

Georgia State University  
**ScholarWorks @ Georgia State University**

---

Philosophy Theses

Department of Philosophy

---

7-31-2006

# The Impossibility of Evil Qua Evil: Kantian Limitations on Human Immorality

Timothy Alan Crews-Anderson

Follow this and additional works at: [https://scholarworks.gsu.edu/philosophy\\_theses](https://scholarworks.gsu.edu/philosophy_theses)

 Part of the [Philosophy Commons](#)

---

## Recommended Citation

Crews-Anderson, Timothy Alan, "The Impossibility of Evil Qua Evil: Kantian Limitations on Human Immorality." Thesis, Georgia State University, 2006.

[https://scholarworks.gsu.edu/philosophy\\_theses/12](https://scholarworks.gsu.edu/philosophy_theses/12)

This Thesis is brought to you for free and open access by the Department of Philosophy at ScholarWorks @ Georgia State University. It has been accepted for inclusion in Philosophy Theses by an authorized administrator of ScholarWorks @ Georgia State University. For more information, please contact [scholarworks@gsu.edu](mailto:scholarworks@gsu.edu).

THE IMPOSSIBILITY OF EVIL *QUA EVIL*:  
KANTIAN LIMITATIONS ON HUMAN IMMORALITY

by

TIMOTHY A. CREWS-ANDERSON

Under the Direction of Melissa M. Merritt

ABSTRACT

Kant denies that evil *qua evil* can be an incentive to human beings. Is this a fact about what sorts of reasons human beings find interesting? Or, is it rooted entirely in Kant's notion of human freedom? I focus on key facets of Kant's system: human freedom, immorality and incentives. With an understanding of these concepts based in Christine Korsgaard's reading of Kant's moral theory, I argue that the impossibility of acting solely from evil *qua evil* is not rooted in human incentives and that if we were able to represent an unconditioned principle of immorality, we would have as powerful an incentive to act in accordance with it as we do to act in accordance with the categorical imperative. Finally, I argue that the impossibility of human beings' having evil *qua evil* as an incentive is grounded in the limited nature of our positive conception of freedom.

INDEX WORDS: Kant, ethics, morals, metaphysics, freedom, immorality, evil, psychology, incentive, moral feeling

THE IMPOSSIBILITY OF EVIL *QUA EVIL*:  
KANTIAN LIMITATIONS ON HUMAN IMMORALITY

by

TIMOTHY A. CREWS-ANDERSON

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of

Master of Arts

in the College of Arts and Sciences

Georgia State University

2006

Copyright by  
Timothy A. Crews-Anderson  
2006

THE IMPOSSIBILITY OF EVIL *QUA EVIL*:  
KANTIAN LIMITATIONS ON HUMAN IMMORALITY

by

TIMOTHY A. CREWS-ANDERSON

Major Professor:	Melissa Merritt
Committee:	Andrew Altman
	Andrew J. Cohen

Electronic Version Approved

Office of Graduate Studies  
College of Arts and Sciences  
Georgia State University  
August 2006

DEDICATION

To my wife,  
Victoria Anne Crews-Anderson

## ACKNOWLEDGEMENTS

This thesis has been greatly improved by comments and suggestions from my thesis director, Dr. Melissa Merritt, and thesis committee members, Dr. Andrew Altman and Dr. Andrew J. Cohen. I would especially like to thank Dr. Merritt, whose patient instruction in the art of philosophical writing has been indispensable. Additionally, I want to thank Dr. Sandra Dwyer, who has contributed greatly to my philosophical knowledge and experience. Finally, I want to thank Sean Aas who has offered helpful criticism of my ideas at every turn.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS.....	v
CHAPTER	
1 INTRODUCTION.....	1
2 THE FREEDOM, IMMORALITY, AND INCENTIVES OF HUMAN BEINGS.....	3
Human Freedom.....	3
Negative Concept to Positive Conception.....	3
<i>Wille</i> and <i>Willkür</i> .....	4
Human Immorality.....	7
The Basics.....	7
Fundamental Maxims.....	9
Human Incentives.....	13
The Predisposition to Good.....	13
The Moral Feeling.....	15
Note on Inclination and Freedom.....	21
The Moral Feeling and the Original Determination.....	22
Summary of Chapter 2.....	26
3 EVIL <i>QUA EVIL</i> .....	28
Diabolical and Malicious Beings.....	28
The Conceivability of a Human Psychological Incentive to Act from Evil <i>Qua Evil</i> .....	29
The Impossibility of a Human Agent's Legislation of an Alternative to the Moral Law.....	34



## CHAPTER 1: INTRODUCTION

Kant explicitly stakes out conceptual space for “evil *qua evil*” in *Religion within the Boundaries of Mere Reason* when he invokes the idea of a “diabolical being” or “evil reason” for which “resistance to the [Moral] law [is] itself...elevated to incentive.” He denies, however, that this is “applicable to the human being.”<sup>1</sup> Why is this the case? According to Kant, it is only possible for human beings (or any imperfectly rational beings) to act either from respect for the Moral Law or from sensual inclinations. What is the reason for this? Is this a psychological matter and simply a fact about what sorts of reasons human beings find interesting? Or, is it rooted entirely in Kant’s notion of human freedom? To answer these questions, attention will be focused on key facets of Kant’s ethical system: the precise nature of human freedom, how it is that human beings act immorally, human incentives, and finally the idea of diabolical and malicious beings. With a plausible understanding of these concepts based in Christine Korsgaard’s reading of Kant’s moral theory, I will argue that the impossibility of our acting solely from resistance to the Moral Law is not rooted in the psychology of human incentives and that if we were able to represent an unconditioned principle of immorality, we would have as powerful an incentive to act in accordance with it as we do to act in accordance with the categorical imperative. Finally, I will argue that the impossibility of human beings’

---

<sup>1</sup>Kant, Immanuel. *Religion within the Boundaries of Mere Reason And Other Writings*. Translated by Allen Wood and George di Giovanni. Cambridge: Cambridge University Press, 1998. 35. Hereafter referenced internally with the abbreviation “Rel.”

having evil *qua evil* as an incentive is grounded in the limited nature of our positive conception of freedom.

CHAPTER 2: THE FREEDOM, IMMORALITY, AND INCENTIVES  
OF HUMAN BEINGS

Section 1: Human Freedom

**Negative Concept to Positive Conception:** In the third section of *Grounding for the Metaphysics of Morals*, Kant defines the will as “a kind of causality belonging to living things insofar as they are rational.” He describes freedom as a “property of this causality that makes it effective independent of any determination by alien causes.” This is contrasted with “natural necessity,” which is the “property of the causality of all non-rational beings by which they are determined to activity through the influence of alien causes.” This “negative explanation” of freedom offers little insight, but, “there does arise from it a positive concept, which as such is richer and more fruitful.”<sup>2</sup> Kant proceeds analytically from the negative concept of freedom to a positive concept. The concept of freedom includes that of causality, and the concept of causality carries the concept of “laws according to which something that we call cause must entail something else – namely, the effect” (Gr. 446). Thus, on Kant’s view, the negative concept of freedom implicitly contains a concept of lawfulness. From this we can infer that freedom cannot be “lawless,” and that a free will, despite having the capacity to determine its actions without influence from the causal order of nature, must act in accordance with an unchanging rule (Gr. 446). Continuing, Kant argues that since natural necessity is a “heteronomy of efficient causes,” freedom of the will must be “autonomy, i. e., the property that the will has of being a law to itself” (Gr. 447).

---

<sup>2</sup> Kant, Immanuel. *Grounding for the Metaphysics of Morals*. Translated by James W. Ellington. Indianapolis: Hackett Publishing Company, Inc., 1981. 446. Hereafter referenced internally with the abbreviation “Gr.”

From a negative concept of freedom, then, we move to a positive conception that Lewis White Beck describes as “the effectiveness of the legislation of pure practical reason [*Wille*] and the ability to undertake actions in accordance with and because of (out of respect for) this law.”<sup>3</sup> Thus, we have established two important features of an agent that possesses a free will. First, such an agent must be able to determine actions independently from sensory (alien) influences, and second, she must be able to legislate for herself a law according to which she has an incentive to act.

***Wille and Willkür:*** With these characteristics in hand, we may easily proceed to the constitutive elements that an agent must possess if her will is to be free. Kant describes two separate elements of the faculty of will, *Wille* and *Willkür*, which I will also refer to as capacities.<sup>4</sup> Henry Allison offers a helpful interpretation. The term *Wille* has two senses, broad and narrow. In its narrow sense, *Wille* signifies the legislative function or capacity of a “unified faculty of volition,” while *Willkür* is the executive capacity of this faculty – the power to choose. *Wille* in its broad sense indicates the faculty taken as a whole and includes both *Wille* in its narrow sense and *Willkür*.<sup>5</sup> In a passage from *Metaphysics of Morals*, Kant states:

A principle that makes certain actions duties is a practical law. A rule that the agent himself makes his principle on subjective grounds is called his *maxim* [. . . which is] a *subjective* principle of action, a principle which the subject himself makes his rule. A principle of duty, on the other hand, is a principle that reason prescribes to him absolutely and so objectively [. . . .] Laws proceed from the will, *maxims* from choice. In man the latter is a capacity for free choice; the will, which is directed to nothing beyond the law itself, cannot be called either free or unfree, since it is not directed to

---

<sup>3</sup> Beck, Lewis White. “Five Concepts of Freedom in Kant.” In *Philosophical Analysis and Reconstruction: Festschrift for Stephen Körner*, edited by T.J. Szrednicki. Dordrecht: Martinus Nijhoff, 1987, 35-51. p. 36.

<sup>4</sup> For my purposes, I will use the term “will” to mean the unified faculty and the German terms to indicate its constitutive elements.

<sup>5</sup> Allison, Henry. *Kant’s Theory of Freedom*. New York: Cambridge University Press, 1990. p. 129. Hereafter referenced with the abbreviation “Allison.”

actions but immediately to giving laws for the maxims of actions (and is, therefore, practical reason itself). Hence the will directs with absolute necessity and is itself *subject to* no necessitation. Only choice can therefore be called *free*.<sup>6</sup>

In addition to showing the distinction between *Wille* and *Willkür*, the passage also nicely expresses the difference between a practical law and a maxim. The maxim as a “subjective...principle which the subject himself makes his rule [...and which] proceeds from *Willkür*,” is a matter of choice while the practical law, which is “prescribe[d]” to an agent “absolutely and so objectively [...] proceeds from the *Wille*,” is not. Kant identifies *Wille* with “practical reason itself.”

In Kant’s system to say that reason is practical is to say that it is a faculty, “which is to have influence on the will” (Gr. 396). In contrast to Hume, who denies that reason can directly influence the will, it is the faculty by virtue of which an agent may, through an exercise of rational thought, determine her will. Practical reason itself has two fundamental employments with which an agent may make such determinations. When an agent becomes aware of a desired object or end, she may employ this faculty together with her *a posteriori* theoretical knowledge of how the phenomenal world is ordered in order to fulfill that desire. This is practical reason in its empirical employment, and by virtue of this application an agent is presented with hypothetical imperatives or “practical precepts.”<sup>7</sup> These hypothetical imperatives “determine the conditions of the causality of a rational being – as an efficient cause – merely in regard to the [desired] effect” (CpR 20). Their validity is conditioned by the desires of the agent, the agent’s theoretical knowledge, and the sufficiency of the agent’s efficient causal powers in the phenomenal

---

<sup>6</sup> Kant, Immanuel. *Metaphysics of Morals*. Translated by Mary Gregor. Cambridge: Cambridge University Press, 1991. 226. Hereafter referenced internally with the abbreviation “MM.”

<sup>7</sup> Kant, Immanuel. *Critique of Practical Reason*. Translated by Werner S. Pluhar. Indianapolis: Hackett Publishing Company, Inc., 2002. 20. Hereafter referenced internally with the abbreviation “CpR.”

world (CpR 20). Practical reason in its pure employment is an exercise of practical reason in which the agent determines her will “by the mere form of the practical rule without presupposition of any feeling, and hence without presentations of the agreeable or disagreeable as the matter of the power of desire, the matter which is always an empirical condition of principles” (CpR 24). Pure practical reason is a strictly *a priori* employment of practical reason by which an agent is presented with practical laws or imperatives that are unconditioned and categorical.

A question arises here. Does Kant mean that *Wille* is practical reason in both its pure and its empirical employment? Since Kant speaks about a practical law as being a “principle that makes certain actions duties” and mentions that *Wille* is “directed to nothing beyond the law itself,” the above passage from *Metaphysics of Morals* seems to suggest that the *Wille* is only *pure* practical reason. Allison, however, reads *Wille* to mean both pure and empirical practical reason. He argues that an agent through the exercise of *Wille* is presented with laws as either practical precepts or categorical imperatives.<sup>8</sup> Regardless of how we resolve this question, it is clear that an agent has two possible employments of practical reason from which she can generate maxims. Whether we ascribe the term *Wille* to both or only to the pure employment is not crucial for my argument. For simplicity, I will use the term *Wille* to mean only pure practical reason.

Though my reading diverts slightly from Allison’s with regards to whether *Wille* should be thought to include practical reason in its empirical employment, his interpretation remains helpful. He suggests a plausible description of the roles these two capacities play in an agent’s determination of her will. It is through an exercise of her

---

<sup>8</sup>Allison 130

*Wille* that the moral law is presented to an agent as a categorical imperative. The agent then has two possible grounds for the generation of maxims: a hypothetical imperative from empirical practical reason and the categorical imperative from *Wille*. It is by merit of her *Willkür* that the agent then generates subjective rules or maxims from these imperatives and then freely chooses the maxim according to which she will act. Though *Wille* allows the agent to legislate the moral law, this capacity does not suffice to allow her to choose whether to act in accordance with it. It is her *Willkür*, rather, that provides her with the ability to generate subjective rules and to choose which of these rules she will follow. *Willkür*, however, cannot play a role in the legislation of the moral law nor can it act as ground for the presentation of practical precepts as hypothetical imperatives. Through *Willkür*, the agent can only choose from those imperatives presented by virtue of *Wille* or empirical practical reason. It is only the agent taken as a whole, then, that can be considered to legislate the moral law for herself, to be autonomous, and to have the power to choose freely.<sup>9</sup>

## Section 2: Human Immorality

**The Basics:** In general, Kant considers an immoral action to be any action that is not in accordance with duty, that is done for any reason other than from respect for duty, or both.<sup>10</sup> This entails that it is possible for an agent to commit an action that is in accordance with duty but still not be moral. This immediately presents a problem, which Kant describes in the Second Section of the *Grounding for the Metaphysics of Morals*. Since we cannot directly observe an agent's thoughts, we can never know the true

---

<sup>9</sup> Allison 130-131.

<sup>10</sup> By "action" or "act" I mean roughly a deliberate, intentional, and morally significant deed performed by an agent.

reasons for her actions. Therefore, we can never in principle know from empirical observation whether another agent is acting morally. This is not, however, merely a problem of a lack of epistemic access to other minds. An agent cannot even say with certainty that she herself is acting morally because there is always a possibility that a reason, hidden even from the agent herself, other than respect for duty is acting as an incentive for an otherwise dutiful action (Gr. 406-407). This lack of certainty, however, does not necessarily arise in saying that an agent is immoral (Rel. 20). An explanation of why this is so will provide a good framework for a discussion of Kant's account of human immorality and "radical evil."

Kant holds that we can indeed know with certainty that an agent is immoral. To explain this somewhat startling claim, we need first to give a brief account of Kant's theory of moral reasoning and free choice. As mentioned above, *Willkür* is a power that an agent exhibits when she freely chooses maxims that she has generated from imperatives presented to her through either empirical practical reason or *Wille*. Important here is that for anything to be an option for choice it must first be incorporated into a maxim through *Willkür*. From this, we can infer from any particular action that an agent commits, that she has generated and selected a maxim. The ground or basis of the selected maxim that dictates a particular action determines whether that action is moral or immoral. Thus, it can be said that the morality of an action is determined by the nature of the maxim that determined it. As Kant indicates in the *Grounding for the Metaphysics of Morals*, "An action done from duty has its moral worth, not in the purpose that is attained by it, but in the maxim according to which the action is determined" (Gr. 399). If the maxim of a particular action has as its ground a categorical imperative, then the maxim is



moral as is the action that it determines. In order to evaluate the moral status of a particular action (and consequently the agent herself), we must be able to infer the ground of the maxim according to which the agent is acting. Since there are two possible grounds for any maxim that determines an action that accords with duty, and since it is impossible to infer with any certainty which of these two is the actual ground, it is impossible to know whether an agent has acted solely from duty. If we are to make good on Kant's claim that it is possible to know with certainty that an agent is immoral, we must show how it is possible to infer the ground of a maxim that determines an act that is not in accordance with duty.

**Fundamental Maxims:** On Kant's view, there are only two possible grounds from which an agent by exercise of her *Willkür* may draw imperatives for maxim creation, empirical practical reason and *Wille*. Korsgaard offers an interpretation of this that nicely illustrates this point. On her reading, given a particular action by an agent, we can ask the agent for the reason that she committed the given act. Suppose that an agent has told a lie. When we ask her why she has lied, she may answer that she did it to conceal the fact that she made a mistake. We may then ask her why she tried to conceal a mistake. Her answer to this question might be something to the effect that she wanted to maintain the appearance of competence or something of the sort. The point here is that any answer she gives would admit of further questioning regarding her reasons, and the possibility of an infinite regress arises. In order to block this regress, a fundamental reason for which there is a strong psychological incentive needs to be identified. In this case, this questioning could continue until she eventually answered that she was acting to protect what she thought was her self-interest. Self-interest, Korsgaard suggests, suffices

to block the regress. In Kantian terms, this appeal to self-interest is the generation and selection of a maxim that expresses what Kant calls a “subjective ground” or “supreme principle” based in self-interest as determined by practical reason in its empirical employment.<sup>11</sup>

Now suppose instead that our agent has told the truth and that in telling the truth, she has angered friends and family and significantly harmed her self-interest. We might again begin the process of questions that will eventually identify the fundamental reason for her act. In this case, her answer would reduce to something like, “It was the right thing to do,” which on Korsgaard’s view also suffices to block the regress. This appeal to the “right thing to do” is the generation and selection of a maxim that has as its source *pure* practical reason. It is a maxim that expresses the supreme principle based in morality as determined by practical reason in its pure employment. The choice that an agent makes between these two fundamental principles reflects the basic moral disposition of the agent.<sup>12</sup> Now, this basic disposition, if the agent is to be held responsible for it, must be a matter of choice, and if it is to be a matter of choice, she must incorporate it into a maxim through *Willkür*. This disposition, then, in addition to being the general basis for maxim generation and selection, is also itself a maxim.

Keeping with Korsgaard’s helpful reading, since there are only two possible grounds from which disposition-defining fundamental maxims may be generated, the difference between a moral basic disposition and an immoral one comes down to how the agent chooses to prioritize these two fundamental maxims, the prescriptions of which often (but not always) conflict. If the agent chooses to place those maxims that she

---

<sup>11</sup> Korsgaard, Christine M. “Morality as Freedom.” In *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996. p. 163-164. Hereafter referenced with the abbreviation “Korsgaard.”

<sup>12</sup> Korsgaard 163-164

generates from hypothetical imperatives in a secondary position to those which are grounded in the moral imperative from *Wille*, then her basic disposition is moral. If however, she places the maxims grounded in the moral imperative in a secondary position to those grounded in empirical practical reason, then the character is immoral. To make this a bit clearer, we can express the fundamental moral maxim as follows: I will act in accordance with morality, and I will act in my own self-interest so long as it does not interfere with my acting in accordance with morality. The fundamental immoral maxim simply reverses the priority: I will act in my self-interest, and I will act morally so long as it does not interfere with my acting in my self-interest.<sup>13</sup> The fact that this foundational maxim affects all other maxims selected through an agent's *Willkür* (indeed it makes all maxims generated by an agent immoral) is why Kant refers to human immorality as "radical evil."

To fulfill Kant's claim that it is possible to identify through inference the presence of the immoral maxim, we need only consider the implications that these two fundamental maxims have for the acts that would (or could) actually be committed by an agent. Let us suppose that an agent is genuinely moral; she has a moral character by virtue of the fact that the subjective ground of her *Willkür* is indeed the fundamental moral maxim. We can, through exhaustive observation of her entire life, establish that she never commits an act that is contrary to duty. But can we from this infer the presence of the moral maxim? Indeed we cannot, for it is entirely possible that she has decided that it is in her self-interest to act *as if* she has a genuinely moral character. Were this to be the case, the maxim under which she appears to be acting is not at all the maxim under

---

<sup>13</sup> Korsgaard 165

which she is really acting. It is impossible in principle to eliminate this possibility through empirical observation, which is, unfortunately, the only method available to us.

Let us suppose, however, that we observe an agent commit even one single act that is contrary to duty. In this case, we can immediately infer that the agent has selected as her “supreme principle” or “subjective ground” the fundamental immoral maxim. The reason for this is simple. If it were the case that she were operating under the fundamental moral maxim, it would be impossible for her to commit an act<sup>14</sup> contrary to duty. The fact that she has committed even one such immoral act reveals her selection of the fundamental immoral maxim. Of interest is the fact that this inference is *a priori*. While the observation of the immoral act is empirical, the inference from the fact of an immoral act’s having been committed to the presence of the fundamental immoral maxim proceeds independently from experience (Rel. 20). This inference is made valid by the fact that, according to Kant’s theory, the presence of the fundamental immoral maxim is the only possible condition under which a human being can commit an action that is contrary to duty.

The next apparent question is why a human being would take an interest in being immoral; that is, why would a human being find an incentive to adopt the fundamental

---

<sup>14</sup> Clearly this does not take into account the possibility of accidents or mistakes in reasoning. Accidents, of course, do not count as “acts” as defined above. However, with regards to mistakes in reasoning or errors that an agent might make about what her duty actually is, there is some cause for concern. In *Metaphysics of Morals* Kant appears to concede that it is possible for an imperfectly rational being to make this kind of mistake (MM 401). This seems to clearly raise an epistemic problem regarding whether a particular action that is not in accordance with duty is committed because of the presence of an immoral maxim (in which case the agent has a correct concept of duty and chooses not to give it priority in her determination) or because of a mistake in reasoning with regards to what her duty is. This would seem to be a possible problem for Kant’s ethical theory but not necessarily for my interpretation. If by hypothesis, we assume that we have a genuine instance of an agent’s committing an action contrary to duty while possessing a correct concept of that duty, then we can infer *a priori* the presence of the fundamental immoral maxim. By contrast, if we assume that we have a genuine instance of an agent’s committing an action that accords with duty while possessing a correct concept of duty, we still cannot infer the presence of the fundamental moral maxim.

immoral principle? The answer to this is as simple as it is obvious: because it is in a person's self-interest to do so. The really interesting question, then, is why would a person take an interest in acting *against* their self-interest, or more precisely, how is it that human beings have an incentive to select a fundamental maxim that has the potential to determine actions that are not in that person's self-interest? What could compel a person to consider the fundamental moral maxim to be a choice? To offer an answer to this question will require a brief discussion of incentive in Kant's system.

### Section 3: Human Incentives

**The Predisposition to Good:** In his account of human evil, Kant describes what he terms the "original predisposition to good in human nature" (Rel. 26). This predisposition may be further divided into three different "elements of the determination of the human being" (Rel. 26). These elements are:

1. The predisposition to the *animality* of the human being, as a *living being*;
2. To the *humanity* in him; as a living and at the same time *rational* being;
3. To his *personality*, as a rational and at the same time *responsible* being. (Rel. 26)

What does it mean for a human being to bear these predispositions? To say that we are predisposed to *animality*, for Kant, means that, as a matter of psychological fact that is due to our being a living thing, we are compelled by such things as self-preservation, species propagation, and community. Kant holds these drives or incentives to be more or less instinctual, describing them as "physical or merely *mechanical*" and as not having "reason at its root at all" (Rel. 26, 28). If we consider this physical self-interest as also involving "comparison (for which reason is required)," we evoke the predisposition to

*humanity*. From the notion of the instinctual self-interest of our animality “originates the inclination *to gain worth in the opinion of others*” (Rel. 27). I understand this to mean that from the physical self-love associated with the predisposition to animality along with the rational capacities of a human being, there arises a more complex second-order notion of self-interest. This notion stems from a human agent’s awareness of the success of other human agents in achieving the ends associated with their own physical self-interests. The agent’s predisposition to humanity psychologically compels her to evaluate her successes in terms that are relative to the successes of others and to the esteem that she receives from others. As Kant says, “only in comparison with others does one judge oneself happy or unhappy” (Rel. 27). If, however, we reflect on ourselves as not only *rational* but also as morally *responsible*, we bring to mind the predisposition to *personality*. Kant defines this predisposition as “the susceptibility to respect for the moral law *as of itself a sufficient incentive to the [Willkür].*” Kant identifies this susceptibility with the “moral feeling” (Rel. 27).

A plausible way to make sense of the elements of the human predisposition to good is to say, first, that the predispositions to animality and humanity are the material of practical reason in its empirical employment. They are incentives for a use of reason that is more or less compatible with the Humean notion in that they provide the ends for ends-means calculations: I want to live, so I will avoid joining the military. Or, I want to score the highest grade on the exam, so I will cheat. These two elements, then, do not necessarily motivate us to commit any particular act; rather, they predispose us to desire certain ends or goals. Moreover, if we choose to allow our desires for these ends to determine our will at the expense of our acting morally, then these predispositions

constitute the psychological basis for our taking an interest in the fundamental immoral principle. Empirical practical reason generates hypothetical imperatives that prescribe the measures we should take in order to achieve the ends that arise from our predisposition. The particulars of these means depend on contingent facts about the environment. If it is cold outside, my empirical practical reason may prescribe that I wear a coat to ensure my physical comfort. If, however, it is a particularly hot day, consideration of the same end would cause my empirical practical reason to prescribe that I wear cool and loose-fitting clothing. That the ends to which we are predisposed by our animality and rationality have influence on us is somewhat obvious. What is somewhat problematic for Kant's theory is why, in the face of such compelling incentives as pleasure and self-interest, an agent would ever opt for a maxim that is generated without consideration of these incentives at all. Put another way, in what sense can the law legislated through *Wille* be seen as an interesting option for *Willkür*? This question goes to the heart of Kant's account of how pure reason can be in itself practical.

**The Moral Feeling:** The third element of the predisposition to good is entirely different from the other two. Kant states that the "idea of the moral law alone, together with respect that is inseparable from it is [...] personality itself" (Rel. 28). The predisposition to *personality* is then a predisposition to respect for the Moral Law. For an agent to have *respect* for the Moral Law is for her to be in a state of having realized that the Moral Law applies to her or, more precisely, to have "consciousness of the subordination of [her] will to a law without the mediation of other influences upon [her] sense" (Gr. 401n). So to say that an agent has respect for something entails that the agent realizes that the object of respect carries determining force on *Willkür* regardless of the

effects of the acts that the object might dictate. Thus, the predisposition to personality or “moral feeling” is not simply respect for the Moral Law; rather, the “moral feeling” is a sort of psychological property that a human agent has that allows respect for the Moral Law to determine the will. In *Metaphysics of Morals*, Kant develops the notion of “moral feeling.” He characterizes it as one of several<sup>15</sup> “moral endowments such that anyone lacking them could have no duty to acquire them” (MM 399). The reason that there is no such duty is because “they lie at the basis of morality, as *subjective* conditions of receptiveness to the concept of duty, not as objective conditions of morality.” Indeed this receptiveness is a necessary psychological feature of a human agent if she is to be a moral agent at all. Kant asserts, “No man is entirely without moral feeling, for were he completely lacking in susceptibility to it he would be morally dead” (MM 400). If as a matter of psychological incentive, human beings were not susceptible to the concept of duty as being sufficient for determining the will, the whole idea of human moral responsibility would be undermined.

Kant further describes the moral feeling as a “susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent with or contrary to the law of duty” (MM 399). He states, somewhat mysteriously, “Every determination of [*Willkür*] proceeds *from* the representation of a possible action *to* the deed through the feeling of pleasure or displeasure, taking an interest in the action or its effect” (MM 399). This language seems on its surface to be somewhat problematic. Talk of feelings and such should raise the hackles of a proponent of an essentially rational ethical theory like that of Kant. He is, however, careful to distinguish between feeling that is “*sensibly*

---

<sup>15</sup> These endowments are “*moral feeling, conscience, love of one’s neighbor, and respect for oneself (self-esteem)*” (MM 399).



*dependent*” and feeling that is “*moral*” (MM 399). The moral feeling is unique in that it is not sensual. Furthermore, while sensibly dependent feeling “precedes the representation of law,” the pleasure associated with moral feeling “can only follow upon [the representation of the law]” (MM 399).

A reasonable interpretation of this is to say that when a human agent carries out a determination of *Willkür* according to the maxim derived from the fundamental immoral principle, she represents the possible action with a focus on its expected *effect* or end. Since her concern for this end is something to which she is predisposed by virtue of her rational animality, this representation is accompanied by a sensibly dependent feeling.<sup>16</sup> This feeling arises without the agent’s having yet considered the Moral Law; thus, it “precedes representation of the law” (MM 399). Now, we must keep in mind that the determination is grounded in the fundamental immoral principle; therefore, it is already the case that the agent by exercise of her *Willkür* has freely chosen the maxim that a sensibly dependent feeling associated with the effect of an act is a sufficient condition for a determination of her will. After this feeling has arisen, a representation of the Moral Law follows. This representation of the Moral Law comes about by virtue of the agent’s *conscience*<sup>17</sup>, which Kant describes as “practical reason holding man’s duty before him for his acquittal or condemnation in every case that comes under a law” (MM 400). He considers conscience to be something that as a matter of psychological fact every human agent has “originally” (MM 400). Following this consciousness of the Moral Law, a

---

<sup>16</sup>I understand those feelings describable as “sensibly dependent” to include not only actual sensual feelings but also expectations of actual sensual feelings. It is consistent to say that the feeling that arises when we expectantly represent ourselves as experiencing an actual sensual feeling is itself sensibly dependent.

<sup>17</sup>On my reading, while conscience plays a role in an agent’s judging whether a particular maxim is in accordance with the Moral Law, it is not itself a factor in the psychology of incentive because it does not have influence on why an agent decides in general to make the Moral Law sufficient for determining her will. I revisit the Kantian conscience below (page 30).

moral feeling of displeasure follows *regardless of whether the act is in accordance with or contrary to duty*. This is because the agent, insofar as she is aware that the act is not done from respect duty, realizes that she has chosen to act immorally. Based on the agent's fundamental maxim, however, this moral displeasure (guilt) is insufficient to counter the conditioning of morality by self-interest.

If on the other hand, a human agent carries out a determination of *Willkür* in accordance with the fundamental moral maxim, she represents the possible action with a focus *only on the action* and not at all on its effect. Since she is considering the action completely independently of its effects, no sensibly dependent feeling arises. Having no regard for the effects of the action, the only remaining basis for decision is whether the action is in accordance with duty. By virtue of her conscience, the agent is then presented with a representation of the Moral Law. If the action under consideration is in accordance with the law, a moral feeling of pleasure arises, and *Willkür* completes the determination with the agent's undertaking of the action. If, however, the action under consideration is contrary to the law, a moral displeasure is experienced, and the agent opts not to carry out the action. A crucial point needs to be made here. What I am describing is not an account of why an agent decides to carry out a moral act and refuse an immoral act. An agent acting from respect for the Moral Law decides one way or another solely because a particular act is in accordance with or contrary to her duty. The account that I am presenting of a feeling associated with acting morally is meant only to offer an essentially psychological explanation of how a human agent takes an interest in acting morally. Put differently, this is a description of why a human being decides to

make a moral imperative a sufficient reason for a determination of her will; it is not an account of why a human being chooses to act from respect for that imperative.

A potentially significant problem arises from this reading. My account shows how the moral feeling arises only in those determinations of the *Willkür* that are derivative of the fundamental maxim, which on my reading is presupposed. I am appealing to the basic disposition of the agent to account for how *Willkür* can have the Moral Law as an incentive. This would seem to be a problem since the nature of a person's basic disposition is itself a matter of choice, and therefore a determination of *Willkür*. Kant, indeed, states “**Every** determination of [*Willkür*] proceeds *from* the representation of a possible action *to* the deed through the feeling of pleasure or displeasure, taking an interest in the action or its effect” (MM 399, my added emphasis). How are we to make sense of this sentence in terms of the original<sup>18</sup> determination of *Willkür*, by which I mean the choosing of the fundamental maxim? Korsgaard offers a compelling account of this original determination that should be of help here.

Korsgaard has us consider a “purely rational will” from “a position of spontaneity,” which is meant as a conception of a will considered independently of space, time, and natural (or phenomenal) causality. While in this “position” the free will is called upon to choose its most fundamental principle.”<sup>19</sup> This choice amounts to a selection by the will of a principle that will determine what sorts of things it will count as reasons. Assuming that the will must have fundamental reasons of some sort to block the

---

<sup>18</sup> By “original,” I do not mean anything like “first in time” because this determination is made by virtue of the fact that a free will is in some sense a member of the intelligible world, and time is a feature of the phenomenal world. This discussion will make use of temporal and spatial terminology, but this should not be taken as descriptions of empirical time and space; rather, they reflect merely a limitation of human language in discussing the subject.

<sup>19</sup> Korsgaard 164

sort of regress mentioned above, on what basis would a free will in such a position decide? Enter Kant's notion of incentive, which Korsgaard characterizes as "something that makes an action interesting to [the agent]." <sup>20</sup> Since the will has not yet decided what sorts of things it will consider as reasons, these incentives do not yet count as such. They do, Korsgaard suggests, limit the options from which this original determination may be made, and while the (human) will cannot choose not to feel compelled by these incentives, it can choose "the order of precedence among the different kinds of incentives to which [it] is subject." These incentives, of course, are the elements of the predisposition to good. On what basis, then, would a free will in the position of spontaneity choose? This brings us back to the question at hand. Why would a human being take an interest in choosing a fundamental maxim that might very well entail the necessity of acts against her self-interest?

A purely rational will must choose a principle upon which to proceed. It has two possible grounds for such a principle, either interest in the self as conceived empirically in terms of the laws of nature or *Wille* and its Moral Law. Korsgaard argues that a purely rational will in the position of spontaneity must choose a principle without phenomenal experience to provide empirical content. The Moral Law is a compelling option because it is a lawful principle without such content. Indeed, if a purely rational will chooses the Moral Law as the primary consideration in its original determination, then, on Korsgaard's view, it remains in the position of spontaneity. If a purely rational will opts for the fundamental immoral principle, it relinquishes its position of spontaneity and subjugates itself to sensory inclination. As Korsgaard puts it, "A constraint on its choice is acquired," and for a purely rational will in the position of spontaneity, there is no

---

<sup>20</sup> Korsgaard 165

incentive for this. The incentives that arise from the predispositions to animality and humanity do not gain their force until the purely rational will departs freely from the position of spontaneity.<sup>21</sup>

**Note on Inclination and Freedom:** Special care should be taken here. Once a purely rational will has departed from the position of spontaneity, it is confronted with the full will-determining force of the phenomenal world. Human beings are predisposed to certain sorts of ends rooted in our rational animality, and considerations of achieving those ends in the world of experience are tremendously powerful to us. How can we say that these inclinations, rooted as they are in empirical practical reason, can then be a matter of free choice? If we assume that these ends are ours by virtue of instinct or psychology, how can we say that we are not causally determined to perform the means to these ends? How can such acts be morally imputable?

Kant, of course, denies that we are causally determined by our sensory inclinations and holds that we are indeed responsible for choosing to act in accordance with them. The basis of this claim is that any determination of the will is an act of an agent by virtue of her *Willkür*. For a particular end to be selected, a particular act to be carried out, or for a particular incentive to influence a determination of the will, an agent through *Willkür* must incorporate that end, act, or incentive into a maxim which she must then choose freely. So while we may be predisposed by virtue of our rational animality to find certain ends compelling as incentives, we must *choose* to act on these incentives. We may feel a powerful impulse to act towards the ends of our own interest, pleasure, or benefit. Indeed, these impulses arise involuntarily from our physical selves as rational animals. Nonetheless, we must first make it our maxim to act from them if they are to

---

<sup>21</sup> Korsgaard 166

play any role in determining our will. Even then, the role that they play is limited. At most, they are only the ends, consideration of the achievement of which generates the contingent content of empirical practical reason, which is only one possible source from which *Willkür* is presented imperatives. Consequently, even when we make it our maxim to achieve these ends, it is not the self-interest, pleasure or benefit nor even the desire for these ends that actually determines the will; rather, it is through our capacity of *Willkür*, our free power of choice, that this becomes possible. Inclination acts as a “constraint on choice,” but it does not act as a *restraint*.

**The Moral Feeling and the Original Determination:** Putting this back into the context of the discussion of moral feeling, we can easily see how it is possible from the position of spontaneity for the original determination to be made in favor of the fundamental moral maxim. In the absence of the incentives of inclination, a purely rational will has no reason to choose other than in favor of the Moral Law. Clearly, however, the human will is not purely rational. By virtue of this fact, we do have a powerful psychological incentive to adopt the fundamental immoral maxim. So why then would a human will opt for the fundamental moral maxim? What can count as a basis for an imperfectly rational human will to take an interest in choosing to identify with the purely rational moral standpoint? Korsgaard suggests that this basis may be found in Kant’s distinction between the phenomenal and noumenal world. By constraining our *Willkür* with determination by the phenomenal world, in a very real sense, we put our will at the disposal of the causally-determined physical world. Assuredly, we choose to do this, but this choice, in effect, is a choice to act as if we are not free. If however, we choose the fundamental moral maxim, we are opting to maintain our will in a position of

freedom, and thereby are we able to fully represent ourselves as members of the noumenal (intelligible) world.<sup>22</sup>

Although Kant famously denies the possibility of theoretical knowledge of the noumenal world, he argues that we can form a conception of it as the ground of the phenomenal world and as determining its nature. Based on this conception, an agent's representation of herself as member of the noumenal world entails a conception of herself as "mak[ing] a real difference to the way the phenomenal world is."<sup>23</sup> Certainly, we are not warranted in any claims about what precisely this difference might be, but by opting for the fundamental moral maxim, an agent can represent herself as contributing not only to the "merely natural, ordering of the sensible world, which" can be "accounted for by other forces in the noumenal world," but also to the "rational" ordering. Adopting the immoral maxim, however, is to accept that "although you are free, you could just as well not have been. Your freedom makes no difference."<sup>24</sup>

It is plausible to suggest the particular sort of pleasure that Kant evokes in his discussion of "moral feeling" is rooted in the fact that acting from respect from the Moral Law uniquely allows an agent to represent herself as being a member of the noumenal world. In his discussion of the predisposition to personality in *Religion within the Boundaries of Mere Reason*, Kant writes that the Moral Law "is the only law that makes us conscious of the independence of our power of choice from determination by all other incentives (of our freedom) and thereby also of the accountability of all our actions" (Rel. 26n). This passage indicates that Kant indeed holds that acting from respect for the Moral Law is *unique* in that it makes us aware of the property that our will has to be

---

<sup>22</sup> Korsgaard 168

<sup>23</sup> Korsgaard 169

<sup>24</sup> Korsgaard 169

determined independently of “all other incentives,” which I understand to mean sensory inclinations. This property of the will to act independently from desires associated with the phenomenal world is the ground for a human agent to represent herself as an efficacious member of the noumenal.

But does the pleasure that accompanies the moral feeling arise because the agent is acting solely from respect for the Moral Law? Or does it arise because the agent is acting in a way that allows her to represent herself as a member of the noumenal world, and it just so happens that the only ground for such a representation is acting solely from respect for the Moral Law? Korsgaard’s argument regarding the position of spontaneity certainly leads in this direction. Her reading suggests that the reason a purely rational will takes an interest in acting according to the Moral Law is not because it is the *Moral Law* but because it satisfies the condition of needing to find *some* principle (law) for determining itself without surrendering its *freedom*. Is the “moral feeling” perhaps better described as a “freedom feeling,” a unique sensibly independent feeling that is associated with morality only because acting solely from respect for the Moral Law is the only avenue by which human beings can represent themselves as free (and therefore members of the noumenal world)? Let us assume for the sake of argument that this reading is incorrect. Suppose that the root of the moral feeling is indeed morality and not freedom. This means simply that the human susceptibility to feel pleasure solely from respect for the moral law is grounded in the fact that respect for the moral law allows us to feel moral. On its surface, this seems to be circular. The human susceptibility to feel pleasure solely from respect for the moral law is grounded in the fact that respect for the moral law uniquely allows us to feel as if we are acting solely from the moral law.



Indeed, since the “moral feeling,” as a psychological incentive, is necessary for even the possibility of human morality, it would seem to be a serious problem to appeal to human morality as the basis of such an incentive.

Bringing all of this together, I understand Kant to be saying that when we act from respect for the Moral Law, we experience a moral feeling of pleasure that is the basis for our having morality as an incentive. This pleasure is not sensibly dependent; rather, it is a purely rational pleasure that can only be experienced by virtue of an agent’s representation of herself as a member of the intelligible world. Since this representation is not possible when an agent is in the grip of incentives of sensory inclination, and since the only way that an agent can conceive of herself as free from this grip is to conceive of herself as acting from respect for the Moral Law, an agent’s conception of herself acting solely from respect for the Moral Law is the only way to experience this special purely rational pleasure. To be precise, an agent who has adopted the moral principle does not determine her will in order to experience this pleasure as a consequent or reward. To act in order to receive an award would violate the maxim to act solely from respect for duty. The moral pleasure follows the representation of the law and acting solely from respect for it. In a sense, it is a reward, but it is the sort of reward that an agent receives only if she acts without consideration of a reward at all. It provides the human will with an incentive to be moral; that is, it makes being moral *interesting* to the human will. It does not, however, count as a reason for a determination to act morally.

### Section 3: Summary of Chapter 2

Insofar as human beings are imperfectly rational beings possessing a free will, we have the capacity to determine our will independently of external causes. This negative concept of freedom suggests a positive conception, which is the idea that our will must possess both a capacity to legislate for itself a law that is completely independent of external influence and an ability to determine itself in accordance with that law. This is the autonomy of a will consisting of *Wille* and *Willkür*. By virtue of her *Wille*, the law-making function of a will, a human agent legislates for herself the Moral Law, which is then presented to her in the form of a moral and categorical imperative. Through her *Willkür*, the executive function of the will, she considers both the categorical imperative presented to her through *Wille* and the hypothetical imperatives presented to her from exercise of empirical practical reason. She takes up these imperatives into maxims and freely chooses whether to condition the categorical moral imperative with the hypothetical (in which case she is immoral) or to condition the hypothetical imperatives with the moral (in which case she is moral).

As possessing a free power of choice, the human agent has two possible reasons for making her choice, self-love and respect for the Moral Law. A choice made solely from respect for the Moral Law is moral while a choice made solely for self-love is immoral. This answers several questions. For what reason must she choose if a determination of the will is to be moral? She must choose solely out of respect for the law. What other possible reasons are there? There is only one other possible reason, and that is self-love. It does not however tell us why it is that a human agent takes an interest in these two possible sources of imperatives. The answer to this question is that she is

predisposed as a matter of psychology to find certain ends compelling. These ends, arising from her rational animality, may be roughly described as self-love. A human agent takes an interest in these by virtue of the fact that she is a member of the sensible world and these ends are connected with sensible pleasure and happiness. We are, as a psychological matter, also “susceptible to respect for the moral law *as of itself sufficient to [Willkür]*” (Rel. 27). The human agent takes an interest in this “moral feeling” by virtue of the fact that she is a member of both the sensible and the intelligible worlds, and exercising her capacity to act solely from respect for the Moral Law is the unique way for her to represent herself fully as also a member of the intelligible world (and therefore free). When she acts morally, she does so in order to do the right thing (otherwise she is not acting morally), but she takes a psychological interest in morality in general not because it makes her feel moral but because it makes her feel (in a uniquely rational way) as if she is a free causal agent with efficacy in the intelligible world. Thus, the *reason* she acts morally is respect for the Moral Law, but the psychological *incentive* for acting morally is that it makes her feel free.<sup>25</sup>

---

<sup>25</sup> Another way to conceive of this distinction is that the psychological incentive offers an explanation for why humans act morally. This is an *empirical* claim that is contingent on the psychological characteristics of human beings. By contrast, the other claim is a necessary *a priori* judgment grounded in an analysis of what attributes an action must possess if it is to count as an instance of a “moral act.” Namely, for act to be considered moral, it must be in accordance with duty, and the agent must commit it solely from respect for duty.

CHAPTER 3: EVIL *QUA EVIL*

## Section 1: Diabolical and Malicious Beings

In *Religion within the Boundaries of Mere Reason*, Kant broaches the subject of what he calls a “*diabolical* being.” He offers this intriguing description:

A reason exonerated from the moral law, an *evil reason* as it were (an absolutely evil will), would [...] contain too much [to provide a ground of moral evil in a human agent], because resistance to the law would itself be elevated to incentive (for without any incentive the power of choice cannot be determined), and so the subject would be a *diabolical* being. (Rel. 35)

Subsequently, Kant refers to “malice,” which he defines as a “disposition (a subjective *principle* of maxims) to incorporate evil *qua evil* for incentive into one’s maxim” (Rel. 38). Kant’s definition of “malice” suggests the notion of what might be called a fundamental evil maxim. I contrast this with the fundamental immoral maxim of a human being, which states roughly: I will act in my self-interest, and I will act morally insofar as this does not interfere with my self-interest. The evil maxim that I am suggesting would be something like: I will act immorally, and I will act in my own self-interest so long as this does not interfere with my acting immorally. Another variation of this evil maxim would be simply: I will act immorally. The former would be the maxim of an imperfectly rational *malicious* being that is at the same time a member of both the intelligible and sensible worlds, and as such, it would have two sources from which imperatives would be presented to it. The latter would be the maxim of a *diabolical* being, an evil parallel to Kant’s Holy Will. Like its moral counterpart which has the Moral Law as its only incentive, such an entity would have no other incentive than the unconditioned principle of immorality.

Since it would have more in common with human beings, an imperfectly rational malicious being is of more interest here. Insofar as such a being is indeed a member of both the intelligible and the sensible worlds, determinations of its free will would have similar dynamics to that of our own. It would need to have a free and spontaneous power of choice, *Willkür*, and it would also need to possess something analogous to a legislative capacity, an evil *Wille*. Now recall that as the executive element of a free will, *Willkür* can only generate and select maxims from imperatives that are presented to it by virtue of its legislative evil *Wille*. The generation and adoption of the fundamental evil maxim by a malicious being would require a legislative capacity that was capable of establishing an unconditioned principle of immorality. This principle would have to be in some sense an opposing alternative to the Moral Law, and once established by an evil *Wille*, it would be presented to *Willkür* as an opposing alternative to the categorical imperative.

Section 2: The Conceivability of a Human Psychological Incentive to act  
from Evil *Qua Evil*

A key point is that if a malicious being is genuinely to determine its will solely from resistance to the Moral Law, then such a determination would have to be an autonomous act of a free will. In the same way that a moral act must be in accordance with the Moral Law and be determined by the will solely from respect for the Moral Law, an act from evil *qua evil* would require that such act be in accordance with the unconditioned principle of immorality (or contrary to the Moral Law) and determined by the will solely from respect for that principle (or resistance to the Moral Law). The involvement of any other considerations in the determination would amount to a choice

to act as if the will were heteronomous and would preclude the possibility of the will having indisputably chosen the fundamental evil maxim. The fact that an act done from the fundamental evil maxim would necessarily be autonomous entails that it would involve only a sort of *a priori* faculty that would be roughly analogous to a variation of pure practical reason, the aforementioned evil *Wille*. By virtue of this fact, a freely chosen adoption of the fundamental evil maxim would warrant a malicious agent to fully represent itself as an efficacious member of the intelligible world.

Two possibilities need to be mentioned here. The first is that the malicious being possesses an evil *Wille* that allows it to legislate only the unconditioned principle of immorality. In this case, a free adoption of the fundamental evil maxim would be the only avenue through which the malicious agent could warrant a full representation of itself as a member of the intelligible world. The other possibility is that the malicious being has a will consisting of a sort of super-*Wille* that is capable of legislating both the Moral Law and the unconditioned principle of immorality. In this case, the agent would have a choice of two possible foundations for full representation of itself as an agent in the noumenal world. Regardless of which of these two possibilities we consider, the important thing is that the adoption of the fundamental evil action offers one possible way for a malicious agent to fully represent itself a free agent in the noumenal world.

This suggests that if a human being were able to legislate the unconditioned principle of immorality, then she would have as powerful an incentive to act in accordance with it as we do to act in accordance with the Moral Law. As indicated above, the human incentive to be moral is not the reason for deciding to be moral; this reason can only be respect for the Moral Law. Rather, this incentive is the basis for

human beings' taking an interest in morality in general. As Korsgaard suggests, it is plausible that this incentive is rooted in the fact that acting solely from respect for the Moral Law is the unique way for a human agent to warrant a full representation of herself as a member of the intelligible world. To reiterate, we take an interest in acting solely from the Moral Law not because it makes us feel moral but because it makes us feel free. This strongly suggests that if there were another maxim, the adoption of which could make us feel free, our incentive to choose this alternative would be as compelling as our incentive to choose to act solely from respect for the Moral Law. *Ceteris paribus*, it is clear from this that were the fundamental evil maxim a possible choice for a human agent, there would be as compelling an incentive to adopt it as there is to adopt the fundamental moral maxim. Acting from evil *qua evil* would make us feel as free as acting from respect for the Moral Law. It would seem, then, that when Kant indicates that raising "resistance to the Moral Law" to an "incentive" is not "applicable to the human being," he cannot mean that human beings could not possibly take an interest in such resistance. There must be another reason that such an incentive is impossible.

There arises here a potentially serious problem for my reading. We may be tempted to assign the *conscience* a role in the explanation of psychological incentive. Since conscience is "practical reason holding man's duty before him for his acquittal or condemnation in every case that comes under a law," it is reasonable to think that by virtue of her conscience, a human agent would find it impossible not to have at least some incentive to act solely from duty (MM 401). Indeed, it seems plausible to suggest that an agent's conscience could provide her with an independent psychological incentive to act morally in addition to the "moral feeling." If this were the case, then a human

being (who had the fundamental evil maxim as a viable choice) would have an added incentive to act solely from respect for duty. With this bolstering incentive to act morally, it would not be the case that her incentive to act from evil *qua evil* would be as powerful as her incentive to act from the Moral Law.

This urge to assign the conscience a role in incentive should, however, be resisted.

In his denial of the possibility of an “*erring* conscience,” Kant states that

“For while I can indeed be mistaken at times in my objective judgment as to whether something is a duty or not, I cannot be mistaken in my subjective judgment as to whether I have submitted it to my practical reason (here in its role as judge) for such judgment” (MM 401).

In this somewhat abstruse passage, Kant seems to indicate that the proper role of conscience is to ensure that an agent is in a psychological position to *judge* whether or not a particular maxim is in accordance with her subjective belief about what her duty is (whether this belief is objectively correct or not). It plays a crucial psychological role in a human agent’s “recognizing duties” and making judgments as to whether a maxim is in accordance with those duties, but it plays a role neither in an agent’s deciding what sorts of things will count as reasons for determinations of her will, nor in how she will prioritize those reasons (MM 400). Put another way, an agent through her *Willkür* decides which reasons will suffice for her determinations and how she will prioritize those reasons. The psychological explanation of incentive offers an account of why it is that the agent takes an interest in duty as sufficient for determining her will. The agent’s conscience, however, offers an account of an agent’s recognizing that a particular situation warrants moral consideration and of her adjudication of whether a maxim that she has selected is in accordance with what she believes her duty to be.



To further develop this, let us suppose that the conscience does indeed play a crucial role in the account of incentives. In other words, assume that an agent by merit of her conscience cannot help but take an interest in moral duty. Conscience is the “holding of man’s duty before him,” but is it moral duty *per se* that conscience holds before the agent? Or, is it the product of the legislation of *Wille* that an agent’s conscience presents to her? The second option seems more plausible. Kant suggests that it is possible for an agent to be wrong about what her duty actually is. This implies that there is no necessary connection between the representation presented to the agent through her conscience and the Moral Law. In contrast to a human agent, consider the nature of the conscience of a malicious being, specifically the sort whose *Wille* can only legislate the unconditional immoral principle. Since such a being would have no concept of the Moral Law, it would be impossible for it to represent the Moral Law to itself either by virtue of its conscience or any other endowment. The malicious being through its conscience would be presented with a representation of the unconditioned immoral principle, which is the product of its *Wille*.

My reading suggests that it would be a mistake to assign the conscience a role in the psychological account of incentives, but even if we did, this would not change the fact that were a human agent capable of legislating an alternative to the Moral Law, she would have as much of an incentive to act from evil *qua evil* as she does to act solely from duty. If this were within her cognitive capabilities, then it is at least possible that her conscience would hold before her the unconditioned immoral principle. And were it the case that the conscience plays a role in incentive, whatever force it would have would be directed towards evil *qua evil*.

Section 3: The Impossibility of a Human Agent's Legislation  
of an Alternative to the Moral Law

In the context of the discussion regarding the evil nature of human beings, Kant indicates that the ground of human immorality cannot be placed in the fact or content of sensible inclination, that is, in the “sensuous nature” of human beings. The reason for this is that human beings cannot be held responsible for what appearances are presented to us by the phenomenal world. We are also not responsible for the fact that natural inclinations exist because, “as connatural to us, [they] do not have us for their author” (Rel. 35). Continuing, Kant states:

The ground of this evil cannot also be placed in a *corruption* of the morally legislative reason, as if reason could extirpate within itself the dignity of the law itself, for this is absolutely impossible. To think of oneself as a freely acting being, yet exempted from the one law commensurate to such a being (the moral law), would amount to the thought of a cause operating without any law at all (for the determination according to natural law is abolished on account of freedom): and this is a contradiction.

Here Kant explicitly denies the possibility of human beings acting immorally just for the sake of being immoral. I interpret this passage to mean that it is impossible for a human agent because of limitations of her capacity of *Wille*, to undermine, alter, or deny that she is bound by the Moral Law.

The reason for this is that the Moral Law that a human agent legislates is grounded in law as such or more precisely in the form of law independent of any empirical content.<sup>26</sup> *Wille* insofar as it is a constituent of an imperfect and finite rational

---

<sup>26</sup> Special care must be taken here. I do not mean to imply that Moral Law or the Categorical Imperative is devoid of *any* content; rather, I mean to say that it is devoid of empirical content. Kant characterizes the Categorical Imperative as an “a priori synthetic practical proposition” (Gr. 420). By virtue of this, it is clear that Kant holds that the Categorical Imperative is ampliative of our knowledge and therefore not entirely empty.

being has only one source for material or empirical content, and that is the sensible world. Recall that for Kant, will is a type of causality, and causality requires laws. The freedom of a will is a property of that causality that gives it a sort of efficacy that is independent of any causes that are not part of itself. For a free will to possess this property means that it must have a constitutive element (*Wille*) that is capable of generating a law according to which it will act. *Wille* must be able to generate a law without consideration of anything that is not a part of it. Since anything phenomenal is “alien” to it, the only possible law that *Wille* can legislate is a law devoid of any empirical content. This leaves the form of law as such as the only candidate, and to consider an undermining, alternation, or denial of the “dignity” of the law as such can, for the imperfectly rational human being, mean only lawlessness. The notion of lawlessness, however, contradicts the concept of causality, which is fundamental feature of Kant’s definition of a will. For this reason, it is impossible for *Wille*, and therefore also for the free will of a human being, to “corrupt” the morally legislative reason. Thus, the reason that a human being cannot act immorally just for the sake of acting immorally is that human *Wille* is incapable of legislating an alternative to the Moral Law.

This restriction on the legislative capacity of human is a consequence of a discursive intellect or, more precisely, the sort of discursive intellect that human beings possess. Without delving too deeply into pure reason in its theoretical employment, we can explain this claim by saying simply that phenomenal experience, which is dependent on the pure forms (space and time) of the human receptive sensibility, is the only source of the empirical content of our theoretical knowledge. We cannot extend our knowledge of the world beyond the possibility of experience, and such an extension is precisely what

would be necessary to establish an alternative to the Moral Law as legislated by virtue of *Wille*. Put another way, the legislation of an alternative to the Moral Law would require knowledge of the ordering principle of the noumenal world or noumenal causality.

On Kant's view, our establishment of the Moral Law is based on the speculative possibility that we are free and on the practical realization that we cannot make choices at all without the assumption that we are. Thus, it is only from the standpoint of reason in its practical employment that we are warranted in making claims with regard to human freedom. As a consequence of this, we are only able to establish the negative concept of freedom, and while it is true that an analysis of the idea of a free will entails a positive conception that grounds our establishment of the Moral Law, this conception is possible only because it is limited to the form of law as such. Christine Korsgaard, following John Rawls' discussions of justice, offers a helpful distinction between a *concept* of freedom and a *conception* of freedom. While a concept of something is "formally or functionally defined," a conception of something is "materially and substantively defined."<sup>27</sup> To use Rawls's terminology, in a real sense, our positive *conception* of freedom is, in fact, only a slightly more robust *concept* of freedom. The only material and substantive definition that we are warranted in establishing is really nothing more than a slightly more informative formal and functional definition. For the *Wille* to be capable of legislating an alternative to the Moral Law, it would need to be a constituent of an intellect that has additional or perhaps different resources by which it could obtain the material of its knowledge, an intellect that is, at least to a certain extent, intuiting. This follows from the fact that the only conception of the law that is available to human beings is essentially the form of law as such, and to think of an opposing alternative immediately entails a

---

<sup>27</sup> Korsgaard 162

contradiction with the very idea of a free will. For an intellect that had epistemic access to the noumenal world and was therefore able to develop theoretical knowledge of the ordering principles of that world (noumenal causality), the legislation of robust and content-laden alternative to the Moral Law would seem to be at least possible. For a being that possessed such an intellect, an alternative to the Moral Law would not immediately or necessarily entail a contradiction with the concept of a free will.

According to my reading, Kant's commitment to the impossibility of human beings acting immorally solely for the sake of immorality arises not from the psychology of human incentive. Since the human predisposition to personality, or "moral feeling," is rooted not in morality but in a human agent's fully representing themselves as a free agent in the noumenal world, it could act as a basis for an incentive to act in accordance with a fundamental evil maxim, were representation of the ground of such a maxim possible for human beings. Holding all other relevant factors equal, the inapplicability of evil *qua evil* to human beings is, therefore, not a psychological fact about human beings; rather, it is a fact about the sort of imperfectly rational beings that human beings are. By virtue of the particulars of the limitations on human cognitive capacities and avenues of epistemic access, it is impossible for us to represent an unconditioned immoral principle. This impossibility is the basis of the inapplicability of evil for its own sake to the human being. Human beings, then, are simply incapable of conceptualizing what it means to be immoral just for the sake of being immoral. We act immorally only because we are willing to put our own self-interests ahead of what is moral. If, however, it were within our capability to legislate the unconditioned principle of immorality, it is plausible that psychologically we would have an incentive to act from it.