*

empirical claim (premise 2 below): 1) Defamation is the intentional spreading of news, gossip, or rumor about the various misdeeds of others; 2) constantly being bombarded with news, gossip, rumor (even if true) about the various misdeeds and immorality or others instills in us (human beings) an attitude of cynicism about and contempt toward the human race as a whole; 3) contempt for humanity is prohibited because humanity has absolute value; 4) a defamer knowingly instills contempt in human beings toward humanity as a whole; so therefore, 5) defamation diminishes respect for human beings as a whole and is therefore a violation of our duty to respect human beings.

On the surface, it might seem Kant is defining defamation in too broad a manner. On this understanding of defamation, for example, many instances of corporate or government whistle blowing, and the criminal investigations that may follow, would be prohibited. Certainly it is sometimes acceptable, if not required, for a person to “go public” with information about another person’s misdeed, even if it may result in a loss of honor for that person in the eyes of other people. Here it is necessary to refine our definition of defamation. Recall Kant’s definition: “By defamation (*obtrectatio*) or backbiting...I mean only the immediate inclination, with no particular aim in view, to bring into the open something prejudicial to respect for others.”³⁰⁷ There is a crucial caveat in this definition: Defamation is an “immediate inclination” that has “no particular aim in view.” There is a distinctly psychological component to defamation in which one person seeks to place others into disrepute and take pleasure in it. Kant goes on to describe the immorality involved in defamation in psychological terms: “It is, therefore, a duty of

³⁰⁷ *Ibid.*

virtue not to take malicious pleasure in exposing the faults of others so that one will be thought as good as, or at least not worse than, others...³⁰⁸ The immorality of defamation ultimately is based on a similar kind of inclination as arrogance—namely, the desire to undermine the basis of respect for others in order to feel better about oneself (or to have others feel better about oneself).

Defamation could be construed as an illicit way of easing one's own conscience (i.e. the awareness that one has inclinations that run against the demands of the moral law or, more strongly, the awareness that one has indeed acted on such inclinations at one time or another). Kant has a similar discussion in the *Critique of Practical Reason* in his discussion of respect: "So *little* is respect a feeling of *pleasure* that we give way to it only reluctantly with regard to a human being. We try to discover something that could lighten the burden of it for us, some fault in him to compensate us for the humiliation that comes upon us through such an example."³⁰⁹ This discussion of our reticence to give respect is particularly noteworthy for our purposes because it is couched in a larger discussion about our attitude toward people in whom we detect a more steadfast commitment to the moral law than we detect in ourselves. Kant claims that, instead of acknowledging our own moral shortcomings, we often seek to alleviate our conscience by undermining positive examples of moral excellence, thereby lessening the starkness of the comparison between their moral discipline and our own.

³⁰⁸ *Ibid.*

³⁰⁹ CPrR 5:77.

There is abundant empirical evidence for such attitudes and behavior. This sort of bad faith rationalization is something that we regularly witness and participate in. Often in times of conflict, each side self-righteously decries the misdeeds of the other side (real as they may be) as a way of drawing attention away from one's own misconduct. We often attend most peculiarly to the faults of others in order to bolster our own sense of moral self-esteem. In the *Vigilantius Lectures* we get a clear rendering of this vicious propensity in us: “[W]e actually hate his [the other person’s] very virtues, and feel an inner joy at viewing him with less respect; we notice the other’s failings more shrewdly than his merits, in order to put him down.”³¹⁰ These sorts of vicious inclinations lie at the heart of defaming actions.

The second part of the caveat—that defamation is an inclination that has “no particular aim in view”—provides a possible basis for legitimate instances of bringing to light the crimes or misdeeds of others. There might be some “particular aims” where morality would allow, and perhaps even require, the dissemination of information regarding the misdeeds of others. Whistle blowing, therefore, does have a proper place in Kant’s account here, even with his strictures on defamation. If it is done, for example, with the aim of protecting consumers from predatory commercial practices or the aim of bringing to light the suffering of people at the hands of some guilty party in order to end the suffering and to ensure reparations are paid. Allen Wood presents a similar interpretation of defamation when he writes, “It is wrong – a violation of a strict duty, and a proper object of blame – to gossip about others, to expose their faults to public censure,

³¹⁰ *Vigilantius Lectures* 27:678.

when this is done not for the purpose of guarding others against their misdeeds but simply in order to bring them (or even human nature in general) into disrepute” (emphasis added).³¹¹ Even though there are instances where public disclosure of another person’s misdeeds would be legitimate, the upshot of Kant’s discussion of defamation is that revealing the misdeeds of others is not *assumed* to be morally admissible. There needs to be a legitimate, “particular aim” that justifies such a disclosure. The target of Kant’s discussion is clearly not whistle blowers or criminal investigators, but rather those who revel in scandal in order to satisfy inclinations that would degrade the humanity of others.

Ridicule

The final particular vice of disrespect Kant discusses is ridicule. Kant defines ridicule as follows: “*Wanton faultfinding and mockery*, the propensity to expose others to laughter, to make their faults the immediate object of one’s amusement...”³¹² Ridicule is first defined as a pattern of behavior—faultfinding and mockery—that aims to make a person’s faults an object of amusement. Kant calls behavior like this a “kind of malice” that differs from “banter” which takes as its object the morally neutral “peculiarities” (as opposed to the faults) of others as the object of amusement.³¹³

Like arrogance and defamation, there is a vicious psychological satisfaction that some get from ridicule. So, like the other vices of disrespect, ridicule is the means to satisfy an inclination. Kant writes, “But holding up to ridicule a person’s real faults, or

³¹¹ Allen Wood, “Duties to Oneself, Duties of Respect to Others” from *The Blackwell Guide to Kant’s Ethics* (edited by Thomas E. Hill Jr.).

³¹² MM 6:467.

³¹³ *Ibid.*

supposed faults as if they were real, in order to deprive him of the respect he deserves, and the propensity to do this, a mania for *caustic* mockery (*spiritus causticus*), has something of fiendish joy in it; and this makes it an even more serious violation of one's duty of respect for other human beings."³¹⁴ The fact that ridiculers derive satisfaction from their vice compounds the immorality of ridicule. The satisfaction, like arrogance and defamation, consists in the demeaning of another's humanity in order to put oneself in a position of superiority.

The Gravity of Disrespect Towards Others

We should note that committing a vice of disrespect is a serious offense in Kant's ethics. Duties of respect are strict or narrow duties, meaning they are absolutely required. Kant writes, "Failure to fulfill mere duties of love is *lack of virtue (peccatum)*. But failure to fulfill the duty arising from the *respect* owed to every human being as such is a *vice (vitium)*. For no one is wronged if duties of love are neglected; but a failure in the duty of respect infringes upon one's lawful claim."³¹⁵

I have argued that the vices of disrespect have as their sources particular sorts of inclinations. Kant, in each of the three vices, discusses the "fiendishness," "malice" or "mania" that animate them. The desire for superiority over others is also present in

³¹⁴ *Ibid.*

³¹⁵ MM 6:464. Kant's formulation here seems to be a bit imprecise. If I completely fail to fulfill, for example, my duty of assistance to the poor, it does seem that individuals are being wronged by me. Namely, those who have a legitimate claim to assistance are being wronged. Kant's larger point, I think, still holds. If I violate a narrow duty of respect, I display malicious disrespect directly to an individual. Failure to assist the poor may be the result of a moral shortcoming, but it need not be a product of the insidious malice that disrespect has. If I withhold assistance to the poor out of disrespect or contempt, then, it seems, I would be doubly guilty of a violation of both a narrow duty and of inadequate adherence to a wide one.

arrogance, while a version of a similar inclination is there in defamation (by attending to the misdeeds of others to prop up one's own sense of moral excellence) and ridicule (by caustically joking about the shortcomings of others). The description of these vices, particularly the sorts of incentives that are behind them are quite similar to the sorts of extremely malignant vices Kant labels as "diabolical" in the *Religion*:

[O]nly in comparison with others does one judge oneself happy or unhappy...from this arises gradually an unjust desire to acquire superiority for oneself over others...in the face of all the anxious endeavor of others to attain a hateful superiority over us, to procure it for ourselves over them for the sake of security, as preventative measure...Hence the vices that are grafted upon this inclination [to be superior to others] can also be named vices of *culture*, and in their extreme degree of malignancy...e.g. in *envy*, *ingratitude*, *joy in others' misfortunes*...they are called *diabolical vices*.³¹⁶

Arrogance, defamation, and ridicule involve the satisfaction of inclinations that seek to place oneself in a position of superiority to others. Joy in others' misfortune is a diabolical vice; defaming another and taking joy in the other's guilt is a vice of disrespect.

The duties of respect, on this reading, have two components, an *inclination* toward disrespect and the *behavior* that satisfies the inclination. The inclinations behind the vices of disrespect are distinctive in their inherent immorality. That is, there can never be a legitimate satisfaction of the inclination to be arrogant, to defame, or to ridicule. Given this fact, the question arises whether 1) the *mere presence* of these inclinations constitutes a moral vice, or 2) an agent is only vicious when they seek to satisfy the inclinations

³¹⁶ Rel. 6:27.

behind the vices of disrespect. *Prima facie* it seems that (1) should be rejected because Kant seems to think that vicious inclinations (i.e. inclinations that, if acted upon, will lead us to deviate from our duty) are simply an inextricable part of human nature. It seems reasonable to lump together inclinations toward disrespect with all the other inclinations and conclude that, though obviously not a *good* thing to have, they do not in themselves constitute a moral vice.

There might be another line one might take, though, regarding inclinations to disrespect others. In the introductory remarks to the duties of respect in the *Metaphysics of Morals*, Kant asserts that we are forbidden from harboring contempt for others. This requirement seems to be a requirement on our *attitude toward* others; it sounds distinctively psychological and only accidentally related to particular modes of behavior. One could clearly harbor contempt for another person without overtly demonstrating the contempt. An inclination, then, to diminish the respect given to others already seems to come dangerously close to an *actual* instance of contempt for others.

How can one accommodate the plausibility of treating inclinations of disrespect like any other inclination with the prohibition of having an attitude of contempt (i.e. disrespect) for others? One way this might be accomplished is to provide a two-tiered explanation for the mental activity of a person who has an inclination for contempt. Perhaps a person's mental state can actually betray a lack of commitment to the moral law. For example, let's say a person has committed an immoral act. One may be tempted to view such a person with contempt (this would be the "inward *looking down* on someone in comparison with others" that Kant describes in §39). At this stage, one can *actually*

view them with contempt (i.e. by saying to oneself “This person is beneath me. This person is of no moral value whatsoever,” etc.), or one can overcome the inclination that would result in a violation of my duty to respect all my fellow human beings (i.e. by saying “Even though this person has committed immoral acts, he is still a human being and on account of that is still owed respect,” or “Even the most vicious of scoundrels, like him, have some connection to the moral law *via* their moral predisposition,” etc.). On this model we still have the inherently vicious inclinations, but we can see how the mere presence of them has not translated to any sort of attitude of contempt for the other.

We get a similar model in Kant’s discussion of radical evil in *Religion Within the Boundaries of Mere Reason*. In a footnote to his discussion of human nature’s propensity to evil, Kant distinguishes between propensities, inclinations, and passions. Propensity is something like the potential and tendency to have specific inclinations. An inclination is the *actual* desire that seeks satisfaction. Kant writes, “*Propensity* is actually only the *predisposition* to desire an enjoyment which, when the subject has experienced it, arouses *inclination*.”³¹⁷ So, using the example of the scoundrel, we have a propensity to disrespect scoundrels before we are ever confronted with one. We develop an inclination toward disrespect when we meet an actual scoundrel. Kant then introduces a third level to this model: “Above inclination there is, finally, still another level of the faculty of desire, *passion* (not *emotional agitation*, for this belongs to the feeling of pleasure and aversion), or an inclination that excludes mastery over oneself.”³¹⁸ When we have actually succumbed to our inclination toward disrespect, the inclination becomes a passion (which

³¹⁷ Rel. 6:29n

³¹⁸ *Ibid.*

is still an inclination, strictly speaking, but one that exceeds our resources for self-control). Passion takes over when one says to oneself, for example, “This scoundrel is beneath me and I have nothing but contempt for him.” Passion need not lead to concrete action. Perhaps circumstances are such that the agent deems it unwise to behave in certain ways, but the action is not averted through moral discipline.

In succumbing to passion in this case, one is already morally culpable. That is, one has already “done” an immoral “deed.” This might seem like an odd claim, since there is no overt behavior yet attached to the passion. Further on in the same section of the *Religion*, Kant claims that there is ambiguity in the term “deed,” and that there are at least two senses a person can commit a deed: “Now the term “deed” can in general apply just as well to the use of freedom through which the supreme maxim (either in favor of, or against the law) is adopted in the power of choice, as to the use by which the actions themselves (materially considered, i.e. as regards the objects of the power of choice) are performed in accordance with that maxim... The former is an intelligible deed, cognizable through reason alone apart from any temporal condition; the latter is sensible, empirical, given in time.”³¹⁹ This is the sense in which a person who, for example, ridicules another in order to satisfy an inclination toward disrespect is doubly guilty. The ridiculer commits two distinct but connected deeds. This model does not place the exercise of choice in the phenomenal level. That is, the difference between having an inclination and succumbing to passion does not result from a choice made in the face of an inclination. Rather, the choice, understood as the choice of incorporating the moral law into one’s supreme

³¹⁹ Rel. 6:31.

maxim, is made by the noumenal self. As a result of that choice one will experience, in this case, one of two phenomenal mental states. That is, if one has incorporated the moral law into one's supreme maxim, then the inclination will not spill over into passion; whereas if one has not chosen to give priority to the moral law, then it is possible one will succumb to passion.

There is a bit of a complication we should clear up in this section of the *Religion* that might appear to rub against the interpretation of the duties of respect I have presented. According to my reading of the duties of respect there is both a psychological component (anchored in inclinations toward disrespect) and a behavioral component, which are the means of satisfying inclinations toward disrespect. That is, we have a duty to avoid certain modes of behavior and a duty to avoid viewing other human beings with contempt or disrespect. Furthermore, I hold that inclinations toward disrespect are immoral (i.e. a violation of our basic duty to respect all human beings) even if circumstances prevent the person from *acting* on them. In the *Metaphysics of Morals* Kant discusses three *vices* that violate duties of respect for other human beings. I interpreted these vices in terms both of the inclinations behind them and in the behavior that is prohibited. In this section of the *Religion*, though, Kant writes, "The propensity to evil is a deed in the first meaning [involving the power of choice in formulating one's maxim], and at the same time the formal ground of every deed contrary to law according to the second meaning [i.e. of a deed] that resists the law materially and is then called vice."³²⁰ Here Kant seems to be

³²⁰ Rel. 6:31

reserving the term “vice” exclusively for the behavioral side of immoral deeds. This seems like it might be in conflict with my mixed interpretation of our duties of respect.

In order to sustain my interpretation it is necessary to examine the precise role of the vices of disrespect with regard to the more general requirement of respect for all human beings. First, though, we should reconstruct our model with the addition of vice at the behavioral end of the spectrum. We have a propensity to disrespect some, if not all, of our fellow human beings.³²¹ This propensity turns into individual inclinations toward disrespect in actual instances, with regard to actual other people. In the face of an inclination to disrespect we either demonstrate commitment to the moral law or we fail to. In the case of failure, we succumb to passion. If we succumb to passion, then, in some circumstances, we will commit particular vices against other human beings (i.e. we may defame others, or ridicule them). One possible interpretation of this model would be to say that moral blame arises only in the last stage, when inclinations get acted out in the exercise of vice. I reject this claim and instead contend that moral blame originates when inclinations turn into passions. That is the moment in which it becomes evident that our fundamental choice does not properly prioritize the requirements of the moral law with the demands of inclinations. We have inverted the moral order in the determination of our maxims and inclinations take precedence over the moral law in the exercise of our will. Whether or not one actually acts on one’s passion is determined by factors other than one’s choice, or commitment to the moral law. In the *Religion*, Kant seems to take a fairly

³²¹ This propensity is not exclusive to the scoundrel, as Kant repeatedly points out. We also may feel a competitive urge with everyone else in which we try to view ourselves as superior to others. See e.g. Rel. 6:27.

radical stance in assigning moral culpability, or more precisely, guilt. Oftentimes, he asserts, people demonstrate a certain moral smugness when their actions are always in conformity with the law, and even when there are no obvious inclinations pulling them to vice.

This is how so many human beings (conscientious in their own estimation) derive their peace of mind when, in the course of actions in which the law was not consulted or at least did not count the most, they just luckily slipped by the evil consequences; and [how they derive] even the fancy that they deserve not to feel guilty of such transgressions as they see others burdened with, without however inquiring whether the credit goes perhaps to good luck, or whether, on the attitude of mind they could well discover within themselves if they just wanted, they would not have practiced similar vices themselves, had they not been kept away from them by impotence, temperament, upbringing, and tempting circumstances of time and place (things which, one and all, cannot be imputed to us).³²²

While in this passage Kant is involved in a more general discussion about radical evil and guilt, the point Kant is making here holds in the case of distinguishing between inclinations and passions. It would be incorrect to argue that a person commits a violation of their duty to respect to other human beings only when they commit an *actual* vice. Moral culpability begins with harboring an attitude of contempt, even if circumstances prevent an actual vice from being carried out.

If one overcomes the inclination toward disrespect, then one will absolutely abstain from any sort of behavior that would degrade the humanity of the other. One will acknowledge the rights afforded to him insofar as he is a human being (i.e. his human

³²² Rel. 6:38.

rights). If one fails to overcome the inclination toward disrespect, then it is impossible that one will have the moral fortitude or will to condemn, or even in some cases to avoid engaging in, behavior that degrades the scoundrel's humanity, to go back to our example. In that case, the person harboring contempt might be doubly guilty: 1) One has already degraded the humanity of the other by contempt, and 2) One will, circumstances permitting, engage in behavior that manifests the contempt one harbors internally (i.e. by defaming the person, or ridiculing the person, or by demanding that the other person abdicate their self-respect in comparison with others). On this reading we could say that it would be *worse* if a person actually acts out on his or her contempt, but the immorality of their choice does not depend on the circumstances in which a person finds him or herself.

III. Merit-Based Respect

There is a third context in which Kant employs the term "respect," in which "respect" clearly has a different meaning from respect for the moral law and the sort of respect that is owed to *all* human beings, expressed negatively through the duties of respect. It is, like the second sense, directed toward other human beings. Unlike the second sense, though, it is not directed to all of humanity equally. It is morally possible to withhold respect from someone in this third, merit-based sense. That is, it is not a moral requirement to have respect for all human beings in the merit-based sense. Merit-based respect is, as we will see, reserved for those who have demonstrated a comparatively high level of commitment to the moral law. It is a respect that can be earned.

This third sense of respect is already contained in a section of the *Critique of Practical Reason* we examined in our discussion of respect for the moral law. In fact, the inclusion of this third sense of respect seems a bit confusing at first. After having established to his satisfaction that respect for the moral law is the only incentive for actions that have moral worth, and then having explained the effect of the moral law on self-love and self-conceit, Kant writes, “*Respect* is always directed only to persons, never to things.”³²³ This statement seems odd. The moral law clearly is not a person, but we also are to have respect for the moral law. The sentence, then, appears to be false. Strictly speaking, of course, it must be false. It is obviously not the case that respect is directed *only* to persons.

The confusion dissipates if we read the sentence as a shift in the discussion; Kant is now discussing a sense of “respect” that is not the same thing as respect for the moral law. Respect *in this sense* is indeed directed only to persons. This shift in meaning, though, is not to the second sense of respect that we have already discussed. Kant writes, “A human being can also be an object of my love, fear, admiration even to amazement and yet not be an object of respect.”³²⁴ But respect in the second sense we have discussed *is owed to all human beings*. So a human being would *always* be an object of respect in the second sense. It is clear, then, that Kant must have some other sense of respect in mind in this section. Respect, in this sense, is reserved for *virtuous* human beings. Kant describes the sort of person who does in fact receive respect when he writes, “[B]efore a humble common man in whom I perceive uprightness of character in a higher degree than I am

³²³ CPrR 5:76.

³²⁴ *Ibid.*

aware of in myself *my spirit bows...*”³²⁵ In saying that his “spirit bows” to the person who has “uprightness of character,” Kant is asserting an attitude of respect for the person. A few sentences later Kant writes, “*Respect is a tribute* that we cannot refuse to pay to merit, whether we want to or not; we may indeed withhold it outwardly but we still cannot help feeling it inwardly.”³²⁶ This sentence is striking for a variety of reasons: Firstly, this respect we “pay to merit” is *involuntary* much in the same way as respect for the moral law is irresistible for human beings. As Kant says, “we...cannot help feeling it [respect] inwardly” to the morally upright character.

Secondly, this respect we pay to merit has a complex, ambivalent effect on us (much like respect for the moral law). In seeing a person who exhibits an uprightness of character that exceeds our own, we become more aware of our own moral shortcomings. Awareness of one’s moral limitations is obviously not a pleasant experience. In fact in the face of a morally upright character, we may be tempted to *defame* the person in order to lessen the burden of our own conscience: “So *little* is respect a feeling of *pleasure* that we give way to it only reluctantly with regard to a human being. We try to discover something that could lighten the burden of it for us, some fault in him to compensate us for the humiliation that comes upon us through such an example.”³²⁷ In being confronted with an example of a morally upright person we experience displeasure in the awareness of our own moral limitations.

³²⁵ *Ibid.*

³²⁶ *Ibid.*

³²⁷ *Ibid.*

Yet we experience some positive feeling in these instances. We witness the *actuality* of the moral law. We see a real instance of it in the world: “His example holds before me a law that strikes down my self-conceit when I compare it with my conduct, and I see observance of that law and hence its *practicability* proved before me in fact.”³²⁸ This can certainly be construed as at least a partially pleasant experience. We see the possibility of observance of the moral law (something we inherently know to be valuable) in the world of “fact.” Furthermore, though Kant does not explicitly mention this here, we might gain a sense of *solidarity* with another in our shared interest in the moral law. Kant does say that even if we simultaneously are aware of a similar, or shared, degree of commitment to the moral law in us, we nevertheless still feel respect for the other: “Now, I may even be aware of a like degree of uprightness in myself, and yet the respect remains.”³²⁹

A third striking feature of this sense of respect is its entirely inward character. Unlike respect in the second sense, which has a behavioral component (expressed negatively through the prohibition of contemptuous modes of behavior), this sense of respect is paid exclusively through an inward feeling. There is an irresistible effect on us when faced with a moral exemplar that mirrors the inherent pull the moral law has on us in our recognition of its universal validity.

This third sense of respect, what I will call “merit-based respect,” is something reserved to morally disciplined people. Kant does not elaborate further on this sense of

³²⁸ *Ibid.*

³²⁹ *Ibid.*

respect in this section of the *Critique of Practical Reason*, but in his *Lectures on Ethics* this sense of respect is examined in more detail. As a result, the picture becomes a bit more complicated. As of now, merit-based respect is reserved for people who fit under the rather broad category of those who possess an “uprightness of character.” We may wonder, though, what constitutes having an “uprightness of character.” It seems like it must have something to do with the fulfilling of duties, even if we have countervailing inclinations. But questions arise at this point: Is it in the performance of *all* our duties? Or is it just in our narrow ones? What about our positive duties?

We get closer to an answer to this line of questioning in the *Collins Lectures*. The context is a discussion of what Kant calls “true honor” as opposed to mere “vanity.”³³⁰ True honor has both negative and positive senses. The negative sense of true honor consists in not violating one’s narrow duties. As Kant says, “If he everywhere acts in accordance with duty, even the greatest scoundrel must pay him respect, and even the tyrant cannot rob him of the idea of self-respect, and the worth arising from his merit.”³³¹ It is clear Kant is talking about narrow duties when he continues by saying, “This concept is negative, because to that extent the agent merely refrains from acting contrary to his humanity and its worth...”³³² So, respect is owed to those who are able to conduct themselves in a way that does not cause them to violate the narrow duties.

Importantly, this is *not* the same respect that we owe to *all human beings* (i.e. our second sense of respect). This third type of respect, unlike respect in the second sense, can

³³⁰ Collins Lectures 27:664.

³³¹ Collins Lectures 27:665.

³³² *Ibid.*

be forfeited if we behave in ways that violate our narrow duties. Elsewhere in the *Collins Lectures*, Kant says, “We acquire respect in virtue of good conduct, but honour in virtue of meritorious actions. We forfeit respect by failure to perform duties incumbent on us.”³³³ “Good conduct” consists in the fulfillment of one’s narrow duties (i.e. avoiding prohibited actions). “Meritorious actions” is the comparatively high level of fulfillment of our wide duties; it is through meritorious actions that we receive, not respect, but honor.

We have discussed some of the narrow duties we have towards others at length, delineated by Kant under the heading of duties of respect owed to others. Failure to comport oneself to those duties will automatically forfeit a rightful claim to respect in this merit-based sense. There is another way, though, that one can forfeit respect in this third sense. Namely, in the violation of one’s narrow duties to oneself.

Perhaps the clearest way a person can abdicate the claim to respect is in the vice of what Kant calls “false humility” or “servility.”³³⁴ Just as we are duty-bound to recognize the humanity in others, we are also obliged to recognize it in ourselves. Kant writes, “[F]rom our capacity [*Fähigkeit*] for internal lawgiving and from the (natural) human being’s feeling himself to revere the (moral) human being within his own person...there comes *exaltation* of the highest self-esteem, the feeling of his inner worth (*valor*), in terms of which he is above any price (*pretium*) and possesses an inalienable dignity (*dignitas interna*), which instills in him respect for himself (*reverential*).”³³⁵ Kant provides a list of prohibitions that, if acted upon, would undermine one’s claim to respect. For example, we

³³³ Collins Lectures 27:410.

³³⁴ MM 6:428.

³³⁵ MM 6:436.

are to be “no man’s lackey;” we cannot “let others tread with impunity” on our rights; we cannot become a “parasite or a flatterer or a beggar;” and we are to avoid “kneeling down or prostrating oneself on the ground.”³³⁶ Violating any of these duties is contrary to our dignity as bearers of humanity. Thus, when we violate these duties we forfeit our claim to respect. In servility, a person, as Kant writes in his discussion of the duties of respect, “makes himself unworthy of it [the respect we are owed by other human beings].”³³⁷ This point is echoed in the *Vigilantius Lectures* when Kant says, “If he wishes others to have respect for his person, he must likewise hold fast to it, and show that he respects himself.”³³⁸ Failure to avoid servility is one way in which others may withhold merit-based respect for us.³³⁹

Respect in this third, merit-based sense, then, is actually narrower than it first appeared in Kant’s discussion in the *Critique of Practical Reason*. We earn merit-based respect through the fulfillment of our narrow duties (i.e. through our refusal to commit acts prohibited by the moral law). Through fulfillment of our wide duties (i.e. of beneficence toward others, or the cultivation of our talents) we receive (positive) honor: “We acquire respect in virtue of good conduct, but honour in virtue of meritorious actions. We forfeit respect by failure to perform duties incumbent on us.”³⁴⁰ The contrast is clearly drawn in Kant’s discussion of the positive concept of true honor: “The positive concept of true honor consists in this, that it essentially takes on actions that involve a merit—more

³³⁶ *Ibid.*

³³⁷ MM 6:463.

³³⁸ *Vigilantius Lectures* 27:606-607.

³³⁹ Kant writes, in discussing a vicious person, “At times one cannot, it is true, help inwardly looking down on some in comparison with others...” (MM 6:463).

³⁴⁰ *Collins Lectures* 27:410.

than what is required; so if a man merely does what is required of him, he has simply averted dishonor...he has not exposed himself to contempt and reproof. But if he has done those things that contain more good than obligation required of him, then merit can arise for him, and this may bring him true honor.”³⁴¹ “Contempt and reproof,” here (not to be confused with contempt in the context of the second sense of respect), mean the withholding of merit-based respect.

The basic picture that emerges when we combine the second and third senses of respect go as follows: 1) All human beings are owed respect. Respect on this level means that no one is to be viewed with contempt. Contempt is the denial of a human being’s humanity, their moral predisposition, and their potential for moral action. 2) Those human beings who fulfill their narrow duties to others and to themselves are owed respect. Respect here means the recognition that, through their actions, they have made themselves worthy of humanity. They conduct themselves with both self-esteem and with respect for others.

IV. Conclusion: Respect for the Law and Respect for Persons

Respect for the moral law must be the starting point for any other sense in which we might have respect for a person. We respect others only insofar as 1) the moral law requires us to, or 2) others provide an example of moral uprightness in general and

³⁴¹ Collins Lectures 27:665. Note that Kant here talks as if positive honor is owed to those who have committed good acts that exceed what is required of them. It seems clear, though, that what Kant really has in mind are wide duties. He also says that it is indeed impossible to do more than is incumbent on us since “everything is required” but we can compare our fulfillment of our wide obligations in comparison with other human beings and can receive positive honor when our achievements in the moral realm are compared to those of others.

commitment to the narrow duties that are required by the moral law in particular. In other words, the two senses of respect that are directed towards other people ultimately have their source in respect for the moral law. Respect for the moral law, then, *grounds* respect for other people (in both senses). That is, respect for the moral law is the most basic kind of moral respect. Without respect for the moral law, respect for other people is actually not rationally anchored in anything. In all likelihood, without respect for the moral law, so-called “respect” for other people is mere sentimental outpouring that has nothing to do with serious morality.

An interesting question to consider in light of Kant’s contention that most, if not all, people are actually evil (that is, they privilege self-love over morality) is whether or not respect for the moral law is possible in imperfectly moral people (that is, people who while affectively responsive to the moral law fail to live up to its dictates when faced with enough resistance from inclinations). In my view, respect for the moral law is an affective state brought about by the striking down of self-love when it conflicts with the requirements of the moral law. It follows, then, that cases of authentic respect for the moral law, according to Kant, will be infrequent, vanishingly so. Respect for the moral law will occur in cases where 1) the person experiences a moral crisis in which nature leads him or her in one direction and freedom in another and 2) the person possesses a good will and thus strikes down self-love in favor of morality.

Respect for persons, therefore, comes *after* possession of a good will. While the duties of respect are unconditionally required of us, only those who have achieved a good will are in a position to respect persons in a truly morally worthwhile fashion.

While each sense of respect is distinct from the other, there are similarities between all three senses of respect's impact on our affective state. Each sense of respect dramatizes, internally (in our own minds), the moral crisis that is an inevitable part of our experience of moral life. In all three senses, there is a dynamic (and sometimes antagonistic) interplay between what the moral law imposes on us and what inclination may pull us toward. In the case of respect for the moral law, we must, at times, restrict the efficaciousness of inclination in our decision making in order to do our duty. In the case of our duties of respect, we may be tempted to act on inclinations that would assert superiority over others. Finally, in the case of merit-based respect, our own conscience may be pained by the respect we irresistibly give to moral exemplars. As a result, we may have an inclination to seek out ways in which we can diminish the degree to which we must afford them respect (though of course this could easily lapse into a serious violation of our duties of respect). Each sense of respect, cultivates in us authentic *humility*.³⁴² In his discussion of the vice of servility in the *Metaphysics of Morals*, Kant writes, "True humility follows unavoidably from our sincere and exact comparison of ourselves with the moral law (its holiness and strictness)."³⁴³ Each sense of respect reminds us, in one way or another, of our propensity toward evil. They engender, like much of Kant's moral theory as a whole, honest self-reflection about our own problematic and complex relation to the moral law.

³⁴² Kant offers a definition of humility in *Metaphysics of Morals* when he writes, "The consciousness and feeling of the insignificance of one's moral worth in comparison with the law is humility (*humilitas moralis*)" (6:435).

³⁴³ MM 6:436.

CHAPTER 4

The Sublime and Its Relation to Kant's Moral Philosophy

Introduction

This chapter examines some issues surrounding Kant's account of the sublime and its relation to his moral philosophy. In particular, it focuses on the relation between the moral feeling of respect and the feeling that constitutes a dynamically sublime experience. I begin with a discussion of Kant's description of the sublime as given in *The Critique of the Power of Judgment*. I then consider the relationship between the sublime and morality. In particular, I discuss some similarities and differences between the moral feeling of respect and the feeling of the sublime. In this section, I also consider the indirect contribution to morality that the sublime makes. I discuss the role sublimity has in Kant's broader moral philosophy and consider the implications that sublimity has on subjects who undergo a sublime experience. I argue that sublimity does not reveal anything new to the subject, but it does dramatize the preexisting receptivity to moral ideas that the subject must have in order to have a sublime experience in the first place.

I. On the Sublime in General

Kant holds that the sublime, like the beautiful, pleases for itself. That is, the pleasure we experience from the sublime does not seek some further satisfaction above and beyond the intrinsic feel of the experience itself. So the satisfaction we experience in the sublime, like the beautiful, does not depend on sensation (like agreeable sensation) nor does it depend on a determinate concept (like judging something to be good of its

kind).³⁴⁴ Early on in his treatment of the sublime Kant writes that judgments about the sublime are, “related to concepts, although it is indeterminate which, hence the satisfaction is connected to the mere presentation or to the faculty for that, through which the faculty of presentation or the imagination is considered, in the case of a given intuition, to be in accord with the **faculty of concepts** of the understanding or of reason, as promoting the latter.”³⁴⁵ Concepts, therefore, may play some role in the satisfaction we experience in the sublime, but they must be indeterminate.

The characteristic mark of the sublime is the experience of a mental movement in which the mind becomes agitated and stirred to excitement. Kant writes, “For since the feeling of the sublime brings with it as its characteristic mark a **movement** of the mind connected with the judging of the object, whereas the taste for the beautiful presupposes and preserves the mind in **calm** contemplation, yet this movement is to be judged as subjectively purposive (because the sublime pleases).”³⁴⁶ There are two sorts of mental movements that can take place in a sublime experience and there are two forms of the sublime that correspond to each variety of sublime mental movement. Kant refers to them as the mathematical sublime and the dynamical sublime. He writes, “Thus this movement is related through the imagination either to the **faculty of cognition** or to the **faculty of desire**, but in both relations the purposiveness of the given representation is judged only with regard to this **faculty** (without an end or interest): for then the first is attributed to

³⁴⁴ CPJ § 23 5:244.

³⁴⁵ *Ibid.*

³⁴⁶ CPJ §24 5:247.

the object as **mathematical**, the second as a **dynamical** disposition of the imagination.”³⁴⁷

The Mathematical Sublime

Kant provides an initial definition of the sublime when he writes, “We call sublime that which is **absolutely great**.”³⁴⁸ Something is absolutely great when it “is great beyond all comparison.”³⁴⁹ The sublime, it seems to follow, must be something that dwarfs all other things. In other words, we say a thing is sublime when it has a vast *magnitude*. Greatness of magnitude is the defining feature of the mathematical sublime. Since, as Kant argues, the sublime is to be found in a formless, limitless object, the magnitude of an object must be, in a sense, limitless. Obviously anything we have a sensory perception of can be quantified in some way (and therefore is not limitless). Does it follow from this that nothing in nature, properly speaking, is actually sublime? This seems to rub against Kant’s theory of the sublime since he argues that objects in nature can, in fact, be rightly called sublime.

Kant, at this point, makes an important distinction that explains why he thinks we can still talk about encountering objects in experience that we can rightly think of as sublime. The distinction Kant draws is between our capacity to attach numerical concepts to an object and our capacity to take in objects aesthetically. Kant writes, “The estimation of magnitude by means of numerical concepts (or their signs in algebra) is mathematical,

³⁴⁷ *Ibid.*

³⁴⁸ CPJ 5:248.

³⁴⁹ CPJ 5:249.

but that in mere intuition (measured by the eye) is aesthetic.”³⁵⁰ Aesthetic estimation involves the presentation of an object in one, unified representation; it does not break up a magnitude into units of measurement and then attach some numerical concept to it. In this aesthetic sense, then, objects can still be sublime because they are not easily unified in intuition.

The mathematical sublime is, as Kant asserts, that which “in comparison with which everything else is small.”³⁵¹ Of the objects we encounter in nature, there will always be the possibility of it being large enough to have a greater numerical concept attached to it. No matter how tall the tallest mountain in the world might be, it is possible that we can imagine a still-taller mountain, which would have a greater numerical magnitude. All the objects we encounter in nature actually have limits, even if they require enormous calculation and imagination to comprehend them. So, in nature, nothing can be sublime in a non-aesthetic sense: “Here one readily sees that nothing can be given in nature, however great it may be judged to be by us, which could not, be amplified for our imagination up to the magnitude of a world... Thus nothing that can be an object of the senses is, considered on this footing, to be called sublime.”³⁵² Though nothing in nature is infinite, there can be, in nature, sublime objects. We can think of objects in nature, then, as sublime when we confront objects within it that cause us to think of

³⁵⁰ CPJ §25 5:220.

³⁵¹ CPJ §25 5:250.

³⁵² *Ibid.*

unbounded magnitude: “Nature is thus sublime in those of its appearances the intuition of which brings with them the idea of its infinity.”³⁵³

Objects bring about, in our minds, the idea of infinity only when the mind encounters vast objects, not easily subsumed within a single representation. The mathematical sublime is only encountered when our sensory capacities are unable to absorb wholly, or take in, some vast object. Phenomenologically, the experience of aesthetic estimation might be understood as our eyes finding the borders, or limits, of the object. Once the borders are taken in, one can reflect on the object as a whole. An experience of aesthetic estimation, therefore, becomes sublime when the eye cannot take in the borders of the object and thus cannot have a representation of the entire object at one time.

Kant describes this experience of being unable to take in the entirety of an object as consisting both of pleasant and unpleasant elements. In other words, the mind experiences a complex feeling, which cannot simply be labeled as wholly pleasurable or displeasurable.³⁵⁴ The mathematical sublime is initially unpleasant because we fail to take in the vast object aesthetically, all at once in one representation. This failure results in frustration when confronted with the limitation of our sensory and imaginative capacities. Thus the mind cannot comprehend the infinite, though. Kant writes, “Comprehension

³⁵³ CPJ 5:255.

³⁵⁴ Paul Guyer (1996) makes the observation that Kant fluctuates between arguing that the experience is, all at once, complex, or rather is a sequence of feelings which each individual feeling being either pleasant or unpleasant. He writes, “Kant now makes it unclear whether what all this produces is a single but complex feeling which is both displeasurable yet pleasurable, or a succession of simple feelings which begins with displeasure but must end in pleasure” (211). In this chapter I will write as though the sublime involves a series of simple feelings, culminating in a pleasurable one, though Guyer rightly points out the lack of clarity in Kant’s exposition.

becomes ever more difficult the further apprehension advances, and soon reaches its maximum, namely the aesthetically greatest basic measure for the estimation of magnitude.”³⁵⁵ So, our comprehension of magnitude is limited and an experience of this limitation is unpleasant. The mind is frustrated by its inability to comprehend the infinite, even as it continually strives to do so.

This displeasure, however, is not permanent because in our frustrated desire to take in the whole, we realize that we must have another mental faculty that causes us to think of an infinite magnitude as a whole in the first place: “But what is most important is that even being able to think of it as **a whole** indicates a faculty of mind which surpasses every standard of sense.”³⁵⁶ The very cause of our initial displeasure, in the end, becomes the revelatory source for an awareness of an “unlimited capacity.”³⁵⁷ Kant describes this progression from frustration to satisfaction as follows: “The **quality** of the feeling of the sublime is that it is a feeling of displeasure concerning the aesthetic faculty of judging an object that is yet at the same time represented as purposive, which is possible because the subject’s own incapacity reveals the consciousness of an unlimited capacity of the very same subject, and the mind can aesthetically judge the latter only through the former.”³⁵⁸ Mathematical sublimity, therefore, is found in vast objects, in things like the Pyramids or the Milky Way. This is not because they are limitless themselves, but rather because our eye’s strain to take in the whole of the Pyramids or in our eye’s (and also in our imagination’s) failure to comprehend the whole Milky Way inevitably gives rise to the

³⁵⁵ CPJ §26 5:251.

³⁵⁶ CPJ 5:255.

³⁵⁷ CPJ §27 5:259

³⁵⁸ *Ibid.*

idea of limitlessness. The mathematical sublime is the kind of sublimity that centers upon vastness. The link between vastness and sublimity was a well-established idea by the time Kant wrote the *Critique of the Power of Judgment*.³⁵⁹ Kant, however, does not limit his discussion of the sublime to vast objects. There is another form of sublimity which is not related, directly, to vastness at all, but rather to notions of *power* and *danger*. Kant calls this sort of sublimity “dynamical.”³⁶⁰

The Dynamical Sublime

Kant defines the dynamical sublime in the following way: “Nature considered in aesthetic judgment as a power that has no dominion over us is **dynamically sublime**.”³⁶¹ The definition is striking because of Kant’s description of nature as a power, but as a power that has no dominion over us. That is, Kant, in discussing the dynamical sublime argues that nature is judged to be a power, but, in some way, it is a power under whose dominion we do not entirely fall.

Kant lists some objects which can elicit an experience of the dynamical sublime: “Bold, overhanging, as it were threatening cliffs, thunder clouds towering up the heavens, bringing with them flashes of lightning and crashes of thunder, volcanoes with their all-destroying violence, hurricanes with the devastation they leave behind, the boundless

³⁵⁹ Joseph Addison is perhaps the most famous example of an aesthetic theorist who thinks of the sublime in terms of vastness. Another thinker who characterizes the sublime as vast include Alexander Gerard in his *Essay on Taste*, published in 1759.

³⁶⁰ Kant’s key predecessor in describing the sublime in terms of an object’s fearfulness seems to be Edmund Burke who argued that sublime feelings involve a safe encounter with a fearful object. Burke defines the sublime objects as follows: “Whatever is fitted in any sort to excite the ideas of pain, and danger, that is to say, whatever is in any sort terrible, or is conversant about terrible objects, or operates in a manner analogous to terror, is a source of the sublime” (39).

³⁶¹ CPJ §28 5:260

ocean set into a rage, a lofty waterfall on a mighty river, etc.”³⁶² Kant seems to think that the paradigmatic cases of dynamically sublime objects are mighty and dangerous objects in nature that carry with them awesome force. Kant says that objects such as these “make our capacity to resist into an insignificant trifle in comparison with their power.”³⁶³ As a result of their immense power to overcome our resistance, dynamically sublime objects are *fearful*: “If nature is to be judged by us as dynamically sublime, it must be represented as arousing fear...For in aesthetic judging the superiority over obstacles can only be judged in accordance with the magnitude of the resistance. However, that which we strive to resist is an evil, and, if we find our capacity to be no match for it, an object of fear.”³⁶⁴ We can imagine what would happen to us if we fell into a mighty river. We would simply be at the mercy of its rapids and force. Furthermore, in such a case, there is little we can do to ensure our very survival. Our lives would be at the mercy of forces that are utterly beyond our control. Objects that possess this immense destructive powers are what Kant holds are capable of giving rise to a dynamically sublime experience.

It is important to keep in mind that these harrowing experiences with those terrifying objects must be *imaginative*. Authentic fear cannot be a constituent of a dynamically sublime experience. Kant writes, “Someone who is afraid can no more judge about the sublime in nature than someone who is in the grip of inclination and appetite can judge about the beautiful.”³⁶⁵ An experience is dynamically sublime only when there is no genuine threat from the sublime object. Observing a raging river from a safe

³⁶² CPJ §28 5:261

³⁶³ *Ibid.*

³⁶⁴ CPJ §28 5:260

³⁶⁵ CPJ §28 5:261

distance, for example, allows us to imagine ourselves confronting the river and to contemplate its power. This is done, though, while also knowing that we are safely removed from it and that it poses no real threat. We recognize the object as fearful without actually *being afraid* of it. This happens when, “We judge it in such a way that we merely **think** of the case in which we might wish to resist it and think that in that case all resistance would be completely futile.”³⁶⁶ Being in the grip of authentic fear is simply terrifying and it would be impossible for the complex feeling of the sublime to come about. A dynamically sublime experience is marked, in the end, by “joyfulness.”³⁶⁷

Imagined fear and joyfulness, then, are the two defining features of the dynamically sublime experience. The sublime, then, is a complex feeling consisting of both positive and negative aspects. Having to imagine the utterly futility of resistance is not a straightforwardly pleasant psychological event. It forces us to accept our finitude and, ultimately, our own mortality. When one imagines resisting a raging river, for example, one cannot help but entertain the possibility that one would die as a result of falling in. These considerations are not pleasant to think about because perhaps our most fundamental natural inclination is to self-preservation. The fearful aspects of dynamically sublime objects, however, are also the very same features that allow us to discover our capacity of reason: “they [frightening objects] elevate the strength of our soul above its usual level, and allow us to discover within ourselves a capacity for resistance of quite another kind, which gives us the courage to measure ourselves against the apparent all-

³⁶⁶ CPJ §28 5:260

³⁶⁷ CPJ §28 5:261

powerfulness of nature.”³⁶⁸ Kant seems to be suggesting that when we are confronted with a dynamically sublime object and realize the futility in resisting its power, we still nevertheless “discover” within ourselves a “capacity for resistance.” This capacity is not subject to the power of the fearful object. The fearful object can do nothing to undermine or diminish this capacity.³⁶⁹

Kant argues that this capacity allows us to see ourselves as *independent of* and *superior to* nature. He writes,

“The irresistibility of its [the dynamical sublime object] certainly makes us, considered as natural beings, recognize our physical powerlessness, but at the same time it reveals a capacity for judging ourselves as independent of it and a superiority over nature on which is grounded a self-preservation of quite another kind than that which can be threatened and endangered by nature outside is, whereby the humanity in our person remains undemeaned even though the human being must submit to that dominion.”³⁷⁰

When confronted with dynamically sublime objects we realize that it is possible to retain our humanity, or our freedom, even in the face of physical danger. Authentic moral action, or adherence to the moral law, is not subject to the dangers and contingencies that

³⁶⁸ CPJ §28 5:261

³⁶⁹ Clewis describes this process as reinforcing and highlighting the distinction between our natural being and our moral being: “The object that elicits a judgment of the dynamical sublime is viewed as having so much power that, insofar as the aesthetic subject is a natural being, she cannot find a power that is comparable to the might demonstrated by the object. Nonetheless, she does contain a power or ability insofar as she is a free being” (68).

Malcolm Budd also argues that the kind of resistance that is opened up here is our capacity to act morally, even in the face of natural destruction: “The second [mode of resistance] is the ability not to abandon our moral principles and commitment to morality, even under the greatest pressure” (The Aesthetic Appreciation of Nature 78).

³⁷⁰ CPJ §28 5:261-2

abound in the natural world. The forces of nature are incapable of undermining our humanity, so long as we continue to adhere to the moral law.

II. Sublimity and Morality

A) *The Moral Feeling of Respect and Sublimity*

There are several similarities between the moral feeling of respect and the experience of sublimity. Perhaps the most striking similarity between respect and sublimity is that both experience involve a complex feeling, consisting of negative and positive aspects. In discussing Kant's discussion of respect in the *Critique of Practical Reason*, Paul Guyer writes, "The structure of Kant's account is the same as that which he was to use two years later in his characterization of the feeling of the sublime. The palpable painfulness of the frustration of the imagination in attempting to grasp the infinite or our fear of physical destruction itself reveals to us the higher power of reason to form the very idea of an infinite in the first place and to grant us a value that is immune from physical destruction."³⁷¹ In Kant's accounts of both respect and sublimity, the subject is left with a complex, though unified, feeling.

Another related similarity between respect and sublimity is that the *source* of the negative aspect of the feeling is the very same as that of the positive feeling. In the case of the mathematical sublime, the imagination undergoes an unpleasant experience of its limits. It attempts to comprehend the infinite, but fails to do so. However, that very failure is what gives rise to an awareness that there must be some faculty within us that is

³⁷¹ Paul Guyer, *Kant and the Experience of Freedom*, 359.

beyond, in some way, the limitations of the senses and of imagination. Similarly, in the case of the dynamical sublime the revelation that we can within ourselves a capacity to withstand physical pressure and to remain resolute in our commitment to morality comes only in the imaginative grappling with a physically unconquerable force in nature. The negative “moment” of the dynamical sublime is the very thing that makes the positive “moment” possible.

Respect also has this positive-negative structure. It involves limiting or infringing upon one’s self-love and also in striking down completely one’s self-conceit. Both of these acts are unpleasant because they require us to adjust and to diminish our natural inclinations. But Kant also describes respect as consisting of a positive element. The positive aspect in respect is, in striking down self-conceit and in infringing upon one’s self-love, we are able to act freely, which for Kant is the only legitimate ground for a claim to self-worth. It is only through the negative aspects of respect that a person can act freely; the negative feeling is the ultimate source for the positive.

Respect and sublimity also highlight the distinction between our natural being and our moral being. With regard to respect, this distinction is fairly obvious. Inclinations are natural occurrences for all human beings, and it is only after one has made the decision to adhere to the moral law does one reconsider and diminish the role of inclinations in forming our behavior. It is only through the exercise of pure practical reason does self-love become rational and in accordance with the demands of morality. In a sense, then, our supersensible freedom must intrude upon the otherwise normal operations of our natural selves. In the case of the dynamical sublime the most fundamental natural

inclination we have is undermined, yet there remains something in us that can still resist. We might think of self-preservation as the most basic and fundamental inclination we have. In the experience of the dynamical sublime, however, we are able to distance ourselves even from this inclination. We identify a part of ourselves that is not attached to self-preservation.

The mathematical sublime too, like respect, causes us to have an experience that reinforces the notion that there are faculties contained in us that transcend our natural, empirical being. In the case of the mathematical sublime, our imagination is unable to comprehend the infinite, but through the failure we are made more aware that some capacity within us causes our mind to *strive* to take in the infinite as a unified whole. Strikingly, this realization causes in us a feeling of *respect*:

Our imagination, even in its greatest effort with regard to the comprehension of a given object in whole of intuition (hence for the presentation of the idea of reason) that is demanded of it, demonstrates its limits and inadequacy, but at the same time its vocation for adequately realizing that idea as a law. Thus the feeling of the sublime in nature is *respect* for our own vocations, which we show to an object in nature through a certain subreption (substitution of a respect for the object instead of for the idea of humanity in our subject), which as it were makes intuitable the superiority of the rational vocation of our cognitive faculty over the greatest faculty of sensibility.³⁷²

Again, we can see how the mathematical sublime might be an indirect contribution to morality because it reveals to us that we have vocations that go beyond, and are

³⁷² CPJ §27 5:257

independent of, our natural being. We feel, in an immediate way, the “superiority of the rational vocations of our cognitive faculty over the greatest faculty of sensibility.”³⁷³

The dynamical sublime also simulates, in a way, the striking down of self-conceit that takes place when one develops consciousness of the moral law. Self-conceit, according to Kant, is wholly incompatible with morality because it attaches some value to the self prior to the self’s adherence to the moral law. Kant writes that self-conceit involves, “claims to esteem for oneself that precede accord with the moral law.”³⁷⁴ These claims, for Kant, must be “null and quite unwarranted because certainty of a disposition in accord with this law is the first condition of any worth of a person, and any presumption prior to this is false and opposed to the law.”³⁷⁵ Self-conceit, then, must be grounded in something other than pure practical reason. Self-conceit might result from thinking about one’s fortune, or one’s intelligence, or one’s good looks prior to moral considerations, but respect strikes down all these considerations and turns the subject’s attention to the moral law, first and foremost.

The dynamical sublime seems to have a similar effect on the subject. In considering the likelihood of being physically overwhelmed by a force in nature, one realizes that no matter how rich one is, or how smart one is, there is no way to resist the force that would be coming upon oneself. However, one becomes more acutely aware that it remains possible to remain morally resolute and to maintain one’s commitment to morality, even in the face of physical or material destruction. The dynamical sublime,

³⁷³ CPJ §27 5:257

³⁷⁴ 5:73

³⁷⁵ 5:73

like respect, seems to belittle, or to diminish, the importance of non-moral values. It makes us experience, firsthand, the idea that the highest and best possibilities for us reside in committing ourselves to the moral law.

While there are interesting similarities between sublimity and respect, it is also important to observe that there are some of the differences between them. One difference between them seems to be in the phenomenological order of the experience. In the case of moral respect, the subject attains awareness of the moral law and then must rein in or diminish one's inclinations in order to abide by the dictates of pure practical reason. Kant writes, "Hence the moral law unavoidably humiliates every human being when he compares with it the sensible property of his nature."³⁷⁶ The negative moment, although in the case of respect culminates in a positive feeling, comes *after* recognition of one's capacity for moral action. Once one has actually subjugated some inclinations in the name of morality one experiences the joy that comes from being free.

The sublime, by contrast, has the phenomenological story going in the opposite direction. The sublime initially has a negative moment in which one struggles against the limits of the imagination (mathematical sublime) or in which one imaginatively struggles with a force that cannot be overcome (dynamical sublime). The positive moment emerges only after the limitations of one's imagination or one's bodily strength are recognized by the subject.

³⁷⁶ 5:74

Another important difference between respect and sublimity is in the differing relations each has to moral action. In the case of the sublime, all that is revealed to the subject is a faculty that he or she has. The faculty need not be exercised in the case of the sublime, rather it is sufficient that the subject be made aware that he or she does indeed have the faculty. The feeling culminates, for Kant, in a conscious awareness of our superiority to nature (just because we have a faculty that, if exercised, will render us superior). He writes, “Thus sublimity is not contained in anything in nature, but only in our mind, insofar as we can become conscious of being superior to nature within us and thus also to nature outside us (insofar as it influences us). Everything that arouses this feeling in us, which includes the **power** of nature that calls forth our own powers, is thus (although improperly) called sublime.”³⁷⁷ The sublime is, according to this passage, a state of mind that comes about when one encounters objects in nature which reveal a person’s superiority over nature *to his or herself*.

Respect, by contrast, is the *effect* on one’s experience of adopting the moral law as one’s fundamental maxim. The attitude of respect emerges when one actually has exercised the faculty of pure practical reason. Kant describes the feeling of respect when he writes:

Freedom, the causality of which is determinable only through the law, consists just in this: that it restricts all inclinations, and consequently the esteem of the person himself, to the condition of compliance with its pure law. This restriction now has an effect on feeling and produces

³⁷⁷ CPJ §28 5:264

the feeling of displeasure which can be cognized a priori from the moral law. It is, however, so far a *negative* effect which...mainly infringes upon the activity of the subject so far as inclinations are his determining grounds...³⁷⁸

Respect only comes about when one *actually has* made a morally relevant decision. That is, the moral law has provided incentive (namely, itself) for a specific course of action.

Respect is a feeling that emerges out of this sort of situation.³⁷⁹

Moral feeling is necessary in order for someone to have a sublime experience, but the subject need not actually make a morally worthy decision to undergo a sublime experience. One needs to have the idea of the moral law or of freedom in order to have a sublime experience, but one does not need to have actually adopted the dictates of the moral law as their fundamental maxim. Kant is skeptical that we will have widespread agreement over which objects in nature are sublime. Kant writes, “We cannot promise ourselves that our judgment concerning the sublime in nature will so readily find acceptance by others. For a far greater culture, not merely of the aesthetic power of judgment, but also of the cognitive faculties on which that is based, seems to be requisite in order to be able to make a judgments about this excellence of the objects of nature.”³⁸⁰

For a person to have a sublime experience at all, they need to be receptive to certain ideas

³⁷⁸ 5:78

³⁷⁹ Clewis seems to have a similar idea in mind when he writes, “[T]he moral feeling of respect is part of the act of making a practical and moral judgment in which the moral law functions as the incentive or reason for morally worthy action” (The Kantian Sublime and the Revelation of Freedom 133).

³⁸⁰ CPJ 5:265.

(i.e. the idea of freedom, the moral law, infinity). He writes, “The disposition of the mind to the feeling of the sublime requires its receptivity to ideas.”³⁸¹

Sublimity and respect, therefore, have some striking shared features, but it would be a mistake to take their similarities too far and to treat them as involved in fundamentally the same role within Kant’s Critical system. Sublimity, even though it shares phenomenological structures with respect, and even though it can indirectly contribute to one’s moral being is not to be taken as identical with the moral feeling of respect.³⁸² We also, should be careful not to treat the sublime as, in anyway, providing incentive for moral action. If we did this, then we would, ultimately, be making morality impossible since a peculiar feeling marks sublimity and a feeling can never be the basis for moral action according to Kant. Rather, we should view sublimity as a peculiar state of mind that is *disinterested* (it takes satisfaction only in the presentation of a faculty), which pleases for its own sake, but nevertheless which might still play an important, albeit indirect, role in the cultivation of one’s moral being.

The dynamical sublime also has long-term effects on a person’s moral life. That is, the significance of the experience of the dynamical sublime can also play an indirect role in the long-term cultivation of one’s moral being. The dynamical sublime can reinforce a more steadfast and brave manner in our adherence to the moral law. Kant writes, “But even tumultuous movements of the mind...can in no way clam the honor of

³⁸¹ *Ibid.*

³⁸² Clewis warns against this tendency to identify the two when he writes, “The moralization of the sublime can arise in various ways. For instance, one might reduce the sublime to the moral feeling or identify sublimity with moral feeling” (The Kantian Sublime and the Revelation of Freedom 134).

being a **sublime** presentation, if they do not leave behind a disposition of mind that, even if only indirectly, has influence on the consciousness of its strength and resolution in regard to that which brings with it intellectual purposiveness (the supersensible)” (CPJ 5:273). The dynamical sublime involves a moment during which a capacity is revealed to us, but it also makes a long-term or indelible mark on our way of thinking. Presumably, the indelible mark made on us will, if only indirectly, affect our future behavior. Namely, it will make us more strong and resolute in our commitment to morality.

III. Sublimity and the Problem of Enthusiasm

The sort of both physical and phenomenal characteristics that accompany a sublime experience seem also to be common to paradigmatic cases of enthusiasm. One example of this similarity comes in Kant’s description of a sublime experience as “**astonishment** bordering on terror, the horror and the awesome shudder.”³⁸³ The description provided here by Kant seems to equally apply to a person in an enthusiastic, ecstatic state. Kant recognizes that sublimity and enthusiasm can easily associated with each other; he writes that enthusiasm is a “state of mind” which “seems to be sublime.”³⁸⁴ Furthermore, both enthusiasm and sublimity seem to be understood by Kant as consisting in intense and potentially overpowering *feelings*. Given these similarities, it is no surprise that one might be tempted to link the two together. Perhaps they have a similar source, or perhaps even they often overlap with each other.

This apparent connection between enthusiasm and sublimity, though, has the potential to complicate the relation between sublimity and morality in a troubling way.

³⁸³ CPJ 5:269.

³⁸⁴ CPJ 5:272.

The problem might arise because, according to Kant, enthusiasm is problematic for morality. In fact, enthusiasm often, if not always, makes true morality *impossible*. Kant argues that people in the thrall of enthusiasm cannot be motivated by respect for the moral law, but rather by an affect, which must be, by definition, pathological to some degree or other. Freedom is impossible when one is in the thrall of enthusiasm because one's actions are governed not by free choice, but rather by the particular emotion one is feeling.³⁸⁵ In the *Critique of Practical Reason* Kant writes, "The moral level on which a human being (and, so far as we can see, every rational creature as well) stands is respect for the moral law...By exhortation to actions as noble, sublime, and magnanimous, minds are attuned to nothing but moral enthusiasm and exaggerated self-conceit; by such exhortations they are led into the delusion that it is not duty...which constitutes the determining ground of their actions...but that it is as if those actions are expected from them, not from duty but as bare merit."³⁸⁶ Enthusiasm, according to Kant, may enthrall people and cause them to do the right thing (the sort of action that is in conformity to the demands of morality), but this makes morality impossible in that instance. Only respect for the moral law can be the motivating source of a truly moral action.

³⁸⁵ It is a peculiar feature of Kant's theory of the sublime to assert that it is capable of assisting us in achieving freedom. Most theorists of the sublime in the 18th century seemed to have thought that autonomy is threatened, if not completely undermined, by the sublime. Burke, for example, argues that the sublime has an arresting effect on the mind. He writes, "In this case the mind is so entirely filled with its object, that it cannot entertain any other, nor by consequence reason on the object which employs it" (57).

³⁸⁶ CPrR 5:84-85.

A person in an enthusiastic state, though, is likely to be acting from a principle that is different than respect for the moral law. Being motivated by enthusiasm, therefore, would have disastrous ramifications for morality. Kant writes:

“For when they [people motivated by being in an enthusiastic state of mind] imitate such deeds...not only have they quite failed to fulfill the spirit of the law, which consists in the disposition subjecting itself to the law, not in the lawfulness of the action; not only do they locate the incentive *pathologically* (in sympathy or self-love), not *morally*; but they produce in this way a frivolous, high-flown, fantastic cast of mind, flattering themselves with the spontaneous goodness of heart that needs neither spur nor bridle and for which not even a command is necessary and thereby forgetting their obligation, which they ought to think rather than merit.³⁸⁷

Kant, in this passage, describes three problems with moral enthusiasm: (1) It goes against the “spirit” of morality because it views our relation to the moral law incorrectly. That is, moral enthusiasts see moral action as flowing from, or consistent with, human nature. There is not, in the moment of enthusiasm, a struggle to do the right thing even though our pathological drives might drive us to do the wrong thing. (2) It leads people to do moral deeds for the wrong sorts of reasons. If moral action is grounded in moral enthusiasm, then it cannot be grounded in respect for the moral law. Interestingly, Kant does not (morally) distinguish between being in an enthusiastic state and being driven into a state of sympathy and acting from self-love. Both sources of motivation are *pathological*. (3) An enthusiast has the wrong “cast of mind” necessary for serious moral action. The enthusiast “flatters” him or herself and self-indulgently revels in their

³⁸⁷ CPrR 5:85.

purported “goodness of heart.” Such a character, while capable of isolated moral action, is not fit for a lifetime of moral behavior.

Given this rather negative treatment of enthusiasm in the *Critique of Practical Reason*, it may seem reasonable to suppose that enthusiasm is entirely contrary to morality and should be avoided in order to avoid the moral difficulties Kant describes. It seems to follow that if it is to serve morality in anyway, it cannot lead to enthusiasm. Enthusiasm and sublimity, in other words, must be treated separately and it must be possible to have a sublime experience without falling into enthusiasm. On this reading, the “seeming” of enthusiasm to be sublime must in fact be inaccurate, an inaccuracy which would be dispelled through further analysis of sublime experience. One would expect, then, that Kant, when turning to his theory of the sublime would unambiguously deny enthusiasm as having any positive relationship with the sublime, even though, on first glance, they seem to share several features.

When we turn to the *Critique of the Power of Judgment*, however, the discussion of enthusiasm in the “General remark on the exposition of aesthetic reflective judgments” is far from straightforward. In fact, Kant’s treatment of the enthusiasm and its relation to the sublime, in the CPJ is rather complicated.³⁸⁸ Kant begins his discussion of enthusiasm and sublimity when he writes, “The idea of the good with affect is called **enthusiasm**. This state of mind seems to be sublime, so much so that it is commonly maintained that without it nothing great can be accomplished.”³⁸⁹ So far, these comments would still be

³⁸⁸ This fact has been attended to both by Kirwan (2004) and Clewis (2009).

³⁸⁹ CPJ 5:272.

consistent with Kant's *ultimate* rejection of enthusiasm having any role in sublime experience. Kant, here, would just be showing the superficial similarities between sublimity and enthusiasm, but will go on to show that they are not important, substantive similarities. Again, this "seeming" would have to be shown to be illusory. Kant apparently then does in fact go on to distance sublimity from enthusiasm when he writes that enthusiasm "cannot in any way merit a satisfaction of reason" because it makes impossible "engaging in free consideration of principles, in order to determine itself in accordance with them."³⁹⁰ Enthusiasm, therefore, can never yield pleasure when considered in terms of satisfying one's reason.

Sublimity, unlike enthusiasm, *is* a satisfaction of reason because it reveals to us a capacity we have to realize our humanity (to freely choose to act from the right sort of principle), even when faced with the prospect of physical destruction. From this it would seem to follow that Kant is completely disassociating the sublime from enthusiasm.

The story, however, does not entirely end here. Kant *does* go on to, in some sense, link enthusiasm with sublimity. He writes, "Nevertheless, enthusiasm is aesthetically sublime, because it is a stretching of the powers through ideas, which give the mind a momentum that acts far more powerfully and persistently than the impetus given by sensory impressions."³⁹¹ Enthusiasm, then, can be considered "aesthetically sublime" because those in the grip of enthusiasm are often driven to heroic resistance to obstacles. Kant writes, "Every affect of the **courageous sort** (that is, which arouses the

³⁹⁰ CPJ 5:272.

³⁹¹ *Ibid.*

consciousness of our powers to overcome any resistance) is **aesthetically sublime**.³⁹²

Enthusiasm, it seems, can be considered sublime when the person who is in an enthusiastic state is driven to a similar sort of total resistance as a person feels when they recognize some capacity in them over which nature does not have dominion. This of course does not amount to an unequivocal endorsement of enthusiasm, but Kant does allow for it to have some *aesthetically* redeeming qualities. In fact, in admitting that enthusiasm can, at times, be aesthetically sublime, Kant is implicitly rejecting the idea that enthusiasm is *necessarily* and *always* a threat to morality. While enthusiasm cannot provide motivating grounds for moral action, it can cause a similar movement of the mind towards sublime feelings, which in turn give rise to an increased awareness of oneself as a moral being.

IV. Sublimity, Enthusiasm, and Fanaticism

We noted earlier how Kant argues that the sublime can indirectly contribute to morality. Kant seems to argue that the sublime can leave with us a newfound resoluteness when we think of ourselves as committed to morality. The person undergoing a sublime experience feels energized or invigorated in their commitment to the moral law because they have, in a sense, *felt* their capacity for transcending the natural in an immediate way.

There are two, related parallels one can draw to this sort of experience that people normally undergo, especially in Kant's time: first, through hearing a rousing, edifying sermon which prompts them in an ecstatic and energetic way to commit themselves to

³⁹² *Ibid.*

God (and to morality and righteousness); and second through a religious experience, often consisting in some sort of encounter with the divine (having beatific visions or hearing the word of God). Both sermons and religious experiences can, like sublime experiences, leave an indelible mark on a person and create in them a newfound resoluteness to be moral. Kant seems to present his account of sublime experience as providing the same sort (or perhaps even a purer, better sort) of moral edification provided by religious seers or by fiery sermons without the lapse into superstition. That is, while traditional religious practices often depend on unjustified claims to knowledge, Kant's account of the sublime can withstand critique and can nevertheless still provide human beings with strong feelings about morality and their capacity to realize their humanity.

Kant describes his account of the sublime as if sublime feeling is a morally superior alternative to more traditional modes of religious experience. Kant asserts that his account of the sublime is superior to traditional, religious presentations of morality. He writes, "This pure, elevating, merely negative presentation of morality carries with it no risk of **visionary rapture**, which is **a delusion of being able to see something beyond all bounds of sensibility**, i.e., to dream in accordance with principles (to rave with reason), precisely because the presentation in this case is merely negative."³⁹³ Here, Kant suggests that his sort of sublime experience is a presentation of morality, but one that does not carry with it unsound claims about what one "sees." In other words, a

³⁹³ CPJ 5:275.

Kantian sublime experience will not devolve into the fantasies a Swedenborgian mystical experience.

At this point, though, one might wonder if Kant's alternative will be disappointing to people who previously have relied upon more fanciful and sensuous experiences when seeking moral edification. Does Kant's account leave the subject feeling a cold, dull chill where the thrall of enthusiasm had previously left them aflame with feeling? Kant seems to anticipate this criticism that one could levy against his account by a religious enthusiast when he writes, "It is utterly mistaken to worry that if it [the representation of the moral law and the predisposition to morality in us] were deprived of everything that the senses can recommend it would then bring with it nothing but cold, lifeless approval and no moving force or emotion."³⁹⁴ In fact, Kant argues that his account is actually more powerful than sermons or visions which attach a sensory element to them (i.e. *seeing* God or *hearing* the word of God). Emotion will run so high that we must take pains to assure ourselves that we are not swept away into enthusiasm. He writes, "For when the senses no longer see anything before them, yet the unmistakable and inextinguishable idea of morality remains, there it would be more necessary to moderate the momentum of an unbounded imagination so as not to let it reach the point of enthusiasm, rather than, from fear of powerlessness of these ideas, to look for assistance for them in images and childish devices."³⁹⁵

³⁹⁴ CPJ 5:274.

³⁹⁵ CPJ 5:272.

Kant's account of sublimity, therefore, can provide edification for a mind that has overcome an unseemly dependence on images and "childish" heuristic devices in its thinking about morality. Sublime objects in nature, such as thunderbolts, storms, cliffs, and mountains no longer are *literally* manifestations of an anthropomorphic God, but rather are seen for what they are: natural objects which, when contemplated, can give rise to an awareness of our capacity for freedom.

V. Sublimity and Moral Precariousness

Since our capacity for freedom consists in our ability to choose our own ends (and not to be determined solely by instinct), Kant's account of sublimity dramatizes, through an aesthetic experience, the conflict between morality and self-love. The sublime is an intense experience where the mind is moved by awareness of the moral law (and its applicability to us as moral agents). Kant's account of the sublime provides an outlet for inflamed, passionate experience without reverting to enthusiasm and fanaticism. Sublimity, unlike enthusiasm and fanaticism, retains the inextricability of moral crisis, while also allowing us to feel ourselves as subject to the moral law (and therefore elevated above nature). On this understanding, Kant's account of the sublime can be an authentic moral experience. The fanatic and enthusiast, by contrast, are actually more prone to inauthentic moral experience and, especially in the case of the fanatic, to immoral behavior. That is, enthusiasm and fanaticism distort the relationship between human beings and the moral law. In doing so, it obscures the inextricably complex and problematic connection human beings have to the moral law.

In the *Critique of the Power of Judgment*, Kant defines enthusiasm in the following way: “The idea of the good with affect is called **enthusiasm**.”³⁹⁶ When a person experiences an intense feeling on account of thinking about the good, that person is in an enthusiastic mindset. Enthusiasm, for Kant, is morally problematic because an action conducted in an enthusiastic mindset is likely not grounded in reason, but rather in feeling. Kant writes, “Now, however, every affect is blind, either in the choice of its end, or, even, if this is given by reason, in its implementation; for it is that movement of the mind that makes it incapable of engaging in free consideration of principles, in order to determine itself in accordance with them. Thus it cannot in any way merit a satisfaction of reason.”³⁹⁷ Enthusiasm can distort our thought process and make us act in a way that reflects our enlivened mindset, but such action would be grounded in feeling, and not reason. Authentically moral action (action from awareness of one’s duty) only comes about through the exercise of one’s practical reason; enthusiasm makes thinking more difficult, if not impossible.

Enthusiasm, while problematic on its own, can lead to states of mind which are even more threatening to morality. In the *Critique of the Power of Judgment*, when comparing sublimity with enthusiasm, Kant warns against enthusiasm because it can lead to “visionary rapture” and the “delusion” that one has the ability to “see” beyond “all bounds of sensibility.”³⁹⁸ Such “visionary rapture” is a more intense, and also more dangerous, passion. Kant writes, “In enthusiasm, as an affect, the imagination is

³⁹⁶ CPJ 5:272.

³⁹⁷ CPJ 5:272.

³⁹⁸ CPJ 5:274.

unreined; in visionary rapture, as a deep-rooted, oppressive passion, it is unruled. The former is a passing accident, which occasionally affects the most healthy understanding; the latter is a disease that destroys it.”³⁹⁹ Enthusiasm, on my reading, is problematic not only in itself, but also because it tends towards fanaticism. The enthusiast is at high risk of becoming a fanatic.

The rapture of the fanatic often has its origins in the intense emotion of the enthusiast. Rachel Zuckert refers to this sort of rapture as “fanaticism” in her paper “Kant’s Account of Practical Fanaticism.”⁴⁰⁰ There, she describes the dangers of fanaticism as follows: “First, the practical fanatic confuses emotional sensibility and sensation: he misinterprets a feeling of his own state (pleasure, or being moved by religious or moral ideas) for a sense experience of or received from the *object* of those ideas...Second...the fanatic understands an effect of the moral, religious articulation of ends (to be brought about by him) itself as an end, indeed as the ultimate accomplishment of his highest aims.”⁴⁰¹ The first problem of fanaticism is a violation one of the basic tenants of Kant’s theoretical philosophy: a claim to experience of a mere idea. Kant expresses his concern for this first problem in the above-quoted remark where he writes that the so-called visionary is under a “**delusion**” in which he or she claims the ability to “**see something beyond all bounds of sensibility.**”⁴⁰²

³⁹⁹ CPJ 5:275.

⁴⁰⁰ Rachel Zuckert, “Kant’s Account of Practical Fanaticism” (pp. 291-318).

⁴⁰¹ Rachel Zuckert, “Kant’s Account of Practical Fanaticism” (p. 311).

⁴⁰² CPJ 5:275.

The second problem of fanaticism involves a fundamentally mistaken view on the moral situation of human beings. Human beings, if they are to possess moral worth, must actually achieve a good will. In the *Religion*, Kant puts this a bit differently, but it is the same line of thought, put in a religious context: “To begin with I accept the following proposition as a principle requiring no proof: *Apart from a good life-conduct, anything which the human being supposes that he can do to become well-pleasing to God is mere religious delusion and counterfeit service to God.*”⁴⁰³ In the case of visionary rapture and fanaticism, there is an acute risk that the mere presence of fervor and intensity in someone is sufficient to make oneself pleasing to god (i.e. is sufficient to make one good). Such a view badly obfuscates the basic relation that human beings have to morality, and to the way in which human beings can become good.

The sublime gives rational people an outlet for the joyful or exhilarating sense of one’s capacity for morality, and thus of one’s capacity to rise above all of nature (in all its vastness and power). The sublime does not yield an image, or a positive representation of our freedom and our capacity for morality. Such an image is impossible. Kant writes, “For the **inscrutability of the idea of freedom** entirely precludes any positive presentation; but the moral law is sufficient in itself in us and originally determining, so that it does not even allow us to look around for a determining ground outside of it.”⁴⁰⁴ Our awareness of the moral law, coupled with a recognition of its validity prepares us for a sublime experience. Without awareness of the moral law, sublimity is impossible. When we do encounter the sublime, we experience a heightened awareness of the moral

⁴⁰³ Rel. 6:171.

⁴⁰⁴ CPJ 5:275.

law because we come to feel (in a way we previously understood) ourselves as elevated above the vast or powerful natural object before us. This presentation of our freedom remains negative because we do not “see” the moral law, or our own freedom, or any kind of deity (the source of all goodness according to religious traditions). While merely negative in its presentation of our freedom, Kant writes that the sublime, in its wake, can “leave behind a disposition of mind that, even if only indirectly, has influence on consciousness of its strength and resolution in regard to that which brings with it intellectual purposiveness (the supersensible).”⁴⁰⁵ Such a disposition does not confuse feeling for authentic moral experience, and it does not lead toward fanatical identification with the good itself or with a privileged and direct access to the supersensible. It leaves us where we started, only with a heightened awareness of what we already knew: that we are capable of morality, if we can only prioritize the moral law over self-love in our supreme maxim (expressed through our actions and their accompanying subordinate maxims).

⁴⁰⁵ CPJ 5:272.

CHAPTER 5

Becoming Good: Moral Evil, Freedom, and Grace in Kant's Moral Philosophy

Introduction

This chapter considers Kant's views on moral evil and on the possibility for moral improvement. Part I provides a brief overview of Kant's basic conception of evil and his claim that evil is a pervasive trait in human beings. Part II provides an account of Kant's notion of intelligible freedom. This section provides the basic framework through which I will analyze Kant's views on moral evil and the possibility of moral improvement. Part III discusses the relationship between freedom and appearances. Here, I develop Kant's claim that, while metaphysically distinct from each other, the empirical and the intelligible have a relationship best understood as one of ground and consequence. Part IV introduces Kant's conception of moral evil. It examines Kant's claim that the basic moral struggle in which human beings find themselves is between competing principles: the principle of self-love and the principle of morality. Part V considers the important notion of the inscrutability of the ultimate ground of moral evil in human beings. Here, I argue that the notion of inscrutability turns largely on Kant's theory of freedom. Part VI begins my discussion of Kant's account of moral improvement. Part VII centers on the relationship between Kant's account of moral improvement and his views on the traditional religious concept of grace. Here I argue that Kant's account of moral improvement is not simply a reformulation of the notion of grace, but it nevertheless bears some striking similarities to it. Finally, in Part VIII, I consider two possible

objections to Kant's account of moral improvement. I offer a qualified defense of Kant against the first objection, but then argue that the second objection brings out a serious issue in Kant's account of moral improvement. I end the chapter by arguing first that the basis on which the objection rests is a foundational feature of Kant's moral theory (namely, intelligible freedom), and then by responding to the claim that Kant's possible failure to account for intelligible moral improvement engenders moral despair.

I. The Pervasiveness of Evil

Kant defines moral evil in human beings in the following way: “[T]he statement, “The human being is *evil*,” cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it.”⁴⁰⁶ Human beings are evil insofar as they are aware of the requirements of morality, but nevertheless act in ways that fail to meet those requirements. More specifically we should think of human evil as residing in the kinds of maxims we adopt: “We call a human being evil, however, not because he performs actions that are evil (contrary to law), but because these are so constituted that they allow the inference of evil maxims in him.”⁴⁰⁷ An immoral, or evil, maxim places self-love above morality as providing the sufficient determining ground for exercising one's power of choice. That is, evil maxims privilege self-love over morality. In cases where the two come into conflict, the evil person, at times, will choose to satisfy an inclination rather than fulfill his or her moral obligation. In the *Religion*, Kant writes, “In order, then, to call a human being evil, it must be

⁴⁰⁶ Rel. 6:32.

⁴⁰⁷ Rel. 6:20.

possible to infer *a priori* from a number of consciously evil actions, or even from a single one, an underlying evil maxim...”⁴⁰⁸

Evil is an especially pressing issue to consider because Kant claims that immorality, rather than morality, is the most prominent disposition in most people. In the *Groundwork*, Kant writes, “One need not be an enemy of virtue but only a cool observer, who does not take the liveliest wish for the good straight-away as its reality, to become doubtful at certain moments (especially with increasing years, when experience has made one’s judgment partly more shrewd and partly more acute in observation) whether any true virtue is to be found in the world.”⁴⁰⁹ In the *Religion*, Kant writes, “We can spare ourselves the formal proof that there must be such a corrupt propensity rooted in the human being, in view of the multitude of woeful examples that the experience of human *deeds* parades before us.”⁴¹⁰ Kant goes on to describe some of the manifestations of moral evil that even so-called “civilized” people exhibit: “If we are however disposed to the opinion that we can have a better cognition of human nature known in its civilized state...we must then hear out a long melancholy litany of charges against humankind—of secret falsity even in the most intimate friendship...of a hearty goodwill that nonetheless admits the remark that “in the misfortunes of our best friends there is something that does not altogether displease us”; and of many other vices yet hidden under the appearance of virtue, let alone those of which no secret is made...”⁴¹¹ Evil is widespread; in fact, Kant seems here to suggest that moral evil is the rule rather than the exception in human

⁴⁰⁸ Rel. 6:20.

⁴⁰⁹ G 4:407.

⁴¹⁰ Rel. 6:33.

⁴¹¹ *Ibid.*

affairs. Thinking about the moral situation of human beings, then, seems to require taking this unhappy fact into account. In other words, given the pervasiveness of evil in human life, it merits careful consideration, particularly how it relates to Kant's broader moral philosophy.

II. Freedom in the Intelligible World

I would argue that any discussion of moral goodness or moral evil, ultimately, should be grounded in Kant's account of freedom. Since moral evil is the result of our freely chosen actions, it seems natural to begin with Kant's account of freedom. Kant's theory of freedom is obviously complicated, but for our purposes, we need focus only on two aspects of it: 1) on Kant's insistence that freedom cannot reside in appearances themselves, and 2) Kant's derivation of freedom from moral obligation.

Freedom and Appearances

In the first *Critique*, Kant famously claims that it is a mistake to take appearances for things in themselves. Various sorts of errors arise from conflating appearances and things-in-themselves, both theoretical mistakes and, crucially for our purposes, mistakes regarding practical philosophy, especially about the possibility of freedom. In the first *Critique*, Kant writes, "For if appearances are things in themselves, then freedom cannot be saved."⁴¹² Why? Because if appearances were also things-in-themselves, and not merely objects of experience, then all things could be completely explained in terms of the mechanistic laws of nature. In the B Preface to the first *Critique*, Kant writes, "Now if

⁴¹² CPR A536/B564.

we were to assume that the distinction between things as objects of experience and the very same things as things in themselves, which our critique has made necessary, were not made at all, then the principle of causality, and hence the mechanism of nature in determining causality, would be valid of all things in general as efficient causes.”⁴¹³

On Kant’s view, this mechanistic framework provides one fundamental way of thinking about causality. But, with the distinction between appearances and things-in-themselves in place, another kind of causality is possible. Kant writes, “In respect to what happens, one can think of causality in only two ways: either according to **nature** or from **freedom**.”⁴¹⁴

Importantly then, Kant argues that the idea of freedom can be preserved if we make the proper distinction between appearances and things-in-themselves. Insofar as we are natural beings, we are “subject to natural necessity” and hence “not free.”⁴¹⁵ But, with the distinction in mind, we can think of ourselves as free. Kant writes, “Now although I cannot **cognize** my soul, considered from the latter side [as subject to the law of nature], through any speculative reason (still less through empirical observation), and hence I cannot **cognize** freedom as a property of any being to which I ascribe effects in the world of sense...nevertheless, I can **think** freedom to myself...so long as our critical distinction prevails between the two ways of representing (sensible and intellectual)...”⁴¹⁶ Kant’s aim in the first *Critique* is to show that freedom is possible (that we can at least think of

⁴¹³ CPR Bxxvii.

⁴¹⁴ CPR A532/B560.

⁴¹⁵ CPR Bxxvii.

⁴¹⁶ CPR Bxxviii.

ourselves as free). Kant draws a contrast between cognizing and thinking: “To **cognize** an object, it is required that I be able to prove its possibility (whether by the testimony of experience from its actuality or *a priori* through reason). But I can **think** whatever I like, as long as I do not contradict myself, i.e., as long as my concept is a possible thought, even if I cannot give any assurance whether or not there is a corresponding object somewhere within the sum total of all possibilities.”⁴¹⁷

In the second *Critique*, Kant tries to show that freedom is not only thinkable, but also that there is in fact an object to which that concept corresponds. We do not cognize our freedom directly, like we do ordinary objects of experience. Kant writes, that we can “neither be immediately conscious of this [freedom]...nor can we conclude to it from experience, since experience lets us cognize only the law of appearances and hence the mechanism of nature, the direct opposite of freedom.”⁴¹⁸ The law of appearances-- mechanistic determinism-- is the “direct opposite of freedom” because any change that occurs, any event that takes place, can be explained solely in terms of a prior state of affairs and the laws of nature. Therefore, anything we, understood only as natural beings, may do will be caused, or determined by, something outside of ourselves (natural laws). How, then, do we arrive at freedom? Freedom, on this line of thinking, is implied by the objective validity of the moral law with regard to our own wills. That is, Kant proves freedom indirectly, by arguing that we know for a fact that morality applies to us, that we have moral obligations. Kant writes, “It is therefore the *moral law*, of which we become immediately conscious (as soon as we draw maxims of the will for ourselves), the *first*

⁴¹⁷ CPR Bxxvi.

⁴¹⁸ CPrR 5:29.

offers itself to us and, inasmuch as reason presents it as a determining ground not to be outweighed by any sensible conditions and indeed quite independent of them, leads directly to the concept of freedom.”⁴¹⁹ Kant here is claiming that as soon as we become aware that, in acting in an intentional way, we are also acting on some volitional principle (and therefore not merely from uncritical, unthinking instinct), we recognize that there are under certain obligations that are absolutely necessary and cannot be overridden by an appeal to “sensible conditions.”

Freedom and Inclination

Perhaps the most pressing and morally relevant “sensible condition” beings like us have to contend with are inclinations. Inclinations are subjective, sensory desires that we often feel impelled to satisfy. Kant writes, “The dependence of the faculty of desire upon feelings is called inclination...”⁴²⁰ Kant holds that freedom in part consists in independence from the dictates of inclination. In the *Critique of Pure Reason*, Kant writes, “**Freedom in the practical sense** is the independence of the power of choice from the **necessitation** by impulses of sensibility. For a power of choice is **sensible** insofar as it is **pathologically affected** (through moving-causing of sensibility); it is called an **animal** power of choice (*arbitrium brutum*) if it can be **pathologically necessitated**.”⁴²¹ He goes on to claim that “The human power of choice” differs from the animal power of

⁴¹⁹ CPrR 5:30.

⁴²⁰ G 4:414.

⁴²¹ CPR A534/B562.

choice because, in the case of human choice, “there is a faculty of determining oneself from oneself, independently of necessitation by sensible impulses.”⁴²²

Without this faculty for “determining oneself from oneself,” freedom would be impossible. Kant takes this as a rather obvious point: “It is easy to see that if all causality in the world of sense were mere nature, then every occurrence would be determined in time by another in accord with necessary laws.”⁴²³ A person’s deeds (and misdeeds) would simply be, on this picture, necessary outcomes, caused by occurrences in accord with necessary (natural) laws. Freedom, by contrast, attaches a causality to our power of choice that can place its origin entirely within our own faculty of choice. Kant writes that practical freedom (our power of choice) “presupposes that although something has not happened, it nevertheless **ought** to have happened, and its cause in appearance was thus not so determining that there is not a causality in our power of choice such that, independently of those natural causes and even opposed to their power and influence, it might produce something determined in the temporal order in accord with empirical laws, and hence begin a series of occurrences **entirely from itself**.”⁴²⁴ Free actions, according to Kant, do precisely this; they set aside completely the determining influence of natural desires and pressures, and are chosen in accordance with the dictates of the moral law (of which we have immediate awareness).⁴²⁵

⁴²² *Ibid.*

⁴²³ CPR A534/B562.

⁴²⁴ *Ibid.*

⁴²⁵ In the famous concluding section of the Critique of Practical Reason, Kant describes the moral law as residing “within me [in mir]” (CPrR 5:162).

Fact of reason

We have seen that freedom must involve a faculty for overriding the influence of natural inclinations in favor of the requirements of morality. Kant claims that we do in fact have the capacity to resist any and all of our inclinations, and that we are responsible for our actions when we fail to (when morality requires that we do so). Kant, in the second *Critique*, seems to take this to be an unproblematic and obvious claim. He makes this point first through an imaginary example:

Suppose someone asserts of his lustful inclination that, when the desired object and the opportunity are present, it is quite irresistible to him; ask him whether, if a gallows were erected in front of the house where he finds this opportunity and he would be hanged on it immediately after gratifying his lust, he would not then control his inclination. One need not conjecture very long what he would reply. But ask him whether, if his prince demanded, on pain of the same immediate execution, that he give false testimony against an honorable man whom the prince would like to destroy under a plausible pretext, he would consider it possible to overcome his love of life, however great it may be... He judges, therefore, that he can do something because he is aware that he ought to do it and cognizes freedom within him, which, without the moral law, would have remained unknown to him.⁴²⁶

In the former case, the person is choosing in accordance with the principle of self-love. Kant assumes here that the desire to stay alive is our most basic and unshakeable inclination, so we will satisfy that physical desire above all others. In the latter case, though, we recognize, through the example, that it is sometimes required of us to set aside even this most fundamental inclination and instead act in a way that morality

⁴²⁶ CPrR 5:30.

requires. Additionally, we also recognize in this example that we are indeed capable of doing so.

After providing this example, Kant describes this realization as a “fact of reason.” He writes, “Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical...”⁴²⁷ Our consciousness of morality, and with it consciousness of our moral obligation to act according to its requirements, is an irreducible fact, not derived from something more basic than this brute awareness itself.

This fact of reason is evident in our practice of imputing responsibility onto others for their actions. When determining the moral worth of an individual who has committed a misdeed, we may begin with an investigation into “the sources of the person’s empirical character, seeking them in a bad upbringing, bad company, and also finding them in the wickedness of a natural temper insensitive to shame, partly in carelessness and thoughtlessness.”⁴²⁸ Nevertheless, we still place blame on the person for his actions: “Now even if one believes the action to be determined by these causes, one nonetheless blames the agent, and not on account of his unhappy natural temper, not on account of the circumstances influencing him, not even on account of the life he has led previously; for one presupposes that it can be entirely set aside how that life was constituted, and that the

⁴²⁷ CPrR 5:31.

⁴²⁸ CPR A555/B583.

series of conditions that transpired might not have been, but rather that this deed could be regarded as entirely unconditioned in regard to the previous state, as though with that act the agent started a series of consequences entirely from himself.”⁴²⁹

From the reality of moral obligation, Kant reasons that we must posit freedom, the faculty of choice that allows us to determine our will independent of natural laws (which determine the content of our inclinations). Without freedom, moral imputation would not be possible.

Since we know that there are moral obligations applying to us, and that this fact implies freedom, we might well ask: Where or what is the realm of freedom? The realm of freedom cannot be the world of appearances. This is a view Kant had expressed in the years leading up to the first *Critique*. In a note on metaphysics, dated from the mid-to-late 1770’s, Kant writes, “Reason is not in the chain of appearances and is with regard to all of that free with regard to its own causality (the actions of reason themselves are also not appearances, only their effects are).”⁴³⁰ Freedom must exist in something other than appearances; it cannot be a property of objects insofar as they exist in the sensible world (i.e. of merely sensible objects), but rather of what Kant calls “intelligible” objects. In the first *Critique*, Kant provides a definition of intelligible objects in the following manner: “I call **intelligible** that in an object of sense which is not itself appearance.” As we saw earlier, Kant argues that it is admissible to *think* of anything so long as one does not run into contradiction in doing so, and that we are free to think of ourselves as free (so long

⁴²⁹ CPR A555/B583.

⁴³⁰ Notes on metaphysics 18:252.

as we distinguish between things in themselves and appearances). We can therefore think of ourselves as subjects as having both a sensible or empirical character (as an appearance) and an intelligible character (that is us, as an object, that is not itself appearance).

III. The Relationship Between Freedom and Phenomena

According to Kant, we can think of ourselves as members of both the sensible world of appearances (or the phenomenal world), and also as members of an intelligible (or noumenal) world. In the *Groundwork* Kant describes this possibility as follows:

Since he does not as it were create himself and does not get his concept a priori but empirically, it is natural that he can obtain information even about himself only through inner sense and so only through the appearance of his nature and the way in which his consciousness is affected—although beyond this constitution of his own subject, made up of nothing but appearances, he must necessarily assume something else lying at their basis, namely his ego as it may be constituted in itself; and thus as regards mere perception and receptivity to sensations he must count himself as belonging to the *world of sense*, but with regard to what there may be of pure activity in him...he must count himself as belonging to the *intellectual world*, of which however he has no further cognizance.⁴³¹

I claim that there is an important consequence of this line of thinking: the free choice, the choice which determines the order of priority between morality and self-love, is not made by the empirical self, but rather by the intelligible self. I mean by “intelligible self” what Kant seems to be describing in *Ground III* when he writes: “As a rational being, and thus a being belonging to the intelligible world, the human being can never think of the

⁴³¹ G 4:451.

causality of his own will otherwise than under the idea of freedom; for, independence from the determining causes of the world of sense (which reason must always ascribe to itself) is freedom.”⁴³² The intelligible self is the genuine ground of human freedom. Human freedom exists in the phenomenal world whenever we impute moral worth (or moral blame) to an agent.

Being free and being a member of the intelligible world are, for Kant, mutually entailing states of being: “[W]hen we think of ourselves as free we transfer ourselves into the world of understanding as members of it and cognize autonomy of the will along with its consequence, morality; but if we think of ourselves as put under obligation we regard ourselves as belonging to the world of sense and yet at the same time to the world of understanding.”⁴³³ In *Groundwork III*, Kant seems to take the former approach, trying to provide an independent proof of freedom. In the second *Critique*, though, Kant takes the latter approach. He first lays out the fact of reason, and then, from that fact, derives our freedom.

As we saw in an earlier quotation, Kant clearly argues that we can have “no further cognizance” of the intelligible world, or ourselves as intelligible agents, beyond merely knowing that we are in fact free, and that freedom needs to reside in the intelligible world. The intelligible world and the sensible world are radically different from each other. The sensible world is the realm of experience, of space and time, of

⁴³² G 4:453. This passage is cited merely as providing a definition of my term “intelligible self” by referring to a concept in Kant’s moral philosophy (that of “a being belonging to the intelligible world”).

⁴³³ G 4:453.

objects to which the categories of the understanding apply (etc.). The intelligible world is rather different; as Kant puts it in a note: “In the intelligible world nothing happens and nothing changes, and there the rule of causal connection disappears.”⁴³⁴

It might seem natural to suppose, given their radical heterogeneity, that the intelligible world and the sensible world have nothing to do with each other. This seems especially plausible when one thinks that the very idea of mechanistic causal connection does not apply to the intelligible world.

Kant presents a picture, though, not of radical rupture, but rather one of ground and consequence. In a section of the first *Critique* called “The possibility of causality through freedom unified with the universal law of natural necessity,” Kant writes:

[E]very effective cause must have a **character**, i.e., a law of its causality, without which it would not be a cause at all. And then for a subject of the world of sense we would have first an **empirical character**, through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws, from which, as their conditions, they could be derived; and thus, in combination with these other appearances, they would constitute members of a single series of the natural order. Yet second, one would also have to allow this subject an **intelligible character**, *through which it is indeed the cause of those actions as appearances*, but which does not stand under any conditions of sensibility and is not itself appearance. The first one could call the character of such a thing in appearance, the second its character as a thing in itself.⁴³⁵

⁴³⁴ Notes on metaphysics 18:254

⁴³⁵ CPR A539/B567, emphasis added.

Later in this section, Kant describes the empirical character of our actions as being “a mere appearance of the intelligible character...”⁴³⁶

The intelligible self has a determining influence over what happens with the empirical self. When a person, in the sensible world, has set aside the influence of inclinations and performed his or her duty, then we know that, in the intelligible world, the power of choice was exercised in a way that will lead him or her to privilege morality over self-love in the empirical world. The exact way in which this choice is made, though, is utterly unknowable. We have no access to the intelligible world; we only know that it must be the world of freedom, and hence of morality. Kant writes in a note: “In free actions reason has influence not merely as a comprehending but also as an effecting and driving *principium*. We have no insight into how it does not merely...judge, but fills the place of a natural cause, let alone how it is itself determined to action or omission by means of impulses.”⁴³⁷

The actions of the empirical self do, though, reveal something of the choice made by the intelligible self. The empirical effects manifested in the sensible world are a bit like a function of the intelligible choice. Kant seems to have something like this in mind when he writes, “The true activity of reason and its effect belong to the intelligible world (*mundo intelligibili*). Hence we do not know to what degree we should impute. Nevertheless we know that the influencing power of reason is not determined and necessitated by any *phenomena*, rather that it is free, and (in the case of imputation) we

⁴³⁶ CPR A541/B569.

⁴³⁷ Notes on metaphysics 18:253.

judge the action merely in accordance with rational laws. The actions here in the world are mere *Schemata* of the intelligible [actions]; yet these appearances (this world already signifies “schema”) are still interconnected in accordance with empirical laws, even if one regards reason itself, in accordance with its expressions, as a *phaenomenon* (of the character). But what the cause of this may be we do not discover in *phaenomenis*.⁴³⁸

This intriguing passage seems to suggest the following: the intelligible choice is not directly apparent in the phenomenal world (i.e. we have no intuition of the choice itself through our inner or outer senses). We do, though, see its effects in phenomena (namely, in the actions to which we can assign moral blame or (potential) praise). So, the intelligible choice determines outcomes in the phenomenal world, but, at the same time, all actions in the phenomenal world are describable in terms of natural laws (even morally good actions in the phenomenal world are consistent with the laws of nature). Nevertheless, insofar as we are able to render moral judgments, some of what we observe in the phenomenal world must also be explainable in terms of a free, noumenal choice.

IV. Radical Evil: Freedom and Self-Love

In the empirical world, the moral struggle in human beings consists in the opposition sometimes experienced between the need to carry out our moral obligations and the desire we have to satisfy our inclinations (to achieve happiness). When a human being chooses to satisfy a desire even when doing so would violate the moral law, he or she is privileging self-love over morality and is therefore evil. There are two competing sets of basic volitional principles: the principle of morality and the principle of self-love.

⁴³⁸ Notes on metaphysics 18:254

Achieving morality in the empirical world is not a matter of extirpating one principle and retaining the other (even very bad people retain some predisposition to the good, and even morally good people are naturally impelled to seek happiness), but rather one of ordering and priority. Kant writes, “The difference, whether the human being is good or evil, must not lie in the difference between the incentives he incorporates into his maxim (not in the material of the maxim) but in their *subordination* (in the form of the maxim): *which of the two he makes the condition of the other.*”⁴³⁹ The morally evil person acts on maxims of the following form: “I will act in accordance with morality so long as doing so does not clash too severely with my inclinations.” The morally good person, by contrast, forms a maxim like the following: “I will seek to satisfy my inclinations when and only when doing so will not interfere with fulfilling my moral obligations.”⁴⁴⁰ Immoral actions, therefore, are grounded in the principle of self-love, and, in particular, in the privileging of self-love over the moral law. In the *Religion* Kant claims that, “self-love...is precisely the source of all evil.”⁴⁴¹ The source of all moral goodness, by contrast, lies in placing morality, and its requirements, as sufficient in determining the will, without acquiescing to the pleadings of self-love.

So within this basic framework, we can understand Kant’s definition of morally meritorious action (or action from duty) given in *Groundwork I*: “Now, an action from duty is to put aside entirely the influence of inclination and with it every object of the will; hence there is left for the will nothing that could determine it except objectively the

⁴³⁹ Rel. 6:36.

⁴⁴⁰ For a similar interpretation of the difference between morally good and morally human beings see: Christine M. Korsgaard *Creating the Kingdom of Ends* p. 165.

⁴⁴¹ Rel. 6:46.

law and subjectively *pure respect* for this practical law, and so the maxim of complying with such a law even if it infringes upon all my inclinations.”⁴⁴² Satisfying one’s inclinations is the basic activity that derives from the principle of self-love, so in choosing morality over self-love as the most fundamental volitional principle, one only acts to satisfy those inclinations that would not result in violations of the moral law. Since human beings always must act on the basis of some kind of motive, in cases where there is not a single pathological inclination that would explain why the person acted in the way he or she did, then all that would be left is respect for the moral law. Respect for the moral law, in other words, is not a characteristic of the intelligible self (we cannot know anything about the intelligible self beside the fact that it exists and it is free), but rather of the phenomenal or empirical self. Respect for the moral law always remains, even in the case of immoral persons, a potential motive for action; it just seems to be the case that the vast majority of people (perhaps everyone) does not in fact act on this motive, but rather on some other inclination (or set of inclinations).

Human beings, on this account, are interminably torn between these two sets of volitional principles. For Kant, so long as one is human, one is potentially subject to inclinations that, if acted upon, would pull us away from our duties. At the same time, however, human beings will always be aware of (and, in the case of immoral agents, tasked by) the moral law and its binding force. In the *Religion* Kant describes this ceaseless back-and-forth in the following manner:

⁴⁴² G 4:400-401.

The human being (even the worst) does not repudiate the moral law, whatever his maxims, in rebellious attitude (by revoking obedience to it). The law rather imposes itself irresistibly, because of his moral predisposition; and if no other incentive were at work against it, he would also incorporate it into his supreme maxim as sufficient determination of his power of choice, i.e. he would be morally good. He is, however, also dependent on the incentives of his sensuous nature because of his equally innocent natural predisposition, and he incorporates them too into his maxim (according to the subjective principle of self-love). If he took them into his maxim *as of themselves sufficient* for the determination of his power of choice, without minding the moral law (which he nonetheless has within himself), he would then become morally evil.⁴⁴³

This is a struggle that has no resolution in the life of human beings. Human beings cannot possess holy wills (a will in which there is no clash between inclination and morality).

The basic structure of the moral life is identical for all people. The details may vary (the kinds of inclinations one has, the strength of those inclinations), but there is always the potential for, and often actual, struggle for volitional priority between self-love and the moral law.

V. Inscrutability and Radical Evil

Whether or not a person fulfills the requirements of morality in the empirical world, is, as we have seen, a function (or effect) of the intelligible exercise of free choice. In the *Religion*, Kant writes, “The human being must make or have made *himself* into whatever he is or should become in a moral sense, good or evil. These two [characters] must be an effect of his free power of choice, for otherwise they could not be imputed to

⁴⁴³ Rel. 6:36.

him and, consequently, he could be neither *morally* good nor evil.”⁴⁴⁴ People are morally responsible for their actions when their actions can be imputed to them. Without the ability to impute actions to people, we cannot have an account of moral goodness or evil. In the *Metaphysics of Morals*, Kant defines moral imputation in the following way: “*Imputation (imputatio)* in the moral sense is the *judgment* by which someone is regarded as the author (*causa libera*) of an action, which is called a *deed (factum)* and stands under laws.”⁴⁴⁵ The idea of an intelligible self, then, is the ground for the possibility of imputing deeds onto an individual person, and not some series of antecedent empirical state of affairs. That means, though, that in cases of human evil, we cannot impute the misdeed to the person as a natural being, but rather as an intelligible agent. Since we do not have experiential access to the intelligible world, we cannot, ultimately, know why a person, as an intelligible being, has chosen moral evil rather than goodness. The ground of moral evil (the faculty of choice) is, in Kant’s terms, *inscrutable*.

Kant seems to be working with a notion of inscrutability similar to that of Alexander Gottlieb Baumgarten, who defined a will as inscrutable when the “impelling causes are incomprehensible.”⁴⁴⁶ In the *Religion* Kant describes this inscrutability when he writes, “Whenever we...say, “The human being is by nature good, or, “He is by nature evil,” this only means that he holds within himself a first ground (to us inscrutable) for the adoption of good or evil (unlawful) maxims...”⁴⁴⁷ One might interject at this point and claim that this inscrutability is not as resounding as it might initially seem; since evil

⁴⁴⁴ Rel. 6:44.

⁴⁴⁵ MM 6:227.

⁴⁴⁶ Baumgarten, *Metaphysica* §900. (Cited in *Religion and Rational Theology* 479).

⁴⁴⁷ Rel. 6:21.

people are evil insofar as they privilege self-love over the demands of morality, then perhaps one could simply say that people are evil because they have such and such an inclination that they could not control when morality required that they do so. This reading of Kant's account of evil, however, is unwarranted. It seems to confuse the ground of moral evil with its effects. The free choice a person makes (as an intelligible being), is independent of and (metaphysically) prior to the normal sensible operations of inclinations. In a footnote from the *Religion*, Kant claims that we cannot find the ground of a free choice in terms of natural incentives.⁴⁴⁸ Evil in human beings is grounded on the intelligible character of one's choice, not in the choice's empirical effects. Kant writes, "Now, the ground of this evil cannot (1) be placed, as it is commonly done, in the sensuous nature of the human being, and in the natural inclinations originating from it."⁴⁴⁹ It is wrong to say that a robber, for instance, is evil because he had a desire for money; he is evil because, as an intelligible being, he exercised his power of choice in such a way that, in the empirical world, he would act to satisfy his desire for money, even when morality requires that he not act on that desire. Our explanation can go no further than this.

This lack of explanatory power in Kant's account of evil has led some commentators to fault Kant's theory of evil as being uninformative, and therefore unhelpful in explaining moral evil. One of the more forceful criticisms on this issue comes from Richard Bernstein. In his book *Radical Evil: A Philosophical Interrogation*, Bernstein writes, "Presumably, the introduction to the concept of radical evil is intended

⁴⁴⁸ Rel. 6:21n.

⁴⁴⁹ Rel. 6:34-5.

to explain *why* (from a practical point of view) we deviate from following the moral law. We do not always follow the moral law *because* as human beings, we have an innate propensity to evil. Our wills are corrupted at their root. But does this “because” really explain anything? Does it do any conceptual work? I do not think so... In short, radical evil—the alleged propensity to moral evil which is a universal characteristic of human beings—does not have *any* explanatory force (practical or theoretical) at all!”⁴⁵⁰ I would agree with Bernstein that Kant’s account of radical evil does not explain why human beings choose evil beyond the obvious claim that human beings are evil because they deviate from the requirements of morality, but would deny that this observation amounts to a significant objection to Kant’s account. Kant’s account of radical evil cannot, and does not presume to, provide an ultimate account for why human beings are evil.⁴⁵¹ It is crucial to his metaphysics that the ultimate “why” of human evil, the free choice, eludes our epistemological grasp.

In his paper “Evil Everywhere: The Ordinariness of Kantian Radical Evil,” Robert B. Louden makes a similar response to this kind of “explanatory emptiness” objection to Kant’s account of radical evil. In that paper, Louden defends Kant’s conception of evil in part by reaffirming Kant’s claim about the inscrutability of moral evil’s ultimate ground. He writes, “For Kant is quite clear in stating that his doctrine of radical evil is in no way

⁴⁵⁰ Richard Bernstein. *Radical Evil: A Philosophical Interrogation*. Blackwell: Malden, MA, 2002, p. 61.

⁴⁵¹ This point is also made by Robert B. Louden when he writes, “For Kant is quite clear in stating that his doctrine of radical evil is in no way intended to explain why human beings choose to adopt evil maxims” (Louden 97).

intended to explain *why* human beings choose to adopt evil maxims.”⁴⁵² Louden even claims that this is a positive aspect of Kant’s account: “Kant’s position regarding the ultimate inscrutability of human motives is a strength rather than a weakness in his doctrine of radical evil.”⁴⁵³

I agree with Louden’s claim that it is a mistake to fault Kant for not providing an explanation for why human beings are evil, but not necessarily with how he argues for that claim. In the previous quotation I provided from Louden, he claims that “human motives” are inscrutable, and that fact informs Kant’s claim that the ground of human evil is inscrutable. This tendency to treat human motives as inscrutable continues when Louden refers to a passage in the Religion where Kant writes “the depths of his [any person’s] own heart (the subjective first ground of his maxims) are to him inscrutable.”⁴⁵⁴ Additionally, Louden makes a similar claim about other people in general: “Our assessment of others’ motives and acts is always fallible.”⁴⁵⁵ It is problematic, though, to ground Kant’s claims about the inscrutability of human evil in claims about the inscrutability of a person’s motives because motives operate in the world of sensibility, and not in the intelligible world. I would argue that we should distinguish between two different kinds of inscrutability: 1) Empirical inscrutability and 2) Intelligible inscrutability. Louden’s response to Bernstein would be more compelling if

⁴⁵² Robert B. Louden, “Evil Everywhere The Ordinarity of Kantian Radical Evil: in Kant’s Anatomy of Evil, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010, p. 97.

⁴⁵³ Louden, p. 98.

⁴⁵⁴ Louden, p. 97.

⁴⁵⁵ *Ibid.*

he were more sensitive to this distinction. Let me briefly discuss each of these, beginning with empirical inscrutability.

In the *Groundwork*, Kant famously makes the claim that we can never be certain that any single act we commit was truly done out of respect for the moral law:

It is indeed sometimes the case that with the keenest self-examination we find nothing besides the moral ground of duty that could have been powerful enough to move us to this or that good action and to so great a sacrifice; but from this it cannot be inferred with certainty that no covert impulse of self-love, under the mere pretense of that idea, was not actually the real determining cause of the will; for we like to flatter ourselves by falsely attributing to ourselves a nobler motive, whereas in fact we can never, even by the most strenuous self-examination, get entirely behind our covert incentives, since, when moral worth is at issue, what counts is not actions, which one sees, but those inner principles of actions that one does not see.⁴⁵⁶

Here, Kant is making an empirical claim about our own inability to plumb completely the depths of our motives. Perhaps behind the good act of assisting those in need there is really the desire for praise and esteem, or perhaps the desire to satisfy an inclination to feelings of self-righteousness and of superiority over others. While the action is still rightly counted as good, it would, if motivated by those sorts of inclinations, not count as *morally* good. In the *Religion*, Kant makes a similar claim when he writes that the “depths of” a person’s heart “are to him inscrutable.”⁴⁵⁷ Also in the *Metaphysics of Morals*, Kant writes, “For a human being cannot see into the depths of his own heart so as to be quite certain, in even a *single* action, of the purity of his moral intention and the

⁴⁵⁶ G 4:407.

⁴⁵⁷ Rel. 6:51.

sincerity of his disposition, even when he has no doubt about the legality of the action.”⁴⁵⁸ Assisting those in need can certainly be a morally legal option, we can be sure of that, but we cannot be sure that, when we are assisting someone, we are acting on a purely moral motive. The inscrutability of our authentic motives described in these passages are about the impossibility of being sure that some sensible inclination is not actually at the root of our actions. It is empirical because, ultimately, these sorts of uncertainties derive from the limits of our capacity for empirical introspection.

Importantly, this kind of skepticism about motives applies also to cases in which a person violates the moral law. We can never be entirely sure why a person commits a misdeed. A robbery, for instance, might be motivated by all sorts of motives that may elude even the perpetrator’s understanding. Empirical inscrutability seems to be something that, at least in principle, is a tractable problem. Even though there are limits to our own capacity for introspection, we can imagine science (psychology) proceeding along and, eventually, being able to track down the actual motives on which a person acts. There must indeed be a particular motive on which any action rests, and, more often than not, it will be some inclination. This idea fits in nicely with Kant’s overall moral theory: Our actions are guided by either morality or self-love. When we are motivated by morality, we are acting out of respect for the moral law and fulfill our moral obligations; when we are guided by self-love, we are motivated by the principle of self-love and act on our desires (given to us through inclination). So, when someone does not fulfill his or her duty, then we can know that he or she was motivated by self-love. There must be

⁴⁵⁸ MM 6:392.

some inclination (or set of inclinations) that he or she is trying to satisfy, at the expense of moral duty. We may be able to investigate into the volitional structure of an agent and uncover the true motive on which he or she acted. Empirical inscrutability, in other words, might someday be overcome. Kant seems to suggest that this possibility exists when he writes, “[A]ll the actions of the human being in appearance are determined in accord with the order of nature by his empirical character and the other cooperating causes; and if we could investigate all the appearances of his power of choice down to their basis, then there would be no human action that we could not predict with certainty, and recognize as necessary given its preceding conditions.”⁴⁵⁹ In instances of evil actions, we can also make investigations into the particular types of inclinations that people seek to satisfy when they deviate from the requirements of morality. Kant himself engages in this sort of theorizing in the *Metaphysics of Morals* when he describes the desire to be superior to others as an inclination that leads to the moral offences of arrogance, defamation, and ridicule.⁴⁶⁰ Empirical inscrutability, therefore, is not in principle insurmountable. While our authentic psychological motivations for our actions may even resist our own, good faith introspective investigation, it remains possible that psychology may be able to provide information about our real motivations.

Empirical inscrutability, however, is not the most important kind of inscrutability, and even if we were to “solve” the problem of empirical inscrutability (whether through introspective, first-person or through scientific, third-person methods), the ground of human evil would remain inscrutable. In addition to empirical inscrutability there is also

⁴⁵⁹ CPR A550-1/B577-8.

⁴⁶⁰ MM 6:465-468.

the far more important issue of intelligible inscrutability. Intelligible inscrutability is the result of the metaphysical gap (or, put less dramatically, ‘distinction’) between the intelligible world and the sensible world. Since the realm of freedom is outside the world of appearances, we do not have access to the choice itself. The choice, though, must not be explainable in terms of an appeal to inclinations.

Louden’s response to Bernstein, as I have claimed, seems to conflate these two kinds of inscrutability. This leads, I think, to two problems: 1) It does not counter the initial objection as precisely as it could have and 2) It fails to do justice to the (understandable) concern that seems to inform Bernstein’s objection.

The “explanatory emptiness” objection claims that Kant cannot explain evil, and that this is a problem for Kant’s account of radical evil. Louden counters this claim by arguing that Kant’s account of radical evil does not purport to explain evil, and that fact is actually a strength of Kant’s view, not a fault. I agree with this basic response, but now, with the distinction between empirical and intelligible inscrutability in mind, we can make this point more forcefully. That is, making explicit the distinction between empirical and intelligible inscrutability makes clear why exactly the “explanatory emptiness” objection is misguided as a criticism of Kant’s discussion of moral evil in human beings.

Thinking in terms of intelligible inscrutability, the objection now seems even more off base because attempting to provide an explanation of radical evil would violate the limits imposed by transcendental idealism itself. Asking for an explanation would be

a request for knowledge about something of which we cannot claim to have knowledge at all (the intelligible world). Louden seems to be making a point similar to this when he writes, “In any given choice situation, we can hypothesize about the roles that various environmental and genetic factors may have played in leading a person to make the choice that he or she made, but ultimately no causal explanation is fully satisfactory, for the simple reason that the choice was free.”⁴⁶¹ Now we can see that, in this sentence, Louden is referring to empirical inscrutability in the beginning and then turns to intelligible inscrutability at the end. It seems that Louden is conflating these two senses of inscrutability. Investigations into “various environmental and genetic factors” do provide information that takes away some of our initial empirical inscrutability, but it has no bearing at all on the action’s intelligible inscrutability.

The distinction between empirical and intelligible inscrutability, I think, also does more justice to what I take to be the concern at the heart of Bernstein’s objection, which I take to be a desire to understand why people do wicked things. Kant’s account may seem to fail to go beyond any account of evil beside the claim that human beings are evil because they knowingly deviate from the requirements of morality and instead act in accordance with the principle of self-love.

Louden suggests that Kant does not attempt to lessen what I am calling empirical inscrutability of human evil. Louden writes, “However, given the indecipherable character of much human action, perhaps it is best not to speak presumptuously about why people commit evil. Those who think they have succeeded in descending into the

⁴⁶¹ Louden, p. 97.

depths here are often mistaken. Unfortunately, it is very difficult to ever reach bottom in this particular line of work, for the depths of human evil are unfathomable.”⁴⁶² While it is certainly a good idea not to “speak presumptuously about why people commit evil,” it nevertheless seems legitimate to inquire into people’s empirical motives to try to fill out our understanding of human behavior (if, for nothing else, perhaps as a way of yielding some practical advice into how we might go about lessening evil in the world). Louden’s criticism of Bernstein seems to suggest that it is not useful or worthwhile to make this kind of inquiry *because* of the inexorableness of empirical inscrutability. I have already suggested, though, that empirical inscrutability is something that can perhaps someday be overcome. We can, with the distinction in mind, make significant progress in understanding motives and incentives. After all, in the case of evildoers, there is *in fact* some incentive on which they are acting. An empirical investigation into the relationship between inclinations and behavior might well yield interesting (albeit perhaps also distressing) results. We should not, though, make the mistaken claim that in improving our naturalistic explanation of people’s behavior, we have made any progress whatsoever in overcoming intelligible inscrutability. Even with a perfected science of human behavior, the inscrutability of moral evil will remain as troubling and perplexing as ever.

VI. Moral Improvement

Kant’s moral theory makes the following claims: 1) Human beings are subject to morality (fact of reason); 2) Most, if not all, human beings are evil (the prevalence of self-love); 3) Human beings, even the worst, take an interest in the moral law that, at the

⁴⁶² Louden, p. 103.

very least, manifests itself as a wish to be morally good. A natural question that arises, then, is the following: How is it that an immoral person can become good? If I find myself habitually and occasionally deviating from the moral law, and, because of my original disposition to the good, I wish I could be moral, what hope can I have regarding the potential for moral improvement? This section examines Kant's account of moral improvement. Taking into account our earlier discussion of freedom and the inscrutability of the ground of moral evil, I consider Kant's statements about moral improvement in the *Religion*.

Kant's Two-Fold Account of Moral Improvement

Before examining Kant's account of moral improvement, let us briefly review Kant's distinction between the empirical and intelligible character of our actions.⁴⁶³ All our imputable actions have both an empirical and intelligible character because they can be described as natural phenomena in the world of sense and as appearances that express the nature of our intelligible use of our power of choice. When a person does an action out of respect for the moral law (say, helps someone in need), there can be an empirical

⁴⁶³ In the *Critique of Pure Reason*, as we saw earlier, Kant writes, [E]very effective cause must have a character, i.e., a law of its causality, without which it would not be a cause at all. And then for a subject of the world of sense we would have first an empirical character, through which its actions, as appearances, would stand through and through in connection with other appearances in accordance with constant natural laws, from which, as their conditions, they could be derived; and thus, in combination with these other appearances, they would constitute members of a single series of the natural order. Yet second, one would also have to allow this subject an intelligible character, through which it is indeed the cause of those actions as appearances, but which does not stand under any conditions of sensibility and is not itself appearance. The first one could call the character of such a thing in appearance, the second its character as a thing in itself" (CPR A539/B567). As we will see as my discussion unfolds, this distinction must be emphasized and centrally regarded in any discussion of moral improvement within a Kantian moral framework.

description of the action. From a psychological perspective, the account would culminate in some reference to the psychological state of pure respect for the moral law as providing the ultimate psychological motivation for the action. This, in turn, may involve a description of the relationship between the respect they feel for the moral law and the nature of the inclinations that may have stood in opposition to the moral law's requirements.

We can imagine an empirical account, therefore, of a person who is perhaps about to spend her money on something that will satisfy an inclination (say, she is on her way to a nice restaurant) when she passes by a collection station for a charity. Instead of satisfying the inclination for delicious food, the person may elect to take some of her money and give it to charity and instead eat someplace cheap. Even though she was excited for the delicious food from the nice restaurant, her commitment to the moral law—expressed here through her imperfect duty to beneficence—overrides her inclinations. An empirical account of this action can provide a good deal of information about her psychological state and the precise nature of the struggle between her inclinations on the one hand and her respect for the moral law on the other.

For her action to be authentically imputable, though, there must also be an intelligible character to her action. That is, for her action to carry any moral worth, her action must be the result of her own choice. She must be the author of the action. As we have previously seen, if we are to be the authors of some of our actions (those for which we bear moral responsibility), then they must result from our freedom. Freedom, in turn, cannot reside in appearances. Appearances are governed by the laws of nature, which

leave no room for moral freedom. Only by maintaining a distinction between things in themselves from appearances is an account of freedom rendered possible.⁴⁶⁴ As Kant writes in the *Critique of Pure Reason*, “[I]f appearances are things in themselves, then freedom cannot be saved.”⁴⁶⁵ Keeping Kant’s “critical distinction” in mind, then, we must say that our morally imputable actions are ultimately the result of the way in which our power of choice is exercised in the intelligible world. It follows, then, that we cannot get much of an account of intelligible choice. All we can say is that a choice is made and our actions in appearances are determined by it.

This distinction between intelligible and empirical characters of our actions must remain central to a discussion of Kant’s account of moral improvement. If Kant wants to describe and, importantly, account for moral improvement in a human being, then the discussion must be limited only to imputable actions that are morally relevant. We can, it seems, speak freely enough of improvement in a more general, amoral sense. We can provide an account for how someone, desiring to become famous and respected in the eyes of others, went from causing unhappiness and misery in people to augmenting and promoting human happiness by only referring to the empirical character of her actions. We may try to trace her inclination to be respected back to its roots and then to provide an account for how such inclinations led her to behave in such and such way.

Importantly, though, such an account would attach no moral worth to her actions (good as

⁴⁶⁴ This, of course, is a version of the famous claim Kant makes in the B Preface of the *Critique of Pure Reason*, when he writes, “...nevertheless, I can think freedom to myself...so long as our critical distinction prevails between our two ways of representing (sensible and intellectual), along with the limitation of the pure concepts of the understanding arising from it...” (CPR Bxxviii).

⁴⁶⁵ CPR A536/B564.

they may be). For an account of *moral* improvement, there needs to be reference to, and assurance of, a change in the intelligible character of a person's actions.

In Part One of the *Religion*, Kant provides an account of moral improvement that tracks this distinction between the empirical and intelligible character of our actions. Here, Kant describes moral improvement as the “restoration of the original predisposition to good in us.”⁴⁶⁶ Most, if not all, people find themselves at some point or other prioritizing self-love over the moral law. Even in cases where people do not find themselves in flagrant violation of the moral law, they can never rest assured with complete confidence that their law-abiding behavior is motivated by pure respect for the law. In fact, it is the more likely hypothesis to suppose that their motives are more opaque than they seem and there is likely to be some influence of inclinations behind their actions. This privileging of self-love over the moral law is moral evil. Moral improvement must involve an inversion of two principles. The moral law, if moral improvement is to occur, must be the ultimate ground for one's actions. Actions must be motivated purely by respect for the moral law. Reinforcing inclinations may accompany the feeling of respect, but they cannot play a causally efficacious role in one's motivation.⁴⁶⁷ That is, an otherwise lawful action has no moral worth if *were it not for*

⁴⁶⁶ Rel. 6:46.

⁴⁶⁷ If inclinations did play even a partial role in the motivating ground of one's action, the person would still be morally evil. Earlier in the *Religion*, Kant describes the three basic “grades” of our “natural propensity to evil” (Rel. 6:29). The second grade of evil is what Kant calls, “the impurity (impuritas, improbitas of the human heart consists in this, that although the maxim is good with respect to its object (the intended compliance with the law) and perhaps even powerful enough in practice, it is not purely moral, i.e. it has not, as it should be [the case], adopted by the law alone as its sufficient incentive, but, on the contrary, often (and perhaps always) needs still other incentives besides it in order to determine the power of choice for what duty requires; in other words, actions conforming to duty are not done purely from duty” (Rel. 6:30).

inclinations pushing the agent to that action, the action would not have occurred. Kant, unsurprisingly, then, describes this moral improvement as “only the recovery of the *purity* of the law, as the supreme ground of all our maxims, according to which the law itself is to be incorporated into the power of choice, not merely bound to other incentives, nor indeed subordinated to them (to inclinations) as conditions, but rather in its full purity, as the self-*sufficient* incentive of that power.”⁴⁶⁸ Any moral improvement must involve the priority of moral over self-love. Anything short of that, it seems, would not be *bona fide* moral improvement, if moral improvement is understood as the “restoration of the original predisposition to good in us.”⁴⁶⁹

Kant’s account of moral improvement, then, seeks to retain this rigorous standard while also acknowledging the empirical and intelligible characters of our actions. An account of moral improvement needs to begin with the state of the moral actor before the moral improvement occurs. Prior to moral improvement, there is an incessant moral push on us, even when we voluntarily violate the moral law. Kant writes, “[T]he command that we *ought* to become better human beings still resounds unabated in our souls... Surely we must presuppose... that there is still a germ of goodness left in its entire purity, a germ that cannot be extirpated or corrupted.”⁴⁷⁰ Without the constant prodding and pressure from the “germ of goodness” in us, we would never feel impelled to alter fundamentally the priority between the moral law and self-love.

⁴⁶⁸ Rel. 6:46.

⁴⁶⁹ *Ibid.*

⁴⁷⁰ Rel. 6:45.

Kant's account of moral improvement is two-fold, involving related discussions of moral improvement: one regarding improvement in the empirical world and the other regarding improvement in the intelligible world. As we will see, Kant's account for moral improvement in the empirical world relies on the possibility for moral improvement in the intelligible world.

Kant describes the apparent trajectory of moral improvement in the empirical world first. Kant writes, "When the firm resolve to comply with one's duty has become a habit, it is called *virtue*...in a legal sense, in its *empirical character (virtus phaenomenon)*. Virtue here has the abiding maxim of *lawful* actions, no matter whence one draws the incentives that the power of choice needs for such actions."⁴⁷¹ Lawful actions conform to the demands of the moral law, but may not necessarily be motivated solely by respect for the moral law. The way a person improves in his or her capacity to engage in lawful actions is piecemeal. Kant writes, "Virtue, in this sense, is accordingly acquired *little by little*, and to some it means a long habituation (in the observance of the law), in virtue of which a human being, through gradual reformation of conduct and consolidation of his maxims, passes from a propensity to vice to its opposite."⁴⁷² Kant cites some examples of this kind of improvement toward legality in the following manner: "[A]n immoderate human being converts to moderation for the sake of health; a liar to truth for the sake of reputation; an unjust human being to civic righteousness for the sake of peace or profit, etc., all in conformity with the prized principle of a

⁴⁷¹ Rel. 6:47.

⁴⁷² *Ibid.*

happiness.”⁴⁷³ These sorts of reforms of one’s actions are the kinds of observable patterns of behavior that seem to track moral improvement. It seems fair to say that we are likely to think of a formerly dishonest (but now honest) person as someone who has undergone moral improvement.

It is clear from Kant’s account of moral worth, though, that there is a crucial element to moral improvement that may be lacking in these cases. It is not enough for one’s action to conform to the moral law; they must also be motivated by pure respect for the moral law. Kant claims that cases like these are possible even when there has been no authentic moral improvement: “But not the slightest *change of heart* is necessary for this; only a change of *mores*.”⁴⁷⁴ Even a person who is improving in the legal sense, they may still lack moral goodness. In the cases given above, a person whose actions are lawful may still be motivated by the principle of self-love. We may also imagine cases where a person commits more fervently to abide by the moral law, but still occasionally deviates from it. Maybe the frequency with which they violate the moral law diminishes, but their maxim still incorporates the principle of self-love in a privileged way. Recall Kant’s definition of human evil: “[T]he statement, “The human being is *evil*,” cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (*occasional*) deviation from it.”⁴⁷⁵ Even if one is violating the moral law less frequently, it is not clear how we can still call such a development *moral* improvement. It

⁴⁷³ *Ibid.*

⁴⁷⁴ *Ibid.*

⁴⁷⁵ Rel. 6:32, emphasis added.

seems, rather that the supreme maxim upon which the person acts is becoming more refined, or perhaps more and more intelligent in achieving one's happiness.

Kant notices a tension here between the stricter standards his account of moral evil requires for moral improvement and the importance we often attach to improvement in one's behavior when he writes, "However, that a human being should become not merely *legally* good, but *morally* good (pleasing to God) i.e. virtuous according to the intelligible character [of virtue] (*virtus noumenon*) and thus in need of no other incentive to recognize a duty except the representation of duty itself...cannot be effected through gradual *reform* but must rather be effected through a *revolution* in the disposition of the human being (a transition to the maxim of holiness of disposition)."⁴⁷⁶ It seems that we can rely on our own naturally possessed qualities to enact a reformation of our character, but, again, this alone cannot count as moral improvement unless it is accompanied by a reordering of the relationship between the principles of self-love and of morality. That is, moral improvement would *already have had to occur* via a "revolution in the disposition of the human being."

If the moral improvement has already happened, though, what are we to make of the gradualism of reform? In his account, Kant attempts to account for, and to incorporate, both the reforming and revolutionary aspects of moral improvement. Kant introduces his two-fold approach to think about moral improvement when he writes, "But if a human being is corrupt in the very ground of his maxims, how can he possibly bring about this revolution by his own forces and become a good human being on his own? Yet

⁴⁷⁶ Rel. 6:47.

duty commands that he be good, and duty commands nothing but what we can do. The only way to reconcile this is by saying that a revolution is necessary in the mode of thought but a gradual reformation in the mode of sense (which places obstacles in the way of the former), and [that both] must therefore be possible also to the human being.”⁴⁷⁷ The “gradual reformation in the mode of sense” is achieved through an increased participation in lawful actions. The idea seems to be the moral convert, as a result of a reordering of the principles of self-love and morality, engages increasingly in lawful behavior. Kant describes the morally improved person as possessing “genuine disposition toward improvement...”⁴⁷⁸ Meanwhile, the “revolution in the mode of thought” is achieved through structuring one’s maxims in such a way that objectively the moral law and subjectively respect for the moral law have priority over the principle of happiness. The real source of moral improvement must be in the revolutionary reordering of self-love and morality. The reformation of one’s actions is the inevitable consequence of the revolution. We can only *see* lawful actions from others, but lawful actions are not sufficient proof for a moral conversion. We may get something closer to a proof over the course of an entire lifetime in which a person moves consistently and constantly towards lawful actions:

If by a single and unalterable decision a human being reverses the supreme ground of his maxims by which he was an evil human being...he is to this extent, by principle and attitude of mind, a subject receptive to the good; but he is a good human being only in incessant laboring and becoming i.e. he can hope – in view of the purity of the principle which he has adopted as the supreme maxim of

⁴⁷⁷ Rel. 6:47.

⁴⁷⁸ Rel. 6:77.

his power of choice, and in view of the stability of this principle – to find himself upon the good (and narrow) path of constant *progress* from bad to better. For him who penetrates to the intelligible ground of the heart... for him to whom this endless progress as a unity, i.e. for God, this is the same as actually being a good human being (pleasing to him); and to this extent the change can be considered a revolution.⁴⁷⁹

Moral improvement requires this fundamental switch in one's attitude towards inclinations and the feeling of respect for the moral law. For this reason, moral education "must begin, not with an improvement of mores, but with the transformation of his attitude of mind and the establishment of a character, although it is customary to proceed otherwise and to fight vices individually, while leaving their universal root undisturbed."⁴⁸⁰ Hence, a merely reformist account of moral improvement will tend toward instances of individuals who cease lying for the sake of reputation or seek civic justice merely for the sake of securing profits. Only the more radical revolutionary inversion of self-love and morality can bring about authentic moral improvement. Reform's place in this picture is as a representation or expression of the more basic change.

Importantly, though, the ultimate ground for this needs to be in the intelligible realm, which, ultimately, determines moral worth. Kant's discussion here seems to be entirely within the empirical world. That is, even the revolutionary change brought about by the reordering of the principles of self-love and of morality are themselves psychological phenomena, determining the content of our maxims. Maxims, it bears

⁴⁷⁹ Rel. 6:48.

⁴⁸⁰ *Ibid.*

emphasizing, are themselves psychological, empirical states. If moral improvement is actually to occur, then it would need to result from a new state of affairs in the intelligible world. Without an account for the ultimate ground of moral improvement, then it seems like Kant's account of moral improvement would be incomplete. It would lack an account of the exact way in which moral improvement actually happens. For human beings, who are often tasked by the moral law and wishful that they could live up to it, but who are nevertheless living in a state of moral evil, an account for how moral improvement occurs at the basic, radical level (that is, the intelligible character of one's actions) may stave off moral despair. For if one cannot even imagine how it could be that one could become good, then it may well dispirit our resolve.

VII. Moral Improvement and Grace

For Kant, authentic moral improvement must have as its ultimate ground the intelligible exercise of our power of choice. Evil people must have exercised their freedom of choice in such a way that, in the empirical world, they would commit misdeeds or would act on the basis of impure motives. That means that people who are now good, but were formerly evil, must have altered in some way the way in which their power of choice was exercised so that their lives are now marked by the "incessant laboring" on the "good (though narrow) path of constant *progress* from bad to better."⁴⁸¹ An account of this change at the intelligible level, though, will, by the limitations imposed on our knowledge of the intelligible by Kant's transcendental idealism, be thin and nearly barren.

⁴⁸¹ Rel. 6:48.

Kant does discuss the nature of this transformation that must take place in cases of authentic moral improvement, but his discussion, I would argue, is puzzling. In the *Religion* Kant writes, “If it is said, The human being is created good, this can only mean nothing more than: He has been created for the *good* and the original *predisposition* in him is good; the human being is not thereby good as such, but he brings it about that he becomes either good or evil, according as he either incorporates or does not incorporate into his maxims the incentives contained in that predisposition (and this must be left entirely to his free choice).”⁴⁸² So far, this is simply a reiteration of the basic features of Kant’s views of moral worth. The next sentence, though, Kant goes on to write, “Granted that some supernatural cooperation is also needed to his becoming good or better, whether this cooperation only consist in the diminution of obstacles or be also a positive assistance, the human being must nonetheless make himself antecedently worthy of receiving it; and he must *accept* this help...in this way alone is it possible that the good be imputed to him, and that he be acknowledged a good human being.”⁴⁸³ Here, Kant claims that “supernatural cooperation” is necessary for a human being to become morally better. There are at least two questions one may ask regarding this notion of “supernatural cooperation.” First, what notion of “supernatural” is Kant using? Second, what is the source of the cooperation?

Before considering a potential Kantian response, it is worthwhile to examine a response that was very likely on the minds of Kant’s more religiously minded readers. There is a single traditional religious response to both these questions: the cooperation is

⁴⁸² Rel. 6:44.

⁴⁸³ *Ibid.*

supernatural in the sense that its source is God, a supernatural being. Furthermore, the process whereby one receives supernatural cooperation in becoming good is called *grace*. Although Kant seems at least skeptical and at most bitterly dismissive of the notion of grace, it seems that Kant's account of intelligible moral improvement shares some characteristics with the traditional notion of grace.

The “supernatural cooperation” discussion quoted in the *Religion* is, in some ways, strikingly similar to the discussion of grace Kant has in his work *The Conflict of the Faculties*. There, Kant writes, “Finally, as far as our will and its fulfillment of God's commands is concerned, the biblical theologian must not rely on nature—that is, on the human being's own moral power (virtue) – but on grace (a supernatural but, at the same time, moral influence), which the human being can obtain only by an ardent faith that transforms his heart – a faith that itself, in turn, he can expect only through grace.”⁴⁸⁴

The mention of “supernatural cooperation” in the quotation from the *Religion* (with its obvious resonances with the Christian doctrine of grace) and the pride of place apparently given to grace in the quotation from *The Conflict of the Faculties* may seem perplexing, given Kant's account of moral worth. It would be important to note here, though, that in *The Conflict of the Faculties*, Kant claims that the “biblical theologian” makes reference to grace in order to explain how a person can become moral. The philosopher, by contrast, cannot overtly rely on a notion that draws its power from revealed religion, and not reason. Nevertheless, the notion of “supernatural cooperation”

⁴⁸⁴ Rel. 7:24.

we encounter in the *Religion* operates in much the same way as grace would within religious discourse.

While they may function similarly in both contexts, the notion of grace creates fundamental problems within Kant's critical moral philosophy. It seems that if the ultimate cause of a human being's moral improvement is the result of a gift from some external source (grace is traditionally described as an (undeserved) gift from God), then moral improvement cannot be imputed exclusively to moral agents. Kant himself seems to be voicing this very same concern in the *Religion* when he writes, "[T]o expect an effect of grace means...that the good (the morally good) is not our doing, but that of another being..."⁴⁸⁵ Furthermore, Kant seems to heap scorn upon the mindset which bemoans one's own immorality and waits to receive grace passively:

Now surely it is not because of the inner nature of the Christian faith, but because of the manner in which people's minds are introduced to it, that a similar charge can be brought against it with respect to those who are the most serious about it but who, starting with human corruption and despairing of all virtue, place their religious principle solely in *piety* (by which is understood the principle of conducting oneself passively in view of the divine blessedness expected through a power from above). For these [individuals] never place any reliance in themselves but constantly look about them in constant anxiety for a supernatural assistance, and even think that in this self-contempt (which is not humility) they possess a means of obtaining favor. The outward expression of this (in pietism or false piety) is indeed a sign of a *slavish* cast of mind.⁴⁸⁶

⁴⁸⁵ Rel. 6:53.

⁴⁸⁶ Rel. 6:184n.

Grace, defined as the intelligible basis for reordering the relationship between self-love and morality, is not something that can occur in the empirical world; hence Kant's obvious disdain for preachers who claim the ability to generate grace like a weaver can generate a garment. Grace cannot be caused by, or occasioned by, events in the natural world. In the religious context, prayer is sometimes held up as a means a receiving grace from God. Kant dismisses this notion when he writes, "*Praying*, conceived as an *inner ritual* service to God and hence as a means of grace, is a superstitious delusion (a fetish-making); for it only is the *declaring of a wish* to a being who has no need of any declaration regarding the inner disposition of the wisher, through which nothing is therefore accomplished nor is any of the duties incumbent on us as commands of God discharged; hence God is not really served."⁴⁸⁷ The person in prayer here is not all that different from the "hardened scoundrel" Kant describes in the *Groundwork* who wishes he could be good.⁴⁸⁸ While it is true that the noumenal choice seems remote from the phenomenal experience of the self (so remote that it may seem to be wholly alien from the phenomenal self), it is crucial that the noumenal choice is nevertheless thought of as self-wrought (that is, I and not God am the agent responsible for it).

So, while Kant's account of moral improvement cannot simply be reduced to a religious account of grace, there are good, heuristic reasons for loosely associating the two. The basic relationship between the empirical and intelligible character is one of ground and consequence. The activity of the intelligible self determines the actions of the

⁴⁸⁷ Rel. 6:194. Readers of Ralph Waldo Emerson may be reminded of Emerson's claim in the essay "Self-Reliance," where he writes: "As men's prayers are a disease of the will, so are their creeds a disease of the intellect" (276).

⁴⁸⁸ G 4:454-5.

sensible self.⁴⁸⁹ We cannot access the working of the intelligible world, nor stake any sort of knowledge claim regarding how the power of choice is exercised. We only know that freedom must reside in the intelligible world and that that exercise of freedom determines the imputable actions we take in the empirical world. From the perspective of our empirical selves, then, it may seem to some that the source of moral improvement must be some outside source (since there is a metaphysical gulf between the intelligible and the sensible worlds). We are somewhat in a position, as empirical beings, of discovering through our (empirically) chosen actions the nature of our intelligible use of freedom. Importantly, also, there does not seem to be any way in which Kant can allow for empirical self to determine the nature of the intelligible self. The intelligible self must be the ground of the empirical self.

It is important to emphasize, though, that this line of reasoning does not imply a kind of pious quietude about moral improvement. Kant is adamant in claiming that the surest sign of an intelligible switch from evil to good is a long and steady movement from immorality to morality, from lawless actions to lawful ones. A “genuine disposition toward improvement” and not mere “self-proclaimed confessions of faith” is the authentic mark of a morally improved human being.⁴⁹⁰ Given the rigid, absolute

⁴⁸⁹ “The true activity of reason and its effect belong to the intelligible world (*mundo intelligibili*). Hence we do not know to what degree we should impute. Nevertheless we know that the influencing power of reason is not determined and necessitated by any phenomena, rather that it is free, and (in the case of imputation) we judge the action merely in accordance with rational laws. The actions here in the world are mere Schemata of the intelligible [actions]; yet these appearances (this world already signifies “schema”) are still interconnected in accordance with empirical laws, even if one regards reason itself, in accordance with its expressions, as a phaenomenon (of the character). But what the cause of this may be we do not discover in phaenomenis” (Notes on metaphysics 18:254).

⁴⁹⁰ Rel. 6:77.

distinction between the phenomenal and the noumenal, one may well press Kant on his account of moral improvement. In what sense can the notion of moral improvement be intelligible so long as the authentic ground for free actions is in a self of whom we have no direct experience? In the next section, I consider two possible concerns that can be raised against Kant on the issue of moral improvement.

VIII. Two Possible Objections to Kant's Account of Moral Improvement

Kant's account of moral improvement, I would argue, may be open to at least two possible objections. The first centers on Kant's rather strict and narrow conception of moral goodness; the second centers on the possibility of intelligible change. I will deal briefly with each objection. I will argue that the first objection can be met, but not without first accepting a paradoxical feature of Kant's account of moral improvement. I will then argue that the second objection may seem more serious, but can also be met. I will end by arguing that the second objection brings out a logical consequence of basic features of Kant's metaphysics.

Is 'Bad to Better' Any Different from 'Bad to Bad'?

In the *Religion*, Kant is clearly committed to the position that human beings express earnest devotion to the moral law through a whole life of moral toil in which they gradually diminish their immoral acts. After describing the person who engages in a lifelong struggle to be good, in which finds "himself upon the good (though narrow) path of constant *progress* from bad to better," Kant writes, "For him who penetrates to the intelligible ground of the heart (the ground of all the maxims of the power of choice), for

him to whom this endless progress is a unity, i.e. for God, this is the same as actually being a good human being (pleasing to him)...⁴⁹¹ Human beings, in contrast to God, must see our actions as temporally organized, which involves a continuous stream of actions that gets progressively closer to moral goodness, but perhaps without ever fully arriving there: “For the judgment of human beings, however, who can assess themselves and the strength of their maxims only by the upper hand they gain over the senses in time, the change is to be regarded only as an ever-continuing striving for the better, hence as a gradual reformation of the propensity to evil, of the perverted attitude of mind.”⁴⁹² The image seems to be that if a person begins life in a state of frequent violation of the moral law and, over time, they diminish the frequency of their immoral actions (through this constant struggle desire to act out of respect for the moral law), then they will have improved morally over the course of his or her life.

It is not clear, though, how seeing a human’s life as a unity moving from less lawfulness to more lawfulness alone is sufficient to show authentic moral improvement at all. If (1) the sole determination of the moral worth of a human being is whether or not self-love is privileged over morality, and (2) the person ends up in a situation where they are still occasionally violating the moral law or are acting from mixed motives, then it follows that the person was evil in the beginning of the progression and he or she remains evil at the end of it. If this is correct, then it leads to the paradoxical claim that an evil human being can, at the same time, be a good human being (insofar as he or she has improved).

⁴⁹¹ Rel. 6:48.

⁴⁹² *Ibid.*

There is, perhaps, a response that one could make on Kant's behalf here. Perhaps the paradox of a person being evil and good is not as contradictory as it may initially seem because there is a precise sense in which both can be true of the same person. Consider the question: What is the volitional structure of a person who has undergone moral improvement? We can imagine a person who, in previous times (but no longer), would satisfy an inclination to superiority and brag about past accomplishments in an attempt to make others feel inferior. The person used to act on the inclination, and perhaps the inclination is still present as a potential source of psychological satisfaction, but now the person reins in the inclination and refrains from bragging out of respect for others (which the moral law requires). We can further imagine this very same person, shortly after refraining from bragging about past accomplishments now telling a self-serving lie motivated, say, by an inclination to vanity. The person, therefore, remains evil, even though there has been moral improvement with regard to the inclination to superiority. So, while the inclination to superiority (one expression of the desire for happiness) has been brought into a proper relation with the moral law, the inclination to vanity (another expression of the desire for happiness) has not. The supreme maxim upon which the person acts must therefore be evil. It would take the form (partially and roughly) of the following: "When an inclination to superiority arises in me, I will restrain it and not allow for it to deviate from my moral duty, but when an inclination to vanity arises in me I will forego my moral duty not to lie." This is certainly better, in a sense, than a supreme maxim that would privilege both superiority complexes and vanity, but it seems odd to call such a person morally good.

Imagine, though, that motivating influence of this inclination to vanity is itself diminished over time. Say this tendency generalizes over *all* one's inclinations. One acts more and more out of respect for the moral law at the end of the temporal sequence than at the beginning. This does in fact seem to resemble the kind of reforming tendency that Kant described in the *Religion*.

Is Intelligible Improvement Coherent?

One may ask a further question here: What does this tell us about the nature of the intelligible choice one has made? If the progression is in fact a *moral* one, then it follows that the way in which the intelligible choice was exercised had to have changed. If one previously engaged in self-aggrandizement at the expense of others' self-image, then the intelligible choice must be the ultimate source of the action. So, if a person who previously was immoral in that regard, but is no longer, the change must be explained in terms of the exercise of one's intelligible power of choice. Whereas the person's choice was previously exercised in a way to lead to bad moral outcomes in the presence of that inclination, it now must be exercised in a different way. That is, there had to have been moral improvement in the intelligible power of choice itself. How, though, is the notion of moral conversion or improvement in the intelligible world possibly coherent? In a note I had quoted earlier, Kant writes, "In the intelligible world nothing happens and nothing changes, and there the rule of causal connection disappears."⁴⁹³ If nothing changes in the intelligible world, then how can an evil person change at all? Appearances, among which are actions in which we either fail or succeed in carrying out our duty, are simply the

⁴⁹³ Notes on metaphysics 18:254.

effects of the real actions of reasons, which are not found in appearances, but rather in the intelligible world.⁴⁹⁴ It seems that the very notion of “change” cannot apply to the intelligible world without obscuring the distinction between the empirical and intelligible worlds.

While the problem of intelligible change may, at first, seem problematic for Kant, there is, perhaps, a way to respond to the charge. One could argue, on Kant’s behalf here, that the possibility of moral improvement does not lead to problems for Kant’s overall picture because there is nothing problematic in asserting that the intelligible subject’s choice can actually undergo change. In defending the possibility of noumenal change, one can evoke Kant’s principle of the primacy of practical reason. If our use of practical reasoning requires us to posit the possibility of change in the noumenal self, then we are not at fault for doing so. Kant lays out his notion of the primacy of practical reason in the second *Critique* in “On the Primacy of Pure Practical Reason In Its Connection With Speculative Reason.”⁴⁹⁵ There, Kant writes, “[I]n the union of pure speculative with pure practical reason in one cognition, the latter has primacy, assuming that this union is not *contingent* and discretionary but based on reason itself and therefore *necessary*.”⁴⁹⁶ The idea seems to be that if pure practical reason derives a claim from necessary principles (i.e. principles upon which the faculty of practical reason rests), then, so long as the claim is *consistent with*, theoretical reason (i.e. it is not in open contradiction with theoretical

⁴⁹⁴ “Reason is not in the chain of appearances and is with regard to all of that free with regard to its own causality (the actions of reason themselves are also not appearances, only their effects are)” (Notes on metaphysics 18:252).

⁴⁹⁵ CPrR 5:119-121.

⁴⁹⁶ CPrR 5:121.

reason), speculative reason must allow the claim to hold. In a paper on the topic, Sebastian Gardner describes Kant's principle of the primacy of practical reason as follows: "To assert the primacy of practical reason with respect to speculative reason is, therefore, to assert that propositions on which the interest of practical reason depends necessarily, so long as they are not contradicted by theoretical reason, must be accepted by theoretical reason."⁴⁹⁷ In the case of noumenal change, then, if it can be shown that the need to posit the possibility of noumenal change rests on basic, necessary principles of pure practical reason, then we can allow for it, so long as it does not contradict the claims of theoretical reason.

First, then, let us examine what principles we can trace that inform and necessitate the need to allow for the possibility of noumenal change. I would argue that the possibility of noumenal change is an inescapable result of accepting two basic features of Kant's moral philosophy: (1) The universal validity of the fact of reason and (2) The principle of 'ought implies can.' Since both (1) and (2) are taken by Kant to be necessary, then we are compelled to accept (3) The possibility of noumenal change. As we saw earlier, Kant describes the 'fact of reason' in the following way: "Consciousness of this fundamental law may be called a fact of reason because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or

⁴⁹⁷ Sebastian Gardner, "The Primacy of Practical Reason," from *A Companion to Kant* (ed. Graham Bird).

empirical...”⁴⁹⁸ The fact of reason is how we come to know that morality’s requirements apply to *us*. Morality applies to equally to those who have well-ordered wills and those who lack well-ordered wills. It follows, then, that those who currently have a will that prioritizes self-love over morality are morally obligated to reorder their priorities and incorporate the moral law as the supreme ground of all their actions. Where there is a moral obligation, there must also be the capacity to fulfill one’s obligation. That is, ought implies can. Kant regularly invokes the principle that ‘ought implies can.’ See, for instance, in the second *Critique*, where Kant writes, “He judges, therefore, that he can do something because he is aware that he ought to do it and cognizes freedom within him, which, without the moral law, would have remained unknown to him.”⁴⁹⁹ Here Kant is claiming that knowing one ought do something is sufficient to know that one can do something. If an evil person has a moral obligation to become good *and* moral improvement must ultimately be grounded in a change in how noumenal choice is exercised, then noumenal change must be possible.

IX: The Propensity to Evil

Kant holds that there is a propensity to evil in all human beings. With such a claim, it might seem that Kant is veering toward a kind of moral fatalism that dooms all of humanity to lives of immorality without real hope of moral improvement. In discussing this propensity to evil, Kant writes, “[W]e may presuppose evil as subjectively necessary in every human being, even the best.”⁵⁰⁰ In light of statements such as this, it

⁴⁹⁸ CPrR 5:31.

⁴⁹⁹ CPrR 5:30.

⁵⁰⁰ Rel. 6:32.

becomes understandable to be concerned about a creeping moral pessimism in the *Religion* and to worry about how such a bleak picture can be consistent with Kant's other substantive ethical claims about human freedom and dignity. In what follows I will provide a close reading the section of the *Religion* where Kant discusses the propensity to evil. I will argue that, first appearances notwithstanding, Kant's claims about the propensity to evil are consistent with his broader moral philosophy.

Before delving into the minutiae of Kant's discussion of the propensity to evil, it makes good sense to reiterate briefly some basic features of Kant's theory of moral accountability. Doing so will provide basic interpretive guidelines that will need to inform any account of the propensity to evil. That is, if Kant's account of the propensity to evil is to be consistent with his overall moral philosophy, it would need to be in accordance with his more central claims about the nature of moral accountability and moral worth. Any defense of Kant's discussion of the propensity to evil in the *Religion*, therefore, must show that the discussion does indeed square with the more substantive and central features of his moral philosophy.

A person is morally accountable only for actions that result from freely chosen maxims. If we are to assign moral approbation or disapprobation to a person, then we are assuming that his or her actions were freely chosen. Moral goodness (or evil) is expressed through one's freely chosen maxims, the principle of volition that underlies our actions: "Hence the ground of evil cannot lie in any object *determining* the power of choice through inclination, not in any natural impulses, but only in a rule that the power of

choice itself produces for the exercise of its freedom, i.e., in a maxim.”⁵⁰¹ The crucial idea, then, is that, underlying the actual carrying out of a deed, there is some maxim upon which the deed is based. Furthermore, the adoption of that maxim itself must be predicated on some exercise of free choice to take on that maxim in the first place. Kant refers to this choice as the “supreme ground of the human being’s freedom.” When explaining the idea that human beings are evil by nature, Kant writes, “[L]et it be noted that by “the nature of a human being” we only understand here the subjective ground—wherever it may lie—of the exercise of the human being’s freedom in general (under objective moral laws) antecedent to every deed that falls within the scope of the senses.”⁵⁰² Kant’s account of the propensity to moral evil, then, should be consistent with this more general account of moral responsibility.

The Propensity to Evil, Defined

Kant first defines propensity in the following manner: “By *propensity* (*propensio*) I understand the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*), insofar as this possibility is contingent for humanity in general.”⁵⁰³ A propensity, on this rendering, is the thing that makes us susceptible to particular inclinations. In a note on this passage Kant adds, “*Propensity* is actually only the *predisposition* to desire an enjoyment which, when the subject has experienced it, arouses *inclination* to it.”⁵⁰⁴ A person can have a propensity to something without actually

⁵⁰¹ Rel. 6:21.

⁵⁰² *Ibid.*

⁵⁰³ Rel. 6:29.

⁵⁰⁴ Rel. 6:29n.

desiring it. Kant's racist example of "savages" with a propensity for "intoxicants" indicates that, on Kant's view, a propensity can exist without an accompanying desire in cases where the person has no concept of the object that would satisfy the merely potential inclination.⁵⁰⁵

Kant's discussion of propensity, though, becomes a bit confused in the very next sentence. He writes, "It [propensity] is distinguished from a predisposition in that a propensity can indeed be innate yet *may* be represented as not being such: it can rather be thought of (if it is good) as *acquired*, or (if evil) as *brought* by the human being *upon* himself."⁵⁰⁶ In the previously quoted note, Kant seems to be claiming that a propensity *is* a predisposition. Here, Kant is distinguishing between a propensity and a predisposition, claiming that a propensity, unlike a predisposition, can be represented as not being innate.

Kant alleviates some of this confusion a bit later in the *Religion*. After laying out the three "grades" of this natural propensity to evil, Kant provides a useful expansion of his notion of a propensity: "Every propensity is either physical, i.e. it pertains to a human's power of choice as natural being or moral, i.e. it pertains to a human's power of choice as a moral being."⁵⁰⁷ Kant then insists that "In the first sense [the physical sense] there is no propensity to moral evil, for the latter must originate from freedom; a physical propensity (one based on sensory inducements) to whatever use of freedom, be it for good or evil, is a contradiction."⁵⁰⁸ Physical propensities are morally neutral for the same

⁵⁰⁵ *Ibid.*

⁵⁰⁶ Rel. 6:29.

⁵⁰⁷ Rel. 6:31.

⁵⁰⁸ *Ibid.*

reason we refrain from rendering moral judgments onto a ferocious lion that kills a human being. There cannot be moral responsibility without an accompanying capacity of choice to guide whether or not a being (whether it be a human or another kind of animal) will act to satisfy an inclination. Kant then concludes, “Hence a propensity to evil can only attach to the moral faculty of choice. Nothing is, however, morally (i.e. imputably) evil but that which is our own deed.”⁵⁰⁹

Kant then enters into a confusing discussion about the two meanings of the word “deed.” While it might, on a first look, appear insignificant, I would argue that it is crucial in coming to an understanding of Kant’s concept of propensity to evil. Kant gets into his discussion of the term “deed” by noting that “the concept of a propensity is understood a subjective determining ground of the power of choice *that precedes every deed*, and hence is itself not a *deed*.”⁵¹⁰ If a propensity is not itself a deed, but rather the determining ground of the power of choice, and, as Kant claims, only deeds are morally (or imputably) evil, then it would seem to follow that the propensity to evil would not itself be evil. Kant, though, explicitly states that the propensity to evil is itself evil: “[T]his propensity must itself be considered morally evil, hence not a natural predisposition but something that a human being can be held accountable for...”⁵¹¹ Kant recognizes this potential inconsistency, and provides an illuminating (for understanding his position) discussion of the two possible senses of the word “deed.” He begins by writing, “There would then be a contradiction in the concept of a simple propensity to

⁵⁰⁹ *Ibid.*

⁵¹⁰ *Ibid.*

⁵¹¹ Rel. 6:32.

evil, if this expression could not somehow be taken in two different meanings, both nevertheless reconcilable to the concept of freedom.”⁵¹² Both “meanings” of deed must be “reconcilable” to freedom because deeds, if they are to carry any moral content (if can be said to be good or evil) must be imputable, and therefore must be grounded in the free exercise of the power of choice.

Kant introduces the dual meaning of the phrase “deed” with the following long sentence: “Now, the term “deed” can in general apply just as well to the use of freedom through which the supreme maxim (either in favor of, or against, the law) is adopted in the power of choice, as to the use by which the actions themselves (materially considered, i.e. as regards the objects of the power of choice) are performed in accordance with that maxim.”⁵¹³ Kant here is distinguishing between two kinds of imputable actions (or deeds). On the one hand, we have ordinary, observable actions to which we attach moral significance. We may observe someone performing a perfect duty, e.g. telling the truth even when it seems to be against their immediate self-interest to do so. On the evil side, we may witness a person treating another person with contempt by ridiculing them. Both these observable actions are deeds insofar as we can impute those actions to the moral agent. On the other hand, we have the freely chosen of a supreme maxim, which consists, in Kant’s starkest rendering of the choice, as the prioritizing of self-love over the moral law (or vice versa) as the overriding principle of volition. An evil supreme maxim privileges self-love over morality such that whenever there is a clash between the demands of morality and the pleadings of self-love, the agent will formulate his or her

⁵¹² Rel. 6:31.

⁵¹³ *Ibid.*

lower level maxims in a way that is consistent with the principle of self-love. This, ultimately, will lead to deeds in the other sense that will be at odds with one's duty. The propensity to evil is not a deed in the sense of observable actions, but rather in the sense of the choosing of a supreme maxim. Kant writes, "The propensity to evil is a deed in the first meaning (*peccatum originarium*), and at the same time the formal ground of every deed contrary to law according to the second meaning, [i.e. of a deed] that resists the law materially and is then called vice (*peccatum derivatum*)."⁵¹⁴ Both sorts of deeds carry moral significance, but the propensity to evil is based on a deed that cannot be directly observed: "The former [the sense of deed involving the propensity to evil] is an intelligible deed, cognizable through reason alone apart from any temporal condition; the latter [the sense of deed involving actions in the phenomenal world] is sensible, empirical, given in time (*factum phenomenon*)."⁵¹⁵

While these two senses of the term "deed" are conceptually distinguishable, it is important to note that there is a relationship that must exist between them. One cannot commit morally good deeds in the phenomenal world without having committed the supreme good deed in the noumenal world. At best, one can hope for deeds that are in conformity with the law, but not grounded in respect for the law. Similarly, one cannot commit misdeeds in the phenomenal world if one has made the morally correct choice in the noumenal world. If the supreme maxim one adopts is to privilege morality over self-

⁵¹⁴ *Ibid.*

⁵¹⁵ *Ibid.*

love as the foundational point of one's particular maxims, then one will never knowingly violate the moral law.⁵¹⁶

The propensity to evil, then, seems to be nothing other than an intelligible deed in which we exercise our free power of choice in a way that prioritizes self-love over the moral law. This intelligible deed informs our adoption of "supreme maxim" the basic volitional structure with which we formulate more particular maxims in the phenomenal world and then act upon.⁵¹⁷ Such a deed can rightly be thought of as a propensity because it precedes the adoption of particular maxims and instead instills in us, prior to our actions in the phenomenal world, the freely adopted framework within which we act. Like other kinds of propensities, then, the propensity to evil makes us susceptible to experience certain inclinations. In the case of the propensity to evil, these inclinations, if acted upon, would lead us to misdeeds in the second, material sense of "deed."

A formal proof for the universal propensity to evil?

Kant declares that one can presuppose that this propensity to evil is present in every human being. In explaining what he means in claiming that humans are evil by

⁵¹⁶ This is, of course, Kant's very definition for moral evil: "In view of what has been said above, the statement, "The human being is evil," cannot mean anything else than that he is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it" (Rel. 6:32).

⁵¹⁷ My discussion on this point is indebted to Lawrence R. Pasternack, who, in his book *Kant on Religion within the Boundaries of Mere Reason*, points out that the propensity to evil is separate from the adoption of a supreme maxim (pp. 115-117). Our supreme maxim is psychological and therefore phenomenal, whereas the intelligible deed is noumenal. We are nevertheless responsible for our supreme maxim because it is grounded in our noumenal choice. The propensity to evil is whatever makes us choose in such a way that its effects, in the phenomenal world, will be an immoral supreme maxim. The propensity to evil is inscrutable to us, whereas the formation of our supreme maxim can be traced to predispositions that we possess on account of our physical nature.

nature, he writes, ““He is evil *by nature*” simply means that being evil applies to him considered in his species...”⁵¹⁸ Insofar as an individual is a member of the human species, the property of “being evil” can apply to him or her. Kant goes on to assert, “we may suppose evil as subjectively necessary in every human being, even the best.”⁵¹⁹ With such a wide-ranging claim, Kant seems to anticipate a demand for a formal proof, an argument that will support the universality and necessity claimed in his statement that human being, by a sort of subjective necessity, possesses a propensity to evil. The universalistic and apparent a priori character of Kant’s terminology here has led several commentators to argue that Kant is indeed making an argument that is grounded in a priori reasoning, and not merely empirical observation. In reviewing some of Kant’s language, David Sussman writes, “Such claims do not appear to be merely empirical generalizations based on the ubiquity of human vice.”⁵²⁰ Sussman then, on the basis of this claim, constructs an argument from necessity that purports to show that “there is a basis of human evil that is simultaneously a necessary feature of human agency and a kind of active resistance to morality that each of us freely and continuously engages in.”⁵²¹ Similarly, in his paper entitled “An Alternative Proof to the Universal Propensity to Evil,” Pablo Muchnik claims that Kant does indeed provide a “transcendental” proof for the universal propensity to evil, which he draws out in his paper, that, in Muchnik’s

⁵¹⁸ Rel. 6:32.

⁵¹⁹ *Ibid.*

⁵²⁰ David Sussman, “Perversity of the Heart,” *The Philosophical Review*, Vol. 114, No 2 (Apr., 2005), p. 153.

⁵²¹ *Ibid.*

words, “goes a long way to justify the subjective necessity, universality, and a priori character of the propensity to evil...”⁵²²

Kant, at least at first, does not attempt to provide a formal argument for the universality of the propensity to evil. He goes on to write, “We can spare ourselves the formal proof that there must be such a corrupt propensity rooted in the human being, in view of the multitude of woeful examples that the experience of human *deeds* parades before us.”⁵²³ Kant then describes a series of evil deeds that different people exhibit in various parts of the world. He begins by describing “scenes of unprovoked cruelty in the ritual murders of Tofoa, New Zealand, and the Navigator Islands, and the never-ending cruelty...in the wide wastes of northwestern America...”⁵²⁴ Kant also discusses the kind of evil one would encounter in places where human beings exist in a more “civilized state (where its predispositions can be more fully developed).”⁵²⁵ Here, too, Kant contends that, if we look to human beings in the more “civilized” condition, “we must then hear out a long melancholy litany of charges against humankind.”⁵²⁶ The evil we might encounter here would be more subtle, but equally indicative of a propensity to evil as one would find elsewhere. Some of the evil deeds that prevail in more developed locations include: “secret falsity even in the most intimate friendship, so that restraint on trust in the mutual confidence of even the best friends is reckoned a universal maxim of prudence

⁵²² Pablo Muchnik, “An Alternative Proof to the Universal Propensity to Evil,” in Kant’s Anatomy of Evil (eds. Sharon Anderson-Gold and Pablo Muchnik), Cambridge University Press: New York, 2010, p. 116-143. For a similar approach to Kant’s claim for the universality and subjective necessity of evil in human beings, see Seiriol Morgan, “The Missing Formal Proof of Humanity’s Radical Evil in Kant’s ‘Religion,’” *Philosophical Review* 114 (2005), pp. 63-114.

⁵²³ Rel. 6:33

⁵²⁴ *Ibid.*

⁵²⁵ *Ibid.*

⁵²⁶ *Ibid.*

in social dealings; of a propensity to hate him to whom we are indebted, to which a benefactor must always heed; of a hearty goodwill that nonetheless admits the remark that “in the misfortune of our best friends there is something that does not altogether displease us”; and of many other vices yet hidden under the appearance of virtue, let alone those of which no secret is made, for to us someone already counts as good when *his evil is common is common to a class.*”⁵²⁷

There are at least two aspects of Kant’s strategy here that are worth noting. First, the evidence Kant upon which Kant relies is all drawn from experience. That is, Kant mentions different sorts of evil deeds that are based on actual deeds that one can encounter in experience, either through direct observation or through psychological prodding into one’s (or others’) motives. Second, Kant seems to suggest that, whether we admit it or not, we all already believe it to be the case that there exists a universal propensity to evil. Why else, Kant asks, would we deem it wise not to trust other people, even one’s best friend? The only possible response, Kant seems to suggest, is to say that we do not believe that human beings are trustworthy, no matter how well we may think we know someone.

Kant’s discussion of the universality of evil in human beings has generated critical commentary. Some recent scholarship, as noted above, has attempted to provide a formal proof where one seems to lack in Kant’s own work. I would argue, though, that Kant is not making especially robust claims to universality and necessity, as it may seem in these passages from the *Religion*. Further, I argue that more basic elements of Kant’s

⁵²⁷ Rel. 6:33-34

moral philosophy actually preclude us from attaching necessity, though not universality, to all human beings. I will argue for this by showing what Kant *cannot possibly be claiming* in those earlier-quoted passages. I will then show that Kant's broader account of moral evil *does not require* an a priori justification. Finally, I will provide my understanding of the reach and scope of moral evil in Kant's discussion, arguing that, problems of clarity notwithstanding, Kant's discussion of radical evil is not at odds with the basic features of Kant's account of moral responsibility.

Whatever else Kant's account of moral evil is claiming, it cannot possibly claim that human beings, when engaged in evil actions, cannot have acted otherwise. Crucial to Kant's moral philosophy is the idea that moral (or immoral) actions are imputable to the agent engaged in them. If, though, an agent does something that he or she is morally responsible for, then the cause of that action is his or her exercise of free choice at the noumenal level. An action is free only when there is the possibility that the agent could have acted otherwise; denying this would be a clear violation of the 'ought implies can' principle that Kant clearly holds. If all people are in fact evil, then it must mean that everyone has made a free, noumenal choice that accounts for the privileging of self-love over morality in the phenomenal world. Crucially, in each and every case of an immoral person, it must be true that he or she was free to order his or her will otherwise; namely, he or she could have privileged morality over self-love. Kant, therefore, cannot be claiming that all people are necessarily evil. That would imply that a good will (one with the proper volitional structure) is unachievable. If Kant is indeed making a claim of

necessity, it would have to be one consistent with the view that immoral agents could have acted otherwise.

Even so, one might argue, Kant does in fact claim that human beings are evil by necessity. The particular passage to which one might refer in order to make this point clear is the following: “[A]ccording to the cognition we have of the human being through experience, he cannot be judged otherwise [i.e. as moral], in other words, we may presuppose evil as subjectively necessary in every human being, even the best.”⁵²⁸ This, it seems, is the most emphatic claim to necessity that Kant makes about moral evil in human beings. Here, we should examine carefully how the word “necessary” is operating in the sentence. The phrase that Kant uses is “subjectively necessary,” but, importantly in my view, it appears in the sentence as a paraphrase of a preceding clause, namely “according to the cognition we have of the human being through experience, he cannot be judged otherwise.”⁵²⁹ The judgment that “all human beings are evil” rests, we can clearly read here, on the “cognition we have of the human being through experience.”⁵³⁰ In other words, by virtue of the fact that we regularly witness, in ourselves and in others, immoral acts, dispositions, or attitudes, we cannot help but presuppose, in any particular human being, the propensity to evil (or an evil will). This experience-based claim is the context for Kant’s discussion one paragraph later of the particular kinds of evils one may encounter when we observe the evil that happens in the world. It is subjectively necessary, *for us*, to view human beings as evil just as it is subjectively necessary for us

⁵²⁸ Rel. 6:32.

⁵²⁹ *Ibid.*

⁵³⁰ *Ibid.*

to view rocks as solids or ravens as black. This subjective kind of necessity is wholly consistent with the view that immoral human beings could in fact become moral.

With Kant's claim to necessity explicated, we can now turn to his claim that *all* human beings are in fact evil. The second troubling (perhaps more troubling) claim embedded in Kant's discussion of moral evil is that "even the best" human being can be presupposed as being evil. If even the best human being is evil, then what rational hope can anyone have of attaining true virtue? Importantly, as I've stressed above, Kant's claim cannot be that it is impossible for the human beings to be morally good. It is certainly possible that there will one day be a human being who does indeed achieve a good will, even if it has not yet occurred. In claiming that "even the best" human being can be presupposed to be evil, Kant is perhaps being a bit unclear. Kant clearly has not profiled all the likely candidates for the designation "morally good," so his claim cannot justifiably be making any definitive statements about the absence or limited presence of moral goodness in the world. Furthermore, Kant, according to his own account of the metaphysics of freedom, would not be able to determine definitively whether or not someone truly possesses a good will (the ultimate ground for our free choices are not accessible through inner or outer sense).

So what sense can be made out Kant's claim here? Kant, it seems, could be making one of two different claims in writing that even the best human being possesses a propensity to evil. First, Kant could be asserting that even good people (truly good people with a good will) possess the propensity to evil because they have inclinations that, if acted upon, would cause them to violate the moral law. Surely it is a reasonable

generalization to claim that all human beings have such inclinations because of our sensuous nature, which subjects us to inclinations toward happiness and, perhaps at times, against the moral law. Kant writes that human beings are “dependent on the incentives of his sensuous nature” which, if taken into our supreme volitional maxim as “*themselves sufficient* for the determination of his power of choice, without minding the moral law,” we “would then become morally evil.”⁵³¹ This might seem like a promising way to interpret Kant, but I would argue that it is an incorrect approach. Such a reading would contradict Kant’s considered view of the propensity to evil, which I above argued is an intelligible deed that prioritizes self-love over the moral law. Morally good people, though, are good only insofar as they have not made this intelligible choice.

Fortunately, there is a second possible interpretation of Kant that preserves his considered view of the propensity to evil (though perhaps betrays a bit of sloppiness in the writing of this passage). On this reading, Kant is hypothesizing that every living person, even the person who comes closest to conforming to the requirements of the moral law, has actually prioritized, in one way or another, self-love over morality. This would be the culmination of the kind of empirical generalizing that Kant does in the subsequent passage in the *Religion*. While, for reasons stated above, Kant is not really entitled to make such a wide-ranging claim, it is not an especially important claim for his broader account of moral evil. If there is actually someone, or some set of people, who have good wills, then Kant would simply have to adjust his more somber observations to

⁵³¹ Rel. 6:36.

accommodate the new information. This adjustment, though, would not cause his claims about moral imputation, moral evil, or freedom any difficulties.

Conclusion: The Intractability of Crisis

Given Kant's unflinching and negative assessment of human beings as they are, it is especially interesting to examine Kant's remarks about the nature and possibility of moral improvement. In the *Religion*, when Kant turns directly to the prospect of moral improvement in evil people, he expresses a kind of perplexity regarding its very possibility: "How is it possible that a naturally evil human being should make himself into a good human being surpasses every concept of ours. For how can an evil tree bear good fruit?"⁵³² Kant is not calling into doubt evil human beings' capacity to improve morally, but rather is claiming that we have difficulty tracing such a change with concepts. Since we can have no cognition of us as noumenal agents, and therefore have no theoretical access to the noumenal choice itself, there will always be something inexplicable and mysterious about human freedom at its very root. He claims that the "fall from good into evil...is no more comprehensible than the ascent from evil back to good..."⁵³³ Even if we cannot conceptually represent moral improvement (i.e. the change from a propensity to evil to a good will), we can recognize that the moral law applies to us, and, therefore, we must be capable of virtue: "[T]he command that we *ought* to become better human beings still resounds unabated in our souls; consequently, we must

⁵³² Rel. 6:44-45.

⁵³³ Rel. 6:45.

also be capable of it...”⁵³⁴ The ground for individual and collective moral should not, for Kant, be based on sentimental wishful thinking, nor on an inaccurately rosy picture of the real moral state of humanity, but rather on the findings in Kant’s critical philosophy, especially, I would contend, by experiencing directly the force of the moral law (through the fact of reason) and deriving its implications: the dignity and freedom of human beings.

⁵³⁴ Rel. 6:45.

BIBLIOGRAPHY

- Addison, Joseph. *Addison's Essays from the Spectator*. Elibron Classics: London, UK, 2007.
- Allison, Henry E. *Kant's theory of freedom*. Cambridge University Press: New York, NY, 1990.
- . *Kant's Transcendental Idealism: An Interpretation and Defense*. Yale University Press: New Haven, CT, 2004.
- . "On the Very Idea of a Propensity to Evil." *The Journal of Value Inquiry*, Volume 36, 2002, pp. 337-348.
- Americks, Karl. *Kant's Theory of Mind: An Analysis of the Paralogisms of Pure Reason*. Oxford University Press: New York, NY, 2000.
- Atwell, John E. "Kant's Notion of Respect For Persons." *Tulane Studies in Philosophy*, Volume 31, 1982, 17-30.
- Baxley, Anne Margaret. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge University Press: New York, NY, 2010.
- Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. The University of Chicago Press: Chicago, IL, 1960.
- Bernstein, Richard. *Radical Evil: A Philosophical Interrogation*. Blackwell: Malden, MA, 2002.

Budd, Malcolm. *The Aesthetic Appreciation of Nature*. Oxford University Press: New York, NY, 2002.

Burke, Edmund. *A Philosophical Enquiry into the Sublime and Beautiful*. Routledge Classics: New York, NY, 2008.

- - - . "The Sublime in Nature" in Kant's *Critique of the Power of Judgment*, ed. Paul Guyer. Rowland & Littlefield Publishers: Lanham, MD, 2003.

Caswell, Matthew. "The Value of Humanity and Kant's Conception of Evil." *Journal of the History of Philosophy*, Volume 44, Number 4, October 2006, pp. 635-663.

Clewis, Robert R. *The Kantian Sublime and the Revelation of Freedom*. Cambridge University Press: New York, NY, 2009.

Dean, Richard. "Glasgow's Conception of Kantian Humanity." *Journal of the History of Philosophy*, Volume 46, Number 2, April 2008, pp. 307-314.

- - - . *The Value of Humanity in Kant's Moral Theory*. Oxford University Press: New York, NY, 2006.

Denis, Lara. "Humanity, Obligation, and the Good Will: An Argument against Dean's Interpretation of Humanity." *Kantian Review*, Volume 15-1, 2010, pp. 118-141.

- - - . "Kant's Cold Sage and the Sublimity of Apathy." *Kantian Review*, Volume 4, March 2000, pp. 48-73.

Emerson, Ralph Waldo. *Emerson: Essays & Poems*. The Library of America: Library of American College Editions: New York, NY, 1996.

Formosa, Paul. "Kant on the Radical Evil of Human Nature." *Philosophical Forum*, Volume 38 (3), 2007, pp. 221-245.

- - - . "Kant on the Limits of Human Evil." *Journal of Philosophical Research*, Volume 34, 2009, pp. 189-214.

Frierson, Patrick R. *Freedom and Anthropology in Kant's Moral Philosophy*. Cambridge University Press: New York, NY, 2010 (paperback).

- - - . "Kant's Moral Pessimism" in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.

Gardner, Sebastian. "The Primacy of Practical Reason," from *A Companion to Kant*, ed. Graham Bird. Blackwell Publishing: Malden, MA, 2009.

Gerard, Alexander. *An Essay on Taste: To Which Are Annexed Three Dissertations On The Subject By De Voltaire, D'Alembert And De Montesquieu*. Kessinger Publishing's Rare Reprints: LaVergne, TN, 2009.

Guyer, Paul. *Kant and the Claims of Knowledge*. Cambridge University Press: New York, 1987.

- - - . *Kant and the Experience of Freedom*. Cambridge University Press: New York, NY, 1996.

- . *Kant*. Routledge: New York, NY, 2006.
- . “Kant’s Principles of Reflecting Judgment” in Kant’s *Critique of the Power of Judgment*, ed. Paul Guyer. Rowland & Littlefield Publishers: Lanham, MD, 2003.
- . “Problems with freedom: Kant’s Argument in Groundwork II and its subsequent emendation” in *Kant’s Groundwork for the Metaphysics of Morals: A Critical Guide*. Ed. Jens Timmermann. Cambridge University Press: New York, NY, 2013.
- . *Kant on Freedom, Law and Happiness*. Cambridge University Press: New York, NY, 2000.
- . “Setting and Pursuing Ends: Internal and External Freedom,” forthcoming.
- Herman, Barbara. *The Practice of Moral Judgment*. Harvard University Press: Cambridge, MA, 1996.
- Hill, Thomas. *Dignity and Practical Reason in Kant’s Moral Philosophy*. Cornell University Press: Ithaca, NY, 1992.
- . *Respect, Pluralism, and Justice*. Oxford University Press: New York, NY, 2000.
- Kant, Immanuel. *The Cambridge Edition of the Works of Immanuel Kant: Anthropology, History, and Education*, eds. Günter Zöllner and Robert B. Louden. Trans. Mary Gregor, Paul Guyer, Robert B. Louden, Holly Wilson, Allen W. Wood, Günter Zöllner, and Arnulf Zweig. Cambridge University Press: New York, NY, 2007.

- - - . *The Cambridge Edition of the Works of Immanuel Kant: Critique of Pure Reason*, ed. Paul Guyer and Allen Wood. Trans. Paul Guyer and Allen Wood. Cambridge University Press: New York, NY, 1998.
- - - . *The Cambridge Edition of the Works of Immanuel Kant: Critique of the Power of Judgment*, ed. Paul Guyer. Trans. Paul Guyer and Eric Matthews. Cambridge University Press: New York, NY, 2000.
- - - . *The Cambridge Edition of the Works of Immanuel Kant: Lectures on Ethics*, eds. Peter Heath and J.B. Schneewind. Trans. Peter Heath. Cambridge University Press: New York, NY, 1997.
- - - . *The Cambridge Edition of the Works of Immanuel Kant: Notes and Fragments*, ed. Paul Guyer. Trans. Curtis Bowman, Paul Guyer, and Frederick Rauscher. Cambridge University Press: New York, NY, 2005.
- - - . *The Cambridge Edition of the Works of Immanuel Kant: Opus Postumum*. Ed. Eckart Förster. Trans. Eckart Förster and Michael Rosen. Cambridge University Press: New York, NY, 1993.
- - - . *The Cambridge Edition of the Works of Immanuel Kant: Practical Philosophy*. Ed. Mary J. Gregor. Trans. Mary J. Gregor. Cambridge University Press: New York, NY, 1996.

--- . *The Cambridge Edition of the Works of Immanuel Kant: Religion and Rational Theology*. Eds. Allen W. Wood and George Di Giovanni. Trans. Allen W. Wood and George Di Giovanni. Cambridge University Press, New York, NY, 1996.

--- . *The Cambridge Edition of the Works of Immanuel Kant: Theoretical Philosophy 1755-1770*. Ed. David Walford. Trans. David Walford. In collaboration with Ralf Meerbote.

--- . *The Cambridge Edition of the Works of Immanuel Kant: Theoretical Philosophy After 1781*. Eds. Henry Allison and Peter Heath. Trans. Gary Hatfield, Michael Friedman, Henry Allison, and Peter Heath. Cambridge University Press: New York, NY, 2002.

Kirwan, James. *The Aesthetic in Kant*. Continuum: New York, NY, 2004.

Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge University Press: New York, NY, 1996.

Kneller, Jane. "Imaginative Freedom and the German Enlightenment" in Kant's *Critique of the Power of Judgment*, ed. Paul Guyer. Rowland & Littlefield Publishers: Lanham, MD, 2003.

Krueger, James. "Duties, Ends and the Divine Cooperation" in *Kant's Moral Metaphysics*. Eds. Benjamin J. Bruxvoort Lipscomb and James Krueger. De Gruyter: Boston, MA, 2010.

- Lipscomb, Benjamin. "Moral Imperfection and Moral Phenomenology in Kant" in *Kant's Moral Metaphysics*. Eds. Benjamin J. Bruxvoort Lipscomb and James Krueger. De Gruyter: Boston, MA, 2010.
- Louden, Robert B. "Evil Everywhere: The Ordinariness of Kantian Radical Evil: in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.
- Martin, Adrienne. "How to Argue for the Value of Humanity." *Pacific Philosophical Quarterly*, Volume 87, Issue 1, March 2006, pp. 96-125.
- McMullin, Irene. "Kant on Radical Evil and the Origin of Moral Responsibility." *Kantian Review*, Volume 18, Issue 01, March 2013, pp. 49-72.
- Michalson, Gordon E. "Kant, the Bible, and the Recovery from Radical Evil" in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.
- Muchnik, Pablo. "An Alternative Proof of the Universal Propensity to Evil" in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.
- Noggle, Robert. "Kantian Respect and Particular Persons." *Canadian Journal of Philosophy*, Volume 29, Number 3, September 1999, pp. 449-478.
- Pasternack, Lawrence R. *Kant on Religion within the Boundaries of Mere Reason*. Routledge: New York, NY, 2014.

- Paton, H.J. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. University of Pennsylvania Press: Philadelphia, PA, 1971.
- Rossi, Philip J. "Kant's "Metaphysics of Permanent Rupture": Radical Evil and the Unity of Reason" in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.
- Rousseau, Jean-Jacques. *The Basic Political Writings*. Ed. and Trans. Donald A. Cress. Hackett Publishing Company: Indianapolis, IN, 1987.
- Sedgwick, Sally. *Kant's Groundwork of the Metaphysics of Morals: An Introduction*. Cambridge University Press: New York, NY, 2008.
- Sensen, Oliver. "Kant's Conception of Inner Value." *European Journal of Philosophy*. Volume 19, Number 2, 2009, pp. 262-280.
- . "Dignity and the formula of humanity" in *Kant's Groundwork for the Metaphysics of Morals: A Critical Guide*. Ed. Jens Timmermann. Cambridge University Press: New York, NY, 2013.
- . *Kant on Human Dignity*. De Gruyter: Boston, MA, 2011.
- Shaw, Philip. *The Sublime*. Routledge: New York, NY, 2006.
- Sussman, David. "Something to Love: Kant and the Faith of Reason" in *Kant's Moral Metaphysics*. Eds. Benjamin J. Bruxvoort Lipscomb and James Krueger. De Gruyter: Boston, MA, 2010.

- - - . "Perversity of the Heart," *The Philosophical Review*, Vol. 114, No 2, Apr., 2005, p. 153-177.

Wood, Allen W. "Humanity as an End in Itself." In *Kant's Groundwork for the Metaphysics of Morals: Critical Essays*. Ed. Paul Guyer. Rowan and Littlefield, Publishers, Inc.: Lanham, MD, 1998.

- - - . "Kant and the Intelligibility of Evil" in *Kant's Anatomy of Evil*, eds. Sharon Anderson-Gold and Pablo Muchnik. Cambridge University Press: New York, NY, 2010.

- - - . *Kant's Ethical Thought*. Cambridge University Press: New York, NY, 1999.

- - - . *Kantian Ethics*. Cambridge University Press: New York, NY, 2008.

- - - . *Kant's Rational Theology*. Cornell University Press: Ithaca, NY, 1978.

Zuckert, Rachel. "Kant's Account of Practical Fanaticism" in *Kant's Moral Metaphysics*. Eds. Benjamin J. Bruxvoort Lipscomb and James Krueger. De Gruyter: Boston, MA, 2010.