
Kant, Skepticism, and Moral Sensibility

OWEN WARE

A thesis submitted for the degree of
Doctor of Philosophy
Graduate Department of Philosophy
University of Toronto

Supervisor
Prof. PAUL FRANKS

© Owen Ware 2010

Abstract

In his early writings, Kant says that the solution to the puzzle of how morality can serve as a motivating force in human life is nothing less than the “philosophers’ stone.” In this dissertation I show that for years Kant had searched for the philosophers’ stone in the concept of “respect” (*Achtung*), which he understood as the complex effect practical reason has on sensibility.

- I sketch the history of that search in CHAPTERS 1 and 2.
- In CHAPTER 3 I show that Kant’s analysis in *Groundwork* I is incomplete because it does not explain how respect functions as a feeling in motivating choice.
- In CHAPTER 4 I argue that Kant’s subsequent attempt to sidestep this explanation in *Groundwork* III is unsuccessful, and that his position remains open to a skeptical threat.
- The account in the second *Critique*, which I reconstruct in CHAPTERS 5 and 6, overcomes this threat, and in doing so explains how the feeling of respect is both painful and pleasurable.

In the course of defending these claims, I provide an alternative reading of the shift in Kant’s ethical project from the *Groundwork* of 1785 to the second *Critique* of 1788. Against a common view in the literature, I argue that the shift does not concern the direction of Kant’s deduction (from freedom to morality, or morality to freedom); rather, it concerns his view of human sensibility and the resources he thinks we have to make our practical self-understanding intelligible. In the second *Critique*, I argue, Kant develops a novel approach to moral feeling from the perspective of the human agent; and this in turn clears room in his ethics for a new kind of *a priori* knowledge—namely, knowledge of what the activity of practical reason must feel like.

By way of conclusion, I offer a few reasons for why the form of Kant’s argument in the second *Critique* is still relevant today, as it shows why we can only address moral skepticism from a *first-personal* perspective.

Acknowledgements

WHILE writing this thesis, I have received an enormous amount of help from many people, in many different ways. For this I am deeply thankful.

I was fortunate to meet PAUL FRANKS soon upon his arrival at the University of Toronto. Much of what I am presenting here—especially Kant’s dialectical relationship to skepticism—is a result of his mentoring. Paul has an unusual talent for knowing what I want to say even before I know how to say it. This thesis owes much to his encouragement.

SERGIO TENENBAUM and ARTHUR RIPSTEIN first introduced me to Kant’s ethics. The clarity and rigour they bring to the exposition of Kant’s thought has been, for me, something of a “regulative ideal.” Sergio is a tireless conversation partner, and I am truly grateful for his willingness to meet on short notice and discuss obscure passages from *Groundwork* III. Arthur helped shape my project at the beginning, and my ideas on Kant’s theory of moral sensibility—particularly as they differ from Heidegger’s—owe to his acute suggestions.

Over the past two years I have benefitted from a number of individuals. STEVE ENGSTROM was kind enough to give me detailed feedback on my early work on Kant. BARBARA HERMAN also gave me much encouragement, and my work on the reflexive character of “feeling” in Kant is largely the result of a single conversation with her. More recently, I have had the pleasure to share ideas with BOB STERN, and his recent work on Kant’s relationship with moral skepticism has had a major impact on my thinking.

The [philosophical community](#) at the University of Toronto has been incredibly supportive. In particular, I would like to thank DONALD AINSLIE, PHIL CLARK, REBECCA COMAY, ROBERT GIBBS, GUPREET RATTAN, ULRICH SCHLÖSSER, and JENNIFER WHITING. Each offered me a different philosophical perspective and helped me navigate my way through the Continental/Analytic divide. During the 2009-2010 year I was a graduate fellow at the [Jackman Humanities Institute](#), which offered me a great place to work, drink espresso, and share ideas with a wonderful group of people. The [Social Sciences and Humanities Research Council](#) also supported me with a Canada Graduate Scholarship from 2006 to 2009, for which I am grateful.

I thank my parents, ANNE and GRAHAME, for their love and on-going support. As always, my greatest debt is to LEAH HARMS. Her love has sustained me throughout the years. This thesis is dedicated to her.

for Leah

List of Figures

6.1	Kant's Dialectic of Respect	110
-----	---------------------------------------	-----

Contents

1	Introduction	1
1.1	Preliminaries	1
1.2	The Quiet Avoidance of Justification	3
1.3	Rebellion Against Psychology	6
1.4	The Philosophers' Stone	8
1.5	Phantoms of the Imagination	11
1.6	A Roadmap	13
I	The Problem	16
2	What Moral Question?	17
2.1	Preliminaries	17
2.2	Empirical Skepticism	18
2.3	From Metaphysics to Critique	21
2.4	Transcendental Skepticism	23
2.5	A Good (Human) Will	25
2.6	Deduction or Defence?	28
2.7	The "Fact of Reason"	32
2.8	Looking Ahead	35
3	Freedom and Rational Agency	37
3.1	Preliminaries	37
3.2	Acting Under the Idea of Freedom	39
3.3	The Standpoint of Theoretical Reason	43
3.4	Happiness and Rational Heteronomy	44
3.5	The "Hidden Circle"	47
3.6	"Raw" Transcendental Idealism	50
3.7	The <i>Verstandeswelt</i> as an Archetype	54

4	The Moral “Ought”	58
4.1	Preliminaries	58
4.2	The Intelligible Ground of the World of Sense	59
4.3	Intuition and Practical Sensibility	64
4.4	The Case of the Mafioso	67
4.5	Actually or Possibly Motivating?	70
4.6	Looking Ahead	74
II	The Solution	76
5	Moral Feeling	77
5.1	The “Aesthetic” of Pure Practical Reason	79
5.1.1	Ectypal Nature	82
5.1.2	The Typic	83
5.1.3	The Causa Noumenon	86
5.2	Life, Desire, Feeling	89
5.3	Heidegger’s Mistake	91
5.4	Hegel’s Challenge	93
5.5	The “Harmony” of Duty and Inclination	96
6	Humiliation, Elevation	100
6.1	Preliminaries	100
6.2	The Dialectic of Respect	104
6.3	Completing the Experiment	111
6.4	Two Objections	112
6.5	Moral Enthusiasm	113
6.6	Closing Remarks	118
	Abbreviations	121
	Bibliography	123
	Index	131

CHAPTER 1

Introduction

Nobody can or ever will comprehend how the understanding should have a motivating power; it can admittedly judge, but to give this judgment power so that it becomes a motive able to impel the will to performance of an action—to understand this is the philosophers' stone.

— Kant (LE 27:1428).

... it had better have a footing in the heart's desire.

— Bernard Williams (1996).

1.1 Preliminaries

IF we have learned anything from Kant's ethical theory, it's that we cannot ground moral principles on a settled view of human motivational capacities. Ethics must precede psychology. But that is not to say ethics must do away with psychology. Once we understand the demands that moral reasons place on us, we can better understand the psychological conditions of moral choice and action. Ethics can *teach* us psychology. It can teach us what we must be capable of as human beings.

In a trivial sense Kant's ethics is rationalist. It grounds moral principles on reason. How this position differs from ethical rationalists in the tradition—like Samuel Clarke or Christian Wolff—is a question we need not raise here. But one problem Kant shares with his rationalist predecessors is a problem of motivation. If moral principles have their seat in reason, how can those principles move us? Reason does not seem equipped with motive force, and yet this is what Kant and other rationalists are committed to saying. Reason must be capable of directing our choices. It must be practical

in the sense of action-guiding. But how reason can be practical, how it can guide *action*, is precisely the mystery. And many have found this mystery so baffling that, like Hume, they have demoted reason's role in deliberation to that of a slave, a "slave of the passions."

Kant may not have been aware of Hume's skepticism about practical reason, but he was aware of the motivational problem in its general form. He also proposed different ways of responding to it. The most well-known response comes from the *Groundwork of the Metaphysics of Morals* (1785), in the final section.¹ Kant recognized that there is something deeply mysterious about the claim that reason can be practical. The source of the mystery is that if reason is to move sensible beings like us, it must be able to instill a feeling of pleasure in us, pleasure in conforming our actions to reason. This is the only way moral action would be possible for human beings, since Kant acknowledged that we are only capable of forming choices on the basis of feelings. This is part of the structure of *our* psychology. If reason is to be practical, then, it must find its way into this structure. It must influence our capacity to feel.

The problem, of course, is that it seems we cannot explain this. We cannot explain how reason gains access to our motivational capacities in a way that could produce an interest in us, an interest to be moral. There is a gap here that no amount of philosophical reflection can bridge: the gap between our knowledge of morality and our putative interest in its claims. This gap arises for us because our understanding is limited. For Kant, our understanding of causal relations is limited to objects of sensory experience; but moral principles are ideas of reason and, because of that, can never be objects of sensory or intellectual intuition. So how these principles move us, i.e., how they enter into the structure of human motivation, is a causal mystery. Unexpectedly, Kant's solution in the final section of the *Groundwork* is to affirm this mystery as it stands. We cannot explain how reason influences our capacity to feel, such that our sensibility can conform to its principles. But that does not support skepticism. Precisely because we are ignorant of reason's influence, we are able to hold open the possibility that moral principles can elicit feelings from us. This justifies our position. For the claim that reason is practical does not depend on our ability to explain the causal relation between moral principles and our capacity to feel. Ethics can teach us psychology, then, without the help of metaphysics.

This is one way of summarizing the conclusion of the *Groundwork*. But the story does not end there. Just three years later, in the *Critique of Practical Reason* (1788), we find Kant developing an extended line of argument. In Chapter III, he devotes an entire analysis to what he calls, somewhat ob-

¹See p. 121 for more information about the list of abbreviations and translations of Kant's works used in this thesis.

scurely, the “incentives of pure practical reason.” Keeping the metaphysics at bay, Kant wants only to show how the moral law can function as the highest incentive of the will by *depriving* empirical incentives of their influence on us. In particular, he wants to show how the moral incentive can overcome our propensity to prefer self-love over the law, what he calls “self-conceit.” But why does Kant want to offer us this further account? His conclusion in the *Groundwork* involves affirming our ignorance of moral motivation. There is nothing more we can say about the metaphysics of our interest in morality. In the second *Critique*, though, Kant is not concerned with the metaphysics, he is concerned with the *psychology*. But why? And why does Kant think this concern deserves a separate chapter?

In attempting to answer this question, I have come to see that the stakes of Kant’s moral psychology are much higher than we think. I have also arrived at a different view of Kant’s development from 1785 to 1788, a development that concerns, not his theory of practical reason, but his theory of practical sensibility. I believe that Kant came to recognize that his argument in the final section of the *Groundwork* failed because it did not spell out the psychological conclusions that follow from his account of moral obligation. His argument left enough room for skepticism about moral motivation, skepticism he later addressed in the second *Critique*. My short answer to the above question, then, is that Kant wrote a separate chapter on moral psychology in 1788 because he was responding to the worry that we have no way of understanding how reason gains access to sensibility. He was responding to the question, left unanswered in the *Groundwork*, of how morality can function as an incentive for us.

1.2 The Quiet Avoidance of Justification

Most criticisms of Kant get their inspiration, directly or indirectly, from Hegel. In the *Phenomenology of Spirit* (1807), and again in the *Philosophy of Right* (1820), Hegel denounced Kantian morality as an “empty formalism.” This charge has two sides. One side is that the categorical imperative is empty of *content*; the other is that it is empty of *force*. Hegel argued that we can never derive specific duties from the moral law, and that the moral law can only posture as an incentive for us. In any concrete situation, the moral law cannot supply us with guidance for how to act. And even if it did, we would remain unmoved. An abstract law is incapable of entering into the driving mechanism of a human will.²

In the twentieth century, critics and defenders of Kant have focused on the first side of Hegel’s “empty formalism” charge. This is especially true of the Analytic Kantianism that developed among British and American

²I will discuss this side of Hegel’s charge in section 5.4.

philosophers; and a concern with “content” has shaped the interests of Kant scholarship to the present day. If we glance at any major book written on Kant’s ethics during the 1960s and 1970s, we will find extensive discussion of Kant’s formulations of the categorical imperative, the topic of *Groundwork* II, but little—if any—discussion of Kant’s deduction of the categorical imperative, the topic of *Groundwork* III, or his subsequent doctrine of the “fact of reason.”³

In her characteristically lucid manner, Barbara Herman diagnosed the situation as follows:

In the resurgence of work on Kant’s ethics, one notices the quiet avoidance of the issue of justification. This is to some extent the harmless by-product of new enthusiasm generated by success with the substantive ethical theory. But the other thing at work, I believe, is the suspicion that the project of justification in Kantian ethics is intractable. What these two papers do, decisively, is to dispel this kind of anxiety. (p. 131)

The two papers she refers to are John Rawls’ “Themes in Kant’s Moral Philosophy” and Henry Allison’s “Justification and Freedom in the *Critique of Practical Reason*,” conference papers that were later published in *Kant’s Transcendental Deductions*. That book appeared in 1989. Since then, Herman’s words have lasted the test of time. Allison went on to write a number of important studies on Kant’s project of moral justification; as have others. The literature is still small, but it reflects a reshaping of interests in Kant’s seemingly intractable arguments for morality and freedom.

But there is something unexpected here. When we look closely at the new literature on justification—including the work of Thomas Hill Jr., Christine Korsgaard, and Allen Wood—we do not find the second half of Hegel’s “empty formalism” charge. This is not by mistake, as it turns out, for these writers agree that the question of whether moral concepts can *move* us is separable from the question of whether the demands those concepts place on us are *justified*. The problem of moral justification for Kant is distinct from the problem of moral motivation; since the latter is the proper topic of his moral psychology.

What has led to this assumption? One cause, I believe, is the recent attempt to naturalize Kant’s ethics, an attempt that is continuous with the view—put forward most forcefully by Sir Peter Strawson—that we can separate the conceptual face of Kant’s system from its transcendental face, and

³See, for instance, Robert Paul Wolff’s *The Autonomy of Reason* (1973), Onora O’Neill’s *Acting on Principle* (1975), and Bruce Aune’s *Kant’s Theory of Morals* (1979). Notable exceptions include the major commentaries on Kant’s ethics by H.J.Paton, *The Categorical Imperative* (1947), and by Lewis White Beck, *A Commentary on Kant’s “Critique of Practical Reason”* (1960).

do away with the latter. In the last twenty years, there has been an increased interest in Kant's moral psychology because many believe it is compatible with a more naturalistic understanding of human action and agency. Some of the best and most exciting work on Kant's ethics today is carried out in this spirit. In the Introduction to his book of essays *Self to Self* (2006), for example, David Velleman maintains that his commitments to Kant and naturalism do not conflict. The irony, he thinks, is only apparent:

Naturalism in moral psychology has traditionally been associated with Hume. But can we be naturalists without settling for Hume's impoverished conception of human nature. I believe—though I don't pretend to have shown—that we can be naturalists while preserving the moral and psychological richness of Kant. (p. 14-15)

This means, roughly, that we can speak in Kant's vocabulary of moral emotion and motivation—with its emphasis on “respect” —without having to commit ourselves to anything like a noumenal agent making decisions in a timeless realm; in short, without having to commit ourselves to Kant's theory of transcendental freedom. Practical reasoning, as Velleman and others conceive it, is compatible with an empirical account of agency and action, and requires nothing like a rational self independent from the influences of sensibility.

A version of this view is also shared by some of the most prominent interpreters of Kant. Allison and Wood, for example, have presented compelling arguments for why Kant's theory of transcendental freedom is not metaphysical, at least not in any sense that would leave his moral philosophy problematic. I won't try to summarize those arguments or weigh their merits against competing interpretations. What is relevant here is the implication they share for our understanding of Kant's moral psychology. Allison is aware of this implication, and he considers it good news. “Fortunately,” he writes,

it is also possible to “bracket,” as it were, the underlying presupposition of the validity of the moral law and to regard the discussion of respect essentially as a phenomenology of moral experience. Considered as such, it can be taken as complementing the account of moral worth offered in the *Groundwork* without denying the possibility that, for all we know, morality might be nothing more than a “phantom of the brain” (*Hirngespinnst*), which is the methodological assumption of the first two parts of the latter work. (p. 121)

This is to say we can specify the psychological conditions of moral motivation, so that we can explain how acting out of respect for the law is possible, without having to justify the assumption that we are transcendently free. In turn, we needn't make progress toward answering the difficult question

of whether moral concepts place legitimate demands on us. Allison's point is that morality may be an illusion—we haven't settled that yet. But we can still make progress toward understanding the motivational capabilities that *follow* from the “fact” of morality's binding character.

Allison's point, then, is that Kant's moral psychology builds on the argument that actions possess moral worth if, and only if, they are performed from duty. This is the well-known conclusion from the first section of the *Groundwork*. Kant's steps are rather simple. An action doesn't contain intrinsic value if it is the product of, say, one's fear of punishment. We then work our way down the list of possible candidates (ranging from blatant hedonistic considerations of pleasure, to more subtle consideration of prudence), rejecting one after the other, until we arrive at “respect” (*Achtung*). Acting from respect for the law is the sole disposition that confers value to one's actions, for it is the only disposition that meets the necessary and sufficient conditions of “moral worth” that Kant has analyzed. Allison's suggestion is that Kant's account of respect put forward in the second *Critique* is a psychological compliment to this argument. It fills out some of the emotional detail missing from his earlier discussion of moral worth. But that is all. Kant's phenomenology of respect does not carry any argumentative weight: it does nothing to alleviate the worry that morality may be, at bottom, a “phantom of the brain.” This worry is the proper topic of moral justification, and that's another chapter.⁴

We can now begin to see why there is whole literature on Kant's moral psychology that makes no mention of the deduction of the categorical imperative; and a whole literature on the deduction of the categorical imperative that makes no mention of his moral psychology. If we have good textual reasons to separate them, what would stop us (even the naturalists among us) from taking the moral psychology and running with it? Even if we end up a far cry from the transcendental side of Kant's system, including his concept of freedom, we can get along well. For we can still preserve, as Velleman says, “the moral and psychological richness of Kant.” One of my broader aims in this thesis is to show why this separation in the literature cannot be won so easily.

1.3 Rebellion Against Psychology

In the above sketch I have suggested that one cause for the separation between the question of justification and the question of motivation in the literature is an outcome of recent attempts to naturalize Kant's ethics. A more specific cause, I think, is due to the belief that Kant's ethics contains

⁴I will use moral phenomenology and moral psychology more or less interchangeably in this thesis; for my definition of Kant's “phenomenology,” see section 6.1.

resources for overcoming forms of moral skepticism, especially those based in doubts about the extent to which reason can, independently of desire, serve as a motive for action.

Thomas Nagel and Christine Korsgaard both understand Kant along these lines. In the *Possibility of Altruism* (1970), for example, Nagel distinguishes between ethical theories that ground their principles on a pre-defined conception of human motivational capacities, and theories that specify the psychological factors of moral action *after, but in light of*, specifying the requirements that moral reasons place on us. Hobbes and Hume are the best examples of the first approach; Plato and Aristotle of the second—but Nagel points to Kant as the best representative of a “rebellion” against psychology in ethics. “On Hume’s view one begins with psychology, and ethics is an elaboration of it” (p. 11). Kant’s view is just the reverse: his efforts are spent placing moral principles at the foundation of human conduct and determining, from there, what psychological conclusions follow. This is just what we have seen in the sketch I gave of Kant’s argument for moral worth. We ask ourselves what actions conform to the concept of a good will, and from there we eliminate any motives that arise from ordinary human psychology (e.g., an instinct for self-preservation, a fear of punishment, a desire for reward, approval in the eyes of others), until we arrive at a motivational capacity consistent with the idea of duty. That is, until we arrive at “respect” (*Achtung*). This is how we learn something about ourselves, about what we’re capable of, in light of the demands that morality places on us. So Kant’s ethics, “rather than appropriating an antecedently comprehensible motivational foundation on which to build its requirements, actually uncovers a motivational structure which is specifically ethical and which is explained by precisely what explains those requirements” (p. 12). The priority of psychology leaves someone like Hume blind to this structure.

Building on these insights in her 1986 paper “Skepticism about Practical Reason,” Korsgaard shows how the priority of ethics in Kant closes off a particularly troubling form of moral skepticism: skepticism about motivation.⁵ She argues that doubt about the extent to which reason can serve as a motive only arises if we misunderstand the relationship between reasons and motives. The fact that motives are part of our explanation of reasons does not suggest that motives must figure into the conditions necessary for having reasons. Korsgaard’s point is that when we understand this relationship properly, we will see that skepticism about practical reason always *reduces* to skepticism about the scope, or function, or capacity of practical reason. In other words, motivational skepticism carries no independent force from what she calls “content skepticism.” What, then, can we do when we are confronted with someone who points out that moral requirements specify

⁵I will cite from Korsgaard’s (1996b).

reasons for action *but still fail to motivate us*? Korsgaard's answer—echoing the conclusion of the *Groundwork*—is that we can only point to the limits of our understanding. “Philosophy can at most tell us what it would be like to be rational” (p. 332), but it cannot explain how we can be rational. Our only response to skepticism about moral motivation is the affirmation of human finitude.

On the view that I defend in this thesis, Korsgaard's response to moral skepticism, like Kant's response from the *Groundwork*, is not enough. When we come to understand why Kant wrote a separate chapter on moral psychology in the second *Critique*, we will see that contemporary strategies of reducing motivational skepticism to content skepticism are limited. This aspect of my argument is not exactly destructive; it does not show that contemporary responses to moral skepticism are wrong. If anything, it shows that they do not go far enough: they leave room for doubt because they fail to spell out the psychological conclusions of their own arguments. We will see that much of Kant's position from the *Groundwork* is right, and that his later account of moral feeling is consistent with his rejection of metaphysics. But we will also see that Kant's way of thinking about human sensibility changes in the second *Critique*, as does his conception of what philosophy can do in making our experience of morality intelligible. In addition to telling us what it would be like to be rational, Kant came to see that philosophy can tell us something more. It can tell us what it must *feel* like to be rational.

1.4 The Philosophers' Stone

Before going any further, I think it would be wise to look back. Kant's ethical thought underwent a long development, although it's not clear how, or when, Kant arrived at his mature views. I want to provide a sketch of this development between the period of 1764 and 1784—the twenty years leading up to, and including, the *Groundwork*. After this I will provide an outline of my argument as a whole.

In his so-called “semi-critical” period, Kant formulated the principle of moral judgments, the *principium diiudicationis*, which would later become the categorical imperative; but he spent decades struggling to find the *principium executionis*, the principle of how such judgments become effective. Above all else, Kant desired to unify the principle of appraisal (which he spoke of as the “objective side” of ethics) and the principle of execution (as the “subjective side”) under a single principle. He once referred to this unity as “the philosopher's stone.” But what led him to this view? And how much did he retain of it in the *Groundwork*?

It may be an exaggeration to say Kant went through a sentimentalist

phase, but in the “Prize Essay” of 1764 he did say our judgments of the good derive from more basic, unanalyzable feelings; and he praised Francis Hutcheson for “providing us with a starting point from which to develop some excellent observations” under the name “moral feeling” (DG 2:300). Yet the essay ends on a tentative note. Kant says we do not yet know “whether it is merely the faculty of cognition, or whether it is feeling (the first inner ground of the faculty of desire) which decides its first principles” (DG 2:300). Kant seems to lean toward a sentimentalist view in the *Observations on the Feeling of the Beautiful and Sublime*, also published in 1764, where he writes that virtue arises from a “consciousness of a feeling that lives in every human breast,” namely, “the feeling of the beauty and the dignity of human nature” (GSE 2:217). But right before saying this he asserts a rationalist claim, that virtue must be based on “principles” (GSE 2:217). This undecidedness runs throughout other writings of the period.⁶ The only place where Kant appears to defend a strong sentimentalist view is in a reflection written sometime around 1764-68: “The rules of morality proceed from a special, eponymous feeling, upon which the understanding is guided” (R6581; 19:77). But even this is unclear. Does he mean moral feeling *grounds* those rules, or that it simply *guides* their execution?

It is difficult to say how we should read these early writings. There are significant points of agreement between Kant’s early and mature ethical views, although I think it is important to highlight the uncertainty of his thinking prior to 1785. His praise of Hutcheson in the “Prize Essay,” to take one example, does not clarify his agreement with the sentimentalists. Nor do the other remarks I have cited. My own speculation is that during the 1760s Kant saw himself building on the ideas of the British, rather than defending their views. He indicates as much in his “Announcement”

⁶For example, in his “Announcement” to the lecture program he designed from the Winter Semester of 1765-1766, Kant writes: “The distinction between good and evil in actions, and the judgment of moral rightness, can be known, easily and accurately, by the human heart through what is called sentiment, and that without the elaborate necessity of proofs” (N 2:311). But this does not commit him to the view that sentiment grounds moral principles, only that the conclusions of such principles can be known “by the heart.” This may not be that different from Kant’s mature view that moral principles are covertly contained in the most common human understanding. Interestingly, Kant comes much closer to articulating his mature view of moral feeling in *Dreams of a Spirit Seer* (1766). He writes: “We sense within ourselves a constraining of our will to harmonize with the general will. To call this sensed constraining ‘moral feeling’ is to speak of it merely as a manifestation [*Ercheinung*] of that which takes place within us, without establishing its cause.” And a little later: “If the phenomenon of the moral impulses were represented in this way, the moral feeling would be this sensed dependency of the private will on the general will: it would be an effect produced by a natural and universal reciprocal interaction” (2:335). Taken out of its theosophical context, Kant’s definition of moral feeling as a consciousness of constraint or “sensed dependency” comes very close to his later definition of moral feeling as the consciousness of the law’s “effect” on our sensibility.

to the lecture series of 1765-1766, where he says that Shaftesbury, Hutcheson, and Hume have “penetrated furthest in the search for the fundamental principles of all morality,” though their search is “incomplete and defective.” There he promises to supply in his lectures what the sentimentalists lack: “precision and completeness” (N 2:311).

However we read these remarks, it is evident that Kant underwent some kind of change between the publication of the “Prize Essay” and the writing of the *Inaugural Dissertation* (1770).⁷ In the latter work, we do not hear any praise directed towards the sentimentalists. Kant does not even mention Hutcheson (or Shaftesbury or Hume, for that matter), and his brief remark on the foundations of moral principles indicates a clear break from their thinking: “moral concepts,” he writes at §7, “are cognized not by experiencing them but by the pure understanding itself” (2:395). Other reflections from the 1760s foreshadow Kant’s hybrid view, that moral principles are grounded in reason, and that moral feeling is an effect of those principles—not their foundation.

In a letter to Marcus Herz, dated toward the end of 1773, Kant makes his new position more clear: “The highest ground of morality must not simply be inferred from the pleasant; it must itself be pleasing in the highest degree. For it is no mere speculative idea; it must have the power to move. Therefore, though the highest ground of morality is intellectual, it must nevertheless have a direct relation to the primary springs of the will” (Br 10:145). Kant would spend the next eleven years struggling to formulate this “direct relation.” In the *Lectures on Ethics*, for example, we find moral feeling presented as a distinct emotive response to our moral evaluations, i.e., feelings of approval that presuppose our understanding of the good. In *Moral Philosophy* Collins, he says: “The moral feeling is a capacity for being affected by a moral judgment. When I judge by understanding that the action is morally good, I am still very far from doing this action of which I have so judged. But if this judgment moves me to do the action, that is the moral feeling” (LE 27:1428). Passages like this suggest Kant retained a sentimentalist view of moral feeling (i.e., as an emotive response) but made it dependent on the faculty of understanding. This was the general outline of the hybrid view Kant would retain for the rest of his life.

Yet the details of the view were the subject of constant change. Prior to

⁷What brought about this change of view? In his essay “Kant and Greek Ethics,” Klaus Reich (1939) gives us a precise date: When Kant read the *Phaedo*, the dialogue in which Socrates famously denounces moral systems grounded in sense rather than reason, which he believes Kant must have read before embarking on the *Dissertation*. Reich speculates the *Phaedo* must have impressed Kant, not only because it criticized a sentimentalist leaning he may have felt at the time, but because Mendelsohn, in his introduction to the 1767 German translation, made a persuasive case that the *Phaedo* comes closest to representing Plato’s own thinking. This is an interesting suggestion, but for lack of space I cannot pursue it here.

1785, the only claim Kant made with any consistency is that moral feeling *follows* from the understanding. He speaks of moral feeling as an “incentive” and a “motivating principle,” but to my knowledge he never specifies its status.⁸ For the most part, Kant was keen on separating the principle of moral appraisal from the principle of execution:

We first have to take up two points here: (1) The principle of appraisal of obligation, and (2) the principle of its performance or execution. Guideline and motive have here to be distinguished. The guideline is the principle of appraisal, and the motive that of carrying-out the obligation; in that they have been confused, everything in morality has been erroneous. (LE 27:274)

Kant’s disparaging remarks about moral feeling must be understood in light of this distinction. He believed the sentimentalists were mistaken to think that we could derive rules from sentiments of approval and disapproval. Feelings are subjective; they only have private validity. Moral rules must be a product of the understanding, because only the understanding can generate laws of action. “In all moral judgments we frame the thought: What becomes of the action if taken universally? If, when it is made into a universal rule, the intention is in agreement with itself, the action is morally possible; but if it is not, then it is morally impossible” (LE 27:1428). This is a judgment of the understanding, which may even oppose what I feel or what I find privately agreeable. Moral feeling is “empty and null” as a principle of appraisal, and yet the sentimentalists were on the right track. For morality is the subordination of an action under the universal rule produced by the understanding; but Kant thought our responsiveness to this rule is the capacity of moral feeling. Moral feeling is thus the principle of execution: it is how we translate knowledge into action.

1.5 Phantoms of the Imagination

In following the trajectory of Kant’s ethical thought, we get the impression that what requires justification is the above presupposition: that the understanding possesses motive force, or that it connects to our sensibility by way

⁸As far as I can tell, however, the *Lectures on Ethics* never settle the question of whether moral feeling is (i) a distinct emotional response or empirical feeling produced by our consciousness of the moral law, or (ii) whether it is the consciousness of the will’s subjective determination by the moral law. This is consistent with what he later says in the *Critique of the Power of Judgment* (1790), in the first Introduction, that the objective determination of the will contains “a feeling of pleasure.” In his personal copy of the third *Critique*, Kant crossed out “as in fact found to be identical with the former.” By doing so I think he wanted to leave the ontological status of moral feeling undecided. Following the above sentence he said that will-determination either precedes and conditions a distinct feeling of moral pleasure, or is “perhaps nothing than the sensation of the determinability of the will through reason itself, thus not a special feeling” (KU 20: 207). I will return to this point in Part II.

of moral feeling. It is a little surprising, then, that Kant addresses a different question in the final section of the *Groundwork*—a question of moral obligation. In 1785 what Kant thinks is puzzling and in need of an answer is not the question of how we can be *moved* by our rational evaluations of the good, but how those evaluations can appear to us as *imperatives*. Only by answering the question, “How is a categorical imperative possible?” can we overcome the thought that our consciousness of duty may be a “phantom of the imagination” (G 4:447). What happened to *the philosophers’ stone*?

The answer remains elusive. The concept of moral feeling does appear in *Groundwork* I, but there Kant’s concern is different. He wants to clarify the kind of motivational state sufficient to give an action moral worth. This is the eliminative argument we discussed above. Unfortunately, Kant says little about what respect is, either as a feeling or as a feature of moral consciousness, and what he does say reveals what I think is a blind spot in his argument. For example, he adds in a footnote that “it could be objected that I only seek refuge, behind the word *respect*, in an obscure feeling, instead of distinctly resolving the question by means of a concept of reason” (G 4:401n). In response, he says that respect is “immediate determination of the will by means of the law and consciousness of this,” and “the representation of a worth that infringes upon my self-love.” What’s troubling, though, is that Kant does not go on to explain the relationship between the two: he does not explain how I am conscious of the immediate determination of my will through respect, or how respect expresses an effect of practical reason on my sensibility.

The problem, as I see it, is that Kant’s position is caught between two intractable, yet conflicting, commitments. This comes out in the argument from *Groundwork* III. Very roughly, in order to overcome the worry that our consciousness of duty may be illusory, Kant must show how we recognize ourselves “under obligation” as rational beings. But he is also forced to limit the access we have to ourselves, as sensibly affected beings, to the mode of intuition. The problem is that in the mode of intuition we are *passive* to external influences, and so *incapable* of recognizing an obligation. This is what Kant will decisively reject by the time of the second *Critique*. In particular, he will argue that we can think about ourselves, as beings who experience the moral “ought,” in the mode of *feeling*. What is distinctive about our capacity to feel is that we *feel ourselves* in what promotes or detracts from our faculty of desire, and Kant will argue that we can bracket the question of whether feeling must be pathologically driven or whether it can also arise from reason. In my view, this was his major breakthrough. With the concept of feeling a new methodological possibility opened up for Kant that was unavailable in *Groundwork* III. Now he could describe the access reason has to our sensibility in terms of the feelings it must instil in us.

In saying this, I am departing from the familiar view outlined above, that the second *Critique* is merely adding psychological richness to the doctrine of moral worth from the *Groundwork*. The problem with this view is that it rests on the assumption that Kant's theory of moral sensibility is external to his project of moral justification. This is Allison's view. Interestingly, Allison says he is following in the direction of Lewis White Beck (1960), but it is not clear this is Beck's position. Beck says that Kant's theory of moral sensibility is not meant to explain the *metaphysical ground* of our interest in morality, but rather what *effects* this interest must have on us as sensible beings. On this point, Beck suggests the second *Critique* adds a small but significant step over *Groundwork* III, as he writes:

Though it displaces the inexplicable mystery of man by only one step, the *Critique of Practical Reason* does attempt at an explanation, in psychological terms, of how the knowledge of the moral law can be effective in the determination of conduct...It is essential that this mystery be removed from the phenomenological surface, as it were, for the thing is so puzzling that doubts of its reality can have the actual effect of reducing the effectiveness of this incentive. (pp. 210-211)

I do not see what support for his reading Allison gets from this. Kant's account in the second *Critique* presupposes the validity of the moral law (which we recognize in the "fact of reason") but that does not address the further question of how morality itself can function as an incentive for us. For this reason, we can read Beck's remarks about the threat of motivational skepticism to suggest that Ch.III is a continuation, and, perhaps, a *completion* of Kant's deduction. That is, at least, the view I will defend.

1.6 A Roadmap

Most discussions of feeling in Kant fall into one of two classes. The first are critical of a widespread misreading of Kant's ethics as coldly affirming duty at the sacrifice of inclination. Writers in this group are eager to point out that Kant's notorious examples of duty versus desire are not intended to portray a genuine ethical disposition—which requires feeling and a positive relation to human sentiment. The examples are designed for a specific purpose: that of highlighting the role of reason in moral action. Showing the conflict between duty and desire, even at the sacrifice of the most powerful desire of all—our love of life—is a heuristic. It is a way of bringing out the sovereign status of morality over the pursuit of self-love. It is not meant to describe how we should act, as if subduing desire is a requirement of acting out of respect for morality. Kant, we are told, is not an enemy of feeling.

The second class agree. But their concern is not with the question of whether Kant was opposed to feeling and related concepts (desire, emotion, affect), but with the question of how Kant viewed the connection between feeling, morality, and the will. Two subgroups emerge here—the affectivists and intellectualists. The first argue that Kant understood the feeling of respect as a distinct empirical feeling (somehow) produced by our intellectual recognition of the moral law. The second argue that Kant understood the feeling of respect just as this intellectual recognition, which (somehow) manifests a phenomenal character, as pleasure or pain. Thus the debate concerns the status of moral feeling and its location in the genesis of action motivated by duty.

I would like to say that both groups have part of the truth. Yet I do not think they fully explain the role of moral feeling in Kant's ethics. The assumption shared by affectivists and intellectualists is that the question Kant is seeking to answer is a question about the metaphysics of moral action. While Kant speaks in terms that come from a tradition invested in this question, the problem Kant is onto pushes us into new territory. On my reading, Kant's theory of moral feeling is best understood in view of his project of moral justification. Instead of contributing to the intellectualist/affectivist debate, then, I want to propose a new way of thinking about the motivational problem, one that I believe does justice to Kant's attempt to explain the experience of practical reason peculiar to beings like us. This will be my task in Part II.

In making this claim, I run up against an even bigger interpretive question: What is Kant's aim in the final section of the *Groundwork*, and how does it relate to the second *Critique*? My answer to the former will take up most of Part I. By the end of it I suspect the reader will have forgotten why I began this introduction talking about feeling and sensibility. This is not by accident. In *Groundwork* III Kant limits the question of how principles of rational choice connect to our sensibility to a question of how it is possible for morality to elicit feelings of pleasure in us. This is a speculative question—Kant is quick to point out—one that falls outside the limits of human understanding. It is ultimately the same as the question of how freedom is possible, of how we can initiate a new series of conditions. The inability to answer these questions does not count as a failure, since it is not in our capacity to know such things. All of this goes toward explaining why *Groundwork* III pushes the question of moral feeling to the “outermost boundary” of practical philosophy, and why the deduction makes no mention of feeling or sensibility.

My claim is that *Groundwork* III should have made mention to feeling, because it is only through our capacity to feel that can we understand our interest in morality. Of course, for this claim to have any foundation I will

first have to establish what the aim of Kant's deduction really is, after which I will work in detail through the deduction's key steps (again, the reason why the next three chapters are the most technical, and perhaps the most important, part of what's to come). The challenge here, which I just hinted at, is that the reading I want to put forward runs against scholarly consensus, which tells us that the aim of *Groundwork* III is to ground the validity of the moral law for rational beings in general. The ambition of the argument, on this view, is to show that human beings, as imperfect as we are, have access to an intelligible "standpoint" or "world," whereby we recognize our autonomy under the moral law.

I believe this view is incomplete. Kant is concerned with grounding the validity of the moral imperative for *human* beings in particular. As I will argue, the key step of Kant's deduction is not in his discussion of "standpoints" or "worlds" in the third subsection, but in the fourth subsection, entitled: "How is a Categorical Imperative Possible?" Another familiar view is that Kant attempts to provide a strict deduction of the moral law (its objective authority as a principle of rational choice) from non-practical premises—a laborious task he supposedly abandoned by the time of the second *Critique*. In the later work, we are told, Kant gave up on a deduction and instead boldly called moral authority a "fact of reason." I will not be addressing Kant's doctrine of the "fact of reason" in any detail; that would deserve a separate study. What I do have to say, though, moves away from this familiar interpretation. As I see it, the direction of Kant's arguments in *Groundwork* III and the second *Critique* are fundamentally the same: both proceed from morality to freedom, not the other way around. I will make some progress in defending this reading of Kant's deduction, but again, it is not my main focus. My first objective, then, is to clarify the aim of *Groundwork* III (Chapter 2), to analyze its key step in subsection four (Chapter 3), and finally, to diagnose its failure (Chapter 4). The implication that will arise from this is that Kant did not reverse the direction of the second *Critique*, but rather added an essential step: namely, an account of how the moral law appears to us in the form of an incentive.

Whether any of this brings us closer to the philosophers' stone is a question I will leave, in the end, to the reader.

Part I

The Problem

CHAPTER 2

What Moral Question?

Kant's deduction of the supreme principle of morality admittedly does not address (or even appear to take seriously) some of the more extreme forms of skepticism about value which have dominated twentieth-century meta-ethics. To these skeptics, the Kantian argument I have been reconstructing may be a big let down.

— Allen Wood (1999).

His argument, then, amounts to an answer to the contemporary question “Why be moral?” But Kant's aim is easily obscured by the fact that his imagined audience is not the sort of moral skeptic with which we are most familiar today. Kant does not see himself as addressing, for example, those who are indifferent to morality and demand that philosophy supply them with a motive to be moral; for Kant's own theory denies that anyone rational enough to ask the question could really be so indifferent.

— Thomas E. Hill, Jr. (1985).

2.1 Preliminaries

SKEPTIC or not, I am sure most readers share Allen Wood's thought that the final section of Kant's *Groundwork* is a big let down.

Part of the difficulty is that Kant is not clear what the aim of his argument is and what's riding on its success or failure. To make matters worse, what Kant does say about *Groundwork* III can give the reader false hope—for example, when he speaks of establishing “the supreme principle of morality” whatever that means. Our sense of a let down, then, is partly

Kant's fault, and partly our own. For the argument of *Groundwork* III is, despite appearances, less ambitious than we think. This forces us to confront a new set of questions. If Kant isn't trying to refute the skeptic indifferent to morality, and is not interested in the traditional question, "Why be moral?," what question *is* Kant interested in? And who, if not the radical skeptic of normative value, is he speaking to?

In this chapter I will argue that the overwhelming disappointment most readers of *Groundwork* III feel is the result of a false expectation: namely, that Kant is trying to convince the skeptic why he should be moral, using premises only he, the skeptic, would accept. The closest Kant comes to "refuting" skepticism is in *Groundwork* II, but there the issue at stake is methodological. Kant wants to approach the concept of duty from rational sources, and his main target is the popular moralist who develops ethical principles from observation. Interestingly, when the question of justification arises at the end of *Groundwork* II, Kant makes no indication of returning to the skeptical concerns he brushed aside earlier. For he addresses himself to "anyone" who does not think morality is an illusion (G 4:445).

My organization is as follows. I will first elaborate on the *Übergang*, the "shift" or "transition," that Kant wants to bring about in *Groundwork* II, specifically from "popular philosophy" to what he calls a "metaphysics of morals." This will provide a framework for understanding Kant's broader strategy of using skeptical threats to motivate a "shift of framework" in the text, and I will argue that the worry Kant voices at the end of *Groundwork* II is a tactic for motivating his critical distinction between "two worlds" in *Groundwork* III. While these considerations will remain mostly general, I will spend some time clarifying the "synthetic a priori" proposition Kant says he needs to establish. Most Kant-interpreters argue—mistakenly, I believe—that this proposition refers to the *validity* of the moral law. On my reading, it refers to the particular guise morality takes for beings like us: the guise of an *imperative*. Stepping back, I will then devote the rest of the chapter to a number of puzzling interpretive questions, ranging from Kant's use of the term "deduction" to his claim that we can "defend" the authority of the moral law, and its force as an incentive, at the "outermost limit" of practical philosophy.

2.2 Empirical Skepticism

Let's first consider Kant's intended audience in the *Groundwork*. It seems that he is speaking to some kind of empiricist, for example, the two unnamed philosophers that appear at the beginning of *Groundwork* II.¹ The

¹ Kant does not clearly distinguish between the two varieties of empirical skepticism. Regarding the first, he says: "Hence there have at all times been philosophers who have

first does not deny the truth or correctness of the moral law. His doubts concern the human will. In his view, the human will lacks the capacity to live up to the law's commands. His skepticism concerns the existence of a moral disposition—a disposition to act out of respect for the law. Kant describes this philosopher as speaking with “deep regret” that human nature is “noble enough to take as its precept an idea so worthy of respect but at the same time is too weak to follow it” (G 4:406). It is not that he thinks the moral law is impure (for example, that duty is a disguised concept of self-interest), but that the human will is unfit to incorporate respect for the law into its maxims. The thought that troubles him does not concern the content of the law but of our capacity, as human beings, to take the law as our sole incentive for acting. Kant's second skeptic raises a different concern. Rather than express a “deep regret” over the frailty of human nature, this philosopher harbours a malicious wish to “ridicule all morality as a phantom of a human imagination [*Hirngespinnst*] overstepping itself in self-conceit” (G 4:407). In his eyes, the evidence we have of people's conduct is decisive. The fact that we cannot find a single case where the moral law alone determines one's will is proof that the concept of duty is empty.

These skeptics, Kant argues, commit the same philosophical error: both rest their claims on an appeal to experience. For the motivation skeptic, it is the evidence of our weakness of will—the fact that we often lack the requisite energy to act on our rational considerations—that questions our capacity to act out of respect for the law. For the content skeptic,² it is the evidence of our selfishness—the fact that we often disguise our self-interested motives under the guise of morality—that questions moral authority in general.³ In *Groundwork* II, Kant uses a method of exaggeration to destabilize these concerns. It is true, he says, we can never know for certain whether our actions have a moral motive or a motive drawn from self-interest. It may very well be the case that all our actions are but effects of the “dear self,” which is, Kant says, “always turning up” (G 4:407). But Kant's point is

absolutely denied the reality of this disposition in human actions and ascribed everything to more or less refined self-love” (G 4:406). On the next page, he then adds: “Moreover, one cannot better serve the wishes of those who ridicule all morality as a mere phantom of the imagination overstepping itself through self-conceit than by granting them that concepts of duty must be drawn solely from experience (as, from indolence, people like to persuade themselves is the case with all other concepts as well); for then one prepares a sure triumph for them” (G 4:407).

²The distinction between “motivation” and “content” skepticism comes from Korsgaard (1986). I am not saying that the two forms of empirical skepticism Kant addresses in the *Groundwork* map perfectly onto Korsgaard's account. The sense in which Kant's second skeptic raises doubts about content refers to his doubts about the existence of rational requirements of action.

³See my (2009) for a more detailed account of the problem of self-deception in Kant's ethics.

that it would be wrong to draw any conclusions from this. The fact that we cannot find examples of our moral disposition or the necessity of the law does not ground any argument for their non-existence. By exaggerating the skeptic's worry, then, Kant can uncover the skeptic's error. This is the error of empiricism (G 4:419).

Kant's broader point is that we cannot settle questions about morality on the basis of experience. We cannot support claims about how we should govern ourselves (for example, that we should harmonize our inclinations) on the basis of how we act. If we try to respond to the skeptic by offering an example of the law's necessity, or of our strength of will, we will commit the same mistake the skeptic has made. We will develop our stance, which may purport to be anti-skeptical, on the assumption that debates over morality can be carried out in empirical terms. This would draw us back into the very dialectic we wished to escape: our anti-skepticism would leave the fundamental problems of skepticism unexamined. What's at stake in the beginning of *Groundwork* II, then, is a question of methodology. Kant's objective is to show that our approach to the study of morals cannot be based on experience. The shortcoming of "popular philosophy" is that it fails to specify the standards of its own investigation, so that what results is a "hodgepodge of patchwork observations," some drawn from reason, others from experience (G 4:409).⁴

By demanding a priori standards, Kant is urging a "transition" (*Übergang*) from popular philosophy to what he calls a "metaphysics of morals." The refutation of empiricism is thus a preliminary for Kant's a priori investigation of morality. The investigation is "merely analytic"—as Kant puts it, "we leave it undecided whether what is called duty is not as such an empty concept"—yet what emerges is rather unexpected. From the concept of duty we can derive a single formula of the categorical imperative: "act only in accordance with that maxim through which you can at the same time will that it become a universal law." But this formula expresses nothing other than the condition by which we can govern ourselves, i.e., the condition of autonomy. So by "mere analysis" we can show that autonomy of will is the "supreme principle of morality," and that what the categorical imperative commands is "neither more nor less than just this autonomy" (G 4:440). Had we let empiricism set independent limits on our investigation, we would not have been in a position to see this.

Despite Kant's frequent remarks that his method in the first two sections

⁴What disturbs Kant about popular philosophy is that it is incapable of presenting the unifying principle of morals, so that what it accounts for is a mix of various insights: "at one time the particular character of human nature (but along with this also the Idea of a rational nature as such), at another perfection, at another happiness; here moral feeling and there fear of God; something of this and also something of that" (G 4:410).

is “merely analytic,” it is hard not to find his final statement in *Groundwork* II shocking. Kant tells us that it remains possible that morality is nothing more than an illusion, a “phantom of the imagination.” But why is this still a possibility? And why is Kant telling us this now? We might get the impression that the spectre of skepticism from *Groundwork* II has returned. But what Kant says suggests something else. He writes: “whoever holds morality to be something and not a chimerical idea without any truth must also admit the principle of morality brought forward” (G 4:445). Now, this would be a gross distortion of the motivation skeptic discussed earlier, who does not hold morality to be something, at least not for frail creatures such as us. Nor would this represent the content skeptic who takes delight in seeing morality stripped of its grandeur and exposed as a “vain delusion.” *Who*, then, is Kant speaking to?

Let’s step back for a moment. Kant thinks we secure a victory for skepticism if we develop the concept of duty from experience (G 4:407). The problem with popular moral philosophy is not only that it leaves us with a “patchwork” of “half-rationalized” principles, but that its methodology consists of observing others and finding examples of virtuous conduct. As a result, it is left defenceless against those who wish to denounce morality as an illusion, because experience teaches us that people usually act out of selfish motives, that they pursue what they find privately agreeable. By adopting an a priori standard of investigation, Kant thereby *rejects* the presupposition on which the traditional problematic rests, i.e., that experience is our testing ground for the concept of duty. Thus the closing remarks from *Groundwork* II are addressed to “anyone” who has, *with Kant*, made the transition to a metaphysics of morals.

2.3 From Metaphysics to Critique

We know from the first *Critique* (from the section on “Methodology”) that Kant thought of himself as uncovering skeptical claims that no one had raised before him, claims lying dormant at the borders of reason. He was also of the opinion that a skeptical claim never expires, and that it does not matter if someone has yet to put the claim to us, threatening our peace of mind. The claim remains, as a possible source of unrest, which the history of reason may at any time suffer. And that is why we must examine those claims ourselves, Kant says, and trace them to their origin. One way to do this is through a “skeptical method.” It is the method of giving our opponent every tool of argument and providing him the best position in battle. The goal is to aid him, not avoid him. And should he fail to exploit this advantage, Kant says, we must take up his position ourselves. *We must raise a problem the skeptic has yet to think.* Only by this can we trace a skeptical

claim to its origin, and, as Kant says, “extirpate” its root (KrV A779/B806). This is, I believe, what we find Kant doing in *Groundwork* III.⁵

The *Groundwork* is organized into three sections.⁶ Each section contains a “transition” or “shift” of perspective.⁷ Kant uses a skeptical worry to highlight the limitations of the current perspective, and so motivate an *Übergang* to a new one. In *Groundwork* I, we find that common cognition leads us to believe that the highest practical vocation of reason is to produce a good will. When we analyze the concept of a good will by representing it under the concept of duty (Kant says, “under certain subjective limitations and constraints”), we are able to clarify the formula which common reason tacitly has “before its eyes as the norm for its appraisals,” namely, the formula of universal law (G 4:404). And yet, common reason is also susceptible to the demands of sensibility, and it can easily delude itself into thinking that a principle for satisfying one’s wants and wishes should have priority over the moral law. As a result, there arises a propensity within the ordinary person “to rationalize against those strict laws of duty and to cast doubt upon their validity, or at least upon their purity and strictness” (G 4:405). This requires us to “seek help in philosophy” (G 4:405).

As we have seen, however, philosophy leaves us in a “precarious position” if it does not first establish an a priori method of investigation.⁸ This is the downfall of popular moral philosophy. It indirectly secures a victory for the skeptic who wishes to deny either our capacity to act out of respect for the law or the status of the law as a requirement of reason. By developing the concept of duty from experience, it allows the skeptic to conclude, for example, that we should do our best to satisfy our inclinations because observation teaches us that we only act out of self-interest. Thus we need a new framework for philosophical cognition, a shift to *metaphysics*. Kant himself characterizes this shift as our only means of fighting against the “laxity” of human reason “which seeks its principle among empirical motives and laws (G 4:426).⁹

⁵Kant also thought we could use a similar strategy for responding to the dogmatist. In the second *Critique*, for example, he speaks of those who boast of having theoretical proof of freedom: “They want to *prove*: very well, let them prove, and the critical philosophy lays all its weapons at their feet as the victors. *Quid statis? Nolint. Atqui licet esse beatis* [from Horace Satires 1.1.19: a god, having given men the opportunity to change places with each other, says, “what are you waiting for? They are not willing. Yet they might be happy”] (KpV 5:6n).

⁶In the Preface Kant writes: “I have adopted in this work the method that is, I believe, most suitable if one wants to proceed analytically from common cognition to the determination of its supreme principle, and in turn synthetically from the examination of this principle and its sources back to the common cognition in which we find it used” (G 4:392).

⁷This expression comes from O’Neill (1989, ch.3).

⁸Cf., (G 4:425).

⁹See Tenenbaum (2003) for more on the relation between “ordinary temptations” and

All of this brings us to ask: If the propensity to rationalize against the purity of moral requirements motivates the transition from common to philosophical cognition, and if the threat of empiricism motivates the transition from popular philosophy to metaphysics, *what threat* must we confront at the level of metaphysics, and *what shift* of frameworks is required here? Recall that a metaphysics of morals examines the concept of duty on rational grounds. As Kant sees it, this is the only way we can “behold virtue in her proper form,” that is, “stripped of any admixture of the sensible and of any spurious adornments of reward or self-love (G 4:426n). The method of *Groundwork* II also reveals a link between the categorical imperative and autonomy of will—yet Kant continually reminds us that the link is only conceptual. A metaphysical method can only uncover analytic relations between our ordinary pre-philosophical notions of morality. It is unable to show that these notions apply to us. By saying that morality may be empty, then, I think Kant is using a skeptical threat to motivate a new shift of frameworks, from metaphysics to what he calls a “critique of pure practical reason.”

2.4 Transcendental Skepticism

Kant spent years labouring on a “deduction of the categories,” which he says cost him the “most effort” in the first *Critique* (KrV Axiv). Unfortunately, Kant left us few clues for understanding the path he took to arrive at his conception of a deduction in *Groundwork* III. It would be futile to try to compare these texts in any detail, but there are some broad similarities I would like to draw attention to.

In the first *Critique*, Kant proceeds to show, first, that the categories are functions of the understanding (the Metaphysical Deduction), and second, that the categories must apply in experience (the Transcendental Deduction). In the Transcendental Aesthetic, Kant argues that space and time are not external to us, but are pure forms of sensibility, our capacity to receive empirical data. In the Transcendental Analytic, he then wants to show that the understanding plays an active role in unifying this data under pure concepts, or categories. Kant’s first task in the deduction is to show that the understanding really does contain pure concepts, and he does this by deriving the categories from the logical forms of judgment. This would not have been an option if he had let empiricism set antecedent limits on the question of how experience is possible.

“speculative mistakes.” Although Kant is not always clear about this distinction, the temptation that besets ordinary reason involves a *re-ordering* of the claims of self-love or those of morality; whereas the mistake of empiricism, for Kant, is to think that all practical reason *reduce* to those of self-love or happiness.

So we have two questions: “Jurists, when they speak of entitlements and claims, distinguish in a legal manner between the questions about what is lawful (*quid juris*) and that which concerns the fact (*quid facti*), and since they demand proof of both, they call the first, that which is to establish the entitlement or the legal claim a **deduction**” (KrV B117). The question *quid juris* raises trouble for concepts that are independent from all experience. “Fortune” and “fate,” for example, “circulate with almost universal indulgence,” but no one can provide clear grounds for their lawful use, “either from experience or from reason.” By contrast, empirical concepts do not provoke suspicion of their lawfulness because we always have experience at hand to test them (for example, the concept of a dog). The problem is that nothing in experience tests the correctness of the categories, which are completely a priori. This opens up a new skeptical worry for Kant. The concept of cause, for example, may be nothing more than a function of thinking, not of experience:

For appearances could after all be so constituted that the understanding would not find them in accord with the conditions of its unity, and everything would then lie in such confusion, e.g., in the succession of appearances nothing would offer itself that would furnish a rule of synthesis and thus correspond to the concept of cause and effect, so that this concept would therefore be entirely empty, nugatory, and without significance. Appearances would nonetheless offer objects to our intuition, for intuition by no means requires functions of thinking. (KrV B123)

This is why the categories require a special explanation of how they relate to objects. The skeptical worry is “transcendental,” because it arises *after* Kant has rejected empiricism about knowledge and secured the a priori origin of the categories. We can say that if the categories are products of the understanding, and if empirical data is received through intuition, Kant still has to explain how the categories apply—and serve to unify—everything that appears in intuition.¹⁰

What is the parallel task in *Groundwork* III? Kant leaves us two clues. At the end of *Groundwork* II, he says the question we need to answer is

¹⁰Most commentators agree that the transcendental deduction contains two steps. After showing that the categories are valid for a discursive intellect in general (i.e., an intellect that must cognize objects in a manifold), Kant must face the remaining task of showing that the categories apply to a human intellect (i.e., an intellect that must cognize objects in space and time). This may parallel the argument structure of *Groundwork* III. After showing that the categorical imperative is valid for rational agents in general, Kant is faced with the remaining task of establishing the validity of the categorical imperative for human agents in particular (i.e., rational agents who are additionally affected by sensible influences). I would like to thank Christian Onof for helping me clarify this parallel. I will come back to this topic in section 5.1.

how a “synthetic a priori practical proposition” is possible (G 4:445); and the fourth subsection of *Groundwork* III is put as the question, “How Is a Categorical Imperative Possible?” After summarizing the argument, Kant writes: “and this *categorical* ought represents a synthetic proposition a priori ... [T]his is roughly like the way in which concepts of the understanding, which by themselves signify nothing but lawful form in general, are added to intuitions of the world of sense and thereby make possible synthetic propositions a priori on which all cognition of a nature rests” (G 4:454).

If we take Kant at his word, then the equivalent of the deduction of the categories in *Groundwork* III is to be found in Kant’s answer to the question of how our experience of moral constraint is possible. Interestingly, this is not an answer to the question “Why be moral?” if by that we understand “Why abandon self-love?” or other questions that request a reason or incentive to be moral. Indeed, the question Kant thinks is particularly pressing and in need of an answer already assumes one’s position “inside” morality. And what Kant thinks we need to explain is the experience of duty peculiar to beings like us, the experience of “constraint” (*Zwang*) or “necessitation” (*Nötigung*). Thus the transcendental question of *Groundwork* III is, “How is our experience of morality possible?” which is, as Kant puts it, “roughly the same” as the question of the first *Critique*, “How is experience possible?”

2.5 A Good (Human) Will

There is little agreement among interpreters what the task of *Groundwork* III amounts to. I do not want to get caught up in the complexities of this debate, but from what I have said so far my position should be more clear. The “synthetic a priori” proposition Kant says we need to establish concerns the imperatival character of the moral law, not its objective validity. In fact, Kant thinks the objective validity of the moral law follows unproblematically once he can establish our status as rational agents—a point we will see in Chapter 3. For Kant, what’s puzzling is the phenomenon of obligation itself, and part of the puzzle is that it concerns the relation *we* have to morality.

What Kant has to say about moral obligation follows from his definition of the will, so this is a good place to start. In *Groundwork* II, Kant writes: The “will” (*Wille*) is “the capacity to act *in accordance with the representation* of laws, that is, in accordance with principles” (G 4:412). As a capacity to act on the representation of laws or principles, *Wille* is nothing other than “reason” (*Vernunft*) in its practical capacity. So the will is nothing other than practical reason, as Kant sees it (G 4:412). But what does this have to do with moral obligation? For one thing, we can now think about an

ideal being that would *not* experience morality in the guise of an imperative. Just by analyzing the concept of the will (as practical reason), Kant thinks we can think about a perfectly rational or “holy” will. His claim is that a being who perfectly controls its will would recognize laws of action as both “objectively necessary” and “subjectively necessary” (G 4:412). This echoes the law/maxim distinction from *Groundwork* I: a law (*Gesetz*) is an objective practical principle valid for the will of every rational being, whereas a maxim (*Maxime*) is a subjective principle of “volition” (*Wollens*). So, if the will just is practical reason, it follows that a perfectly rational being would form its subjective principles of choice *according to objective principles*. It would be perfect because it would never encounter subjective grounds of action deviating from reason.

Kant’s point is that the *inclinations* present us with such grounds. We are “sensibly affected.” What appears agreeable to us, i.e., a representation that influences us *by a subjective feeling of pleasure*, is not always in conformity with reason. From this it follows, again by definition, that a being with *Wille* but lacking any capacity to be affected by representations of pleasure would always act in the way of reason. For such a being we would say that objective practical laws describe the activity of its will. By contrast, a being with the capacity to be affected by representations of pleasure, a “sensibly affected” being, would not always act in conformity with reason. It would also have a tendency to form maxims on the basis of inclinations, and those would only generalize to its own set of interests. So what a perfectly rational being represents as “subjective necessary” an imperfectly rational being will represent as “subjectively contingent” (G 4:413). In the case of the latter sort of being, by which Kant means a human being, we would say that whatever reason derives from an objective law must appear to it as an imperative, i.e., an action it “ought” to perform. This explains why Kant thinks a perfectly rational being would only stand in a relation of “necessity” (*Notwendigkeit*) to the moral law, rather than one of “necessitation” (*Nötigung*) (G 4:413). In the former case, he says, “the ‘ought’ is out of place here” (G 4:414).¹¹

¹¹ I want to suggest that the holy will is an “ideal” of practical reason. “Ideals” represent individual kinds in complete alignment with an idea. According to Kant, ideas are generated by reason in its search for the supreme unifying principles of concepts. By contrast, ideals exert practical force: they serve as models for emulation. Thus the Sage serves as a model of emulation for the agent who aims toward self-governance. The Sage is in complete alignment with the idea of virtue, which is the perfection of autonomy. Along these lines, we can say the idea of freedom represents the perfection of practical reason. Thus a holy will represents an individual in complete alignment with the idea of freedom. A holy will necessarily forms its subjective rules of action (i.e., its maxims) according to objective principles. And the representation of this activity serves as an archetype or model for beings like us, beings who form maxims on the basis of incentives. In this way, a holy will addresses us with our goal toward self-governance in its completion. For Kant, ideals satisfy a need of reason. By individuating pure concepts, they present us with standards of

While abstract, this distinction throws light on Kant's otherwise confusing remarks from *Groundwork* III. In the first paragraph, he defines freedom as the property of a rational will to be "efficient independently of alien causes determining it" (G 4:446). The definition is negative because we do not yet know what this property is. All we know is that the efficiency of a rational will cannot draw on causes that lie outside of it. What property then makes a rational will efficient? It must be a determining ground of some sort, for the idea of an efficient causality without a determining ground is contradictory. That is why Kant calls the idea of a lawless will an "absurdity." The will as the causality of a rational being needs a determining ground, as any causality does. But given the negative definition above, this causality cannot come from outside the will—from a law of nature, for example. There is only one alternative left in Kant's view: the causality must lie within the will itself. Thus the will's property to be causally efficient is the property to be its own law. And the property of the will to be its own law is nothing other than autonomy, as defined earlier (G 4:440). Kant then points out that the principle of the will to be a self-sufficient law is the principle: "act only on maxims that are fit to serve as universal laws"—and that is the principle of morality. So he concludes, in the second paragraph, that "a free will and a will under moral laws are one and the same" (G 4:447).

This conclusion may appear strange, but I think we are in a position to make sense of it. Kant told us earlier that to establish the principle of morality we must depart from the procedure of conceptual analysis developed in sections I and II. Now, in *Groundwork* III, he claims that morality "follows" from the concept of freedom. Why, then, do we need a deduction of the moral principle? Kant says:

the principle of morality—that an absolutely good will is that whose maxim can always contain itself regarded as a universal law—is nevertheless always a synthetic proposition; for, by analysis of the concept of an absolutely good will *that property of its maxim cannot be discovered*. (G 4:447; my emphasis)¹²

This is even more confusing, and things are quickly getting technical. But in light of Kant's distinction between the holy and human will, I think we can

comparison. "These ideals," he writes, "even though one may never concede them objective reality (existence), are nevertheless not to be regarded as mere figments of the brain; rather, they provide an indispensable standard for reason, which needs the concept of that which is entirely complete in its kind, in order to assess and measure the degree and the defects of what is incomplete" (KrV A569/B597).

¹² "schlechterdings guter Wille ist derjenige, dessen Maxime jederzeit sich selbst, als allgemeines Gesetz betrachtet, in sich enthalten kann, denn durch Zergliederung des Begriffs von einem schlechthin guten Willen kann jene Eigenschaft der Maxime nicht gefunden werden."

sketch an answer. If the good will refers to the will of a rational being in general (i.e., a will unaffected by sensibility), then we could derive the principle of morality directly from its concept. Indeed, for such a being the principle of morality would be self-descriptive (i.e., what it represents as “objectively necessary” will also appear “subjectively necessary”). After all, a good will lacking inclinations will always form its maxims in line with the law. In this way, I think the synthetic “ought” of the law only applies to the will of an imperfectly rational being, i.e., a *human* will. And so the proposition Kant says we need to establish by way of “deduction” only concerns a good will under human conditions. We require a special justification of morality because, from the human perspective, the possibility of an unconditional imperative is so strange that we find ourselves susceptible to doubt.

2.6 Deduction or Defence?

Let me tie together the main strands of this discussion. I have argued that Kant’s response to empirical skepticism in *Groundwork* II is meant to clear room for a new methodological possibility—namely, that of providing an a priori standard for evaluating the categorical imperative. Kant’s rejection of empiricism at this stage in the discussion prepares the way of a metaphysics of morals, which suspends questions of our motivational capacities as human beings. In this light, I argued that Kant’s skeptical remarks at the end of *Groundwork* II are best understood as a consequence of adopting a transcendental method of moral inquiry. For the question arises how the categorical imperative *applies* to us. A parallel question is also at the heart of Kant’s deduction of the categories, which also follows from a metaphysical derivation of the concepts of pure understanding from the forms of logical judgments. In each case, a special explanation is required to show how an a priori category has application, either to objects of experience, or to ourselves as agents. In *Groundwork* III, Kant makes this clear with his distinction between the holy and human will, suggesting that the question of the synthetic “ought” only arises for beings who, like us, are affected by non-rational incentives. These considerations help show why the focus of *Groundwork* III is on *our* relation to morality, and Kant’s worry is that this relation (and so the categorical imperative itself) may be illusory.

At this point one question jumps out. What does Kant mean by a “deduction” (*Deduktion*)? For one thing, he does not mean a strict syllogistic argument, one that adheres only to theoretical premises (a “deduction” in the modern sense of the term). It is, rather, in the sense common to the legal procedures of Kant’s day: an account meant to secure a title whose legitimacy has been called into question.¹³ Such legal deductions account for

¹³For more in-depth discussions of the legal character of Kantian deduction, see Henrich

a title's legitimacy by providing certificates of its origin; for example, a land title. I do not want to follow the juridical metaphor too closely, but I do think we can see Kant adopting a similar approach. For at least part of his strategy of justification in *Groundwork* III to show that the moral "ought" is valid. By tracing, or deducing, the phenomenon of constraint to the activity of our own will, we can then answer the question *quid juris*. In this sense, we can produce a transcendental certificate of origin that shows the moral "ought" is a product of practical reason, not of the imagination.¹⁴

Sadly, Kant is not explicit about the juridical nature of his arguments, but at least two elements are basic to his approach: (i) that the claim under question be held valid until proven otherwise; and (ii) that the final verdict (valid or invalid) be determined to the satisfaction of a judge or impartial witness, not to that of the opposing party. In *Groundwork* III, for example, Kant says we can defend the idea of freedom, even though it presupposes something inexplicable to theoretical cognition—i.e., a noumenal causality. For the skeptic would need to have intellectual intuition to disprove its existence, and that is impossible. What the skeptic requires to undermine our claim is, in fact, unavailable to him or anyone else. So while we cannot offer a theoretical explanation of freedom that would place its status beyond doubt, Kant thinks we can defend the idea against those "who pretend to have seen deeper into the essence of things and therefore boldly declare that freedom is impossible" (G 4:459). The advantage is on our side here. As Kant puts it: "the opponent should therefore prove. But since he no more knows something about the object that is doubted which would establish its non-being than does the former, who asserts its actuality, here an advantage on the the side of he who asserts something as a practically necessary presupposition (*melior est conditio possidentis*)¹⁵ is revealed" (KrV A777/B805). The condition of the possessor is better, in Kant's view, because he is absolved of having to provide any further explanation. This is philosophy in its negative role, as "defence" (*Vertheidigung*).

Yet defence for Kant has both a theoretical and practical use. With the help of transcendental idealism, Kant thinks he can show that the contradiction between freedom of will and the complete determinism of nature is only apparent. The determination of nature only pertains to appearances, but we cannot legitimately infer the status of things-in-themselves from appearances (for example, that every condition must proceed in a causal chain of conditions). We can thus defend the claim to freedom against the charge

(1989), Proops (2003), and Franks (2005)

¹⁴However, speaking of a priori *origins* is misleading, since Kant thinks we can answer the question *quid facti* by tracing the concept of the categorical imperative to the definition of autonomy.

¹⁵"The condition of the possessor is the better."

of logical impossibility, and this is a task Kant thinks “is incumbent upon speculative philosophy only so that it may clear the way for practical philosophy.” “Hence,” he continues, “it is not left to the philosopher’s discretion whether he wants to remove the seeming conflict [between freedom and determinism] or leave it untouched; for, in the latter case the theory about this would be *bonum vacans*, into possession of which the fatalist could justifiably enter and chase all morals from its supposed property, as occupying it without title” (G 4:456). This is a speculative problem. As Kant sees it, what’s at stake here is the path from theoretical to practical philosophy. He thinks it is the philosopher’s responsibility to defend the logical possibility of freedom so that this path may remain *open*. This is essential for all morality, because the idea of freedom only acquires positive significance from a practical point of view.

Kant is more clear about this in the second *Critique*, but all the basic points are at work in *Groundwork* III. In both texts, Kant maintains that we only have reason to think of ourselves “under the idea of freedom” when we are immediately conscious of the moral law. For the moral law requires nothing other than our complete independence from sensible influences. In this way, the presupposition of freedom is consistent with the unconditional imperative of morality. Of course, establishing freedom on the grounds of its consistency may not satisfy the skeptic, as the opposing party, who wants to say our will is dictated by a hidden mechanism of nature. But under the juridical constraints Kant’s argument adheres to, it does not matter if the skeptic is satisfied. What counts is that we can bring our case to the satisfaction of an impartial judge or anyone else who recognizes the right of an undisputed title. Indeed, one of Kant’s recurring strategies to place the ideas of morality (e.g., freedom, practical interest) beyond the reach of skepticism, not by supporting them with theoretical explanations or “proofs,” but by showing their internal consistency from a practical standpoint. His strategy is not simply to neutralize the opposition’s challenge, but is to secure a positive verdict, that our claim is valid.

Stepping back once again, we can see that Kant is struggling with at least *four* questions in *Groundwork* III. I have argued that the main question he is pursuing concerns the particular guise morality takes for sensible rational beings like us, the guise of an imperative. But we have also touched on another “transcendental question” Kant thinks is pressing and in need of answer: the question of whether we possess *Wille*, i.e., whether we really are rational agents. So we have two questions in need of deduction. We can ask whether we really are rational agents, and since this requires a free use of will, the question here takes a transcendental form: How is freedom possible for us? We also have the transcendental question concerning how our experience of moral obligation is possible. I will call these the Question

of Identity and the Question of Obligation respectively.

The Question of Identity: How is freedom possible?

The Question of Obligation: How is moral obligation possible?

There are, however, two further questions that Kant thinks are essential for morality *yet beyond deduction*. They bring us to the “outermost limit” (*außersten Grenze*) of practical philosophy. One concerns how we can take an interest in morality, or how the mere thought of duty can produce a feeling in us to conform our actions to reason. The other concerns how, or why, we prefer the law of reason, as the moral law, over any material principle like self-love. I will call these the Question of Motivation and the Question of Authority respectively.

The Question of Motivation: How is moral motivation possible?

The Question of Authority: How is moral authority possible?

Let me say a bit more about these last two questions. They are “incomprehensible” for Kant, but for different reasons. Kant is not clear about this in *Groundwork* III, and later—in Chapter 4—we will see why this gets his argument into trouble. But without going into that here, I simply want to clarify his position. For Kant, our consciousness of the law’s authority is a “fact” (*Faktum*) in the juridical sense—a “deed,” or “something done.” It is irreducible to the *act* of reasoning from considerations of a maxim’s validity. We are therefore unable to deduce the principle’s authority, not because it lacks an a priori origin in reason, but because it is the form of reason in its practical capacity. As a result, there is no non-moral route of reflection, no “subtle reasoning,” that could produce this *Faktum* in us. In this way, we have a right to figurative speech: we can say, as Kant says, that the authority of the moral law is “given” to us.¹⁶ On the other hand, we are unable to deduce the interest we take in morality for a different reason. The *ground* of our interest in morality is in pure practical reason, but the interest itself is an *effect* manifested at the level of the world of sense. We are unable to provide a deduction of moral interest, then, because we lack insight into its a priori origin: we cannot see how a causal connection between reason and sensibility is possible. So whereas the law’s authority is irreducible to the activity of willing the law, a “fact” of pure reason, the interest we take in the law is connected (in an inexplicable way) to our faculty of desire. The first is “inexplicable” because no process of reflection will lead us to its ground: it is already the ground of pure practical reason. The second is “inexplicable”

¹⁶“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 [*Willkür*] over to it” (R 6:26n).

because the possibility of causal interaction between reason and sensibility is beyond the scope of our understanding. And yet, because they are inexplicable, *they are beyond refutation*. We cannot claim insight into the necessity of the moral law, or the interest we take in it. But for that very reason, we cannot claim insight into their non-existence or impossibility. The condition of the possessor is the better, as Kant says. We can then defend these moral presuppositions because they lie at the “outermost limit” of practical philosophy.

2.7 The “Fact of Reason”

I have been skirting around the question of whether Kant changed or repeated his deductive strategy in the second *Critique*. A common view in the literature is that Kant went to great lengths in *Groundwork* III providing a strict deduction of the moral law from non-moral premises, a laborious task he then abandoned by the time of the second *Critique*. In the later work, we are told, Kant reversed the direction of his argument, starting instead from our consciousness of morality as a “fact of reason.” As a strategy for addressing the worry that morality may be “high flown fantasy,” many commentators believe Kant resorted to dogmatism in 1788.¹⁷ In my view, there are important differences between these two works, but Kant’s answer to the Question of Authority is not one of them.

Much of what I am saying is clarified in the last paragraph of *Groundwork* III, where Kant says:

[B]y constant inquiry after the condition, the satisfaction of reason is only further and further postponed. Hence it restlessly seeks the unconditionally necessary and sees itself *constrained to assume it* without any means of making it comprehensible to itself, fortunate enough if it can discover only the concept that is compatible with this presupposition. (G 4:463; my emphasis)¹⁸

This passage echoes the opening lines of the Preface to the *Critique of Pure Reason*, in the First Edition, where Kant says reason is drawn to questions it cannot avoid, because they arise from its nature, but which it cannot answer, because they exceed its capacity for explanation (KrV A vii). What I

¹⁷This is an old criticism: from Hegel’s (1820) charge that duty is the “undigested lump” of revelation left in the stomach of reason, to Wood’s (2008) recent claim that Kant’s only resource for responding to the skeptic in the second *Critique* is “moralistic bluster” (p. 135).

¹⁸“Auf diese Weise aber wird durch die beständige Nachfrage nach der Bedingung die Befriedigung der Vernunft nur immer weiter aufgeschoben. Daher sucht sie rastlos das Unbedingt=Nothwendige und sieht sich genöthigt, es anzunehmen, ohne irgend ein Mittel, es sich begreiflich zu machen; glücklich genug, wenn sie nur den Begriff ausfindig machen kann, der sich mit dieser Voraussetzung verträgt.”

take Kant to be saying in the above paragraph is that practical reason has a way of avoiding this fate: namely, by adopting its own form as the sufficient determining ground of the will. This is what Kant means when he says reason is “constrained to assume” its own necessity without the means for comprehending it. Kant admits that we cannot say on what grounds the universal validity of a maxim is necessary, i.e., as a requirement of reason. There is no condition under which we can cognize its necessity. From the standpoint of theoretical reason, then, it appears the moral law is a concept without an object, hence “empty.” But the advantage of practical reason, in Kant’s view, is that it can adopt its own form as the condition of its necessity, although how we do this, and why, will remain a mystery.

Now, though, we seem to have raised a new problem. For what I have just described sounds like an arbitrary and unjustified assertion of authority, an assertion that would merely cut short reason’s search for the unconditioned, its search for the answer “Why? Why is the moral law necessary?” Kant seems to concede this point in his discussion of the “fact of reason” from the second *Critique* where he alludes to the tyrannical wife from Juvenal’s Satire 6.¹⁹ This is an odd way to characterize the authority of the moral law, since the wife from Satire 6 makes no attempt to hide the arbitrariness of her demands. She is not involved in a process of giving or asking for reasons. Indeed, the famous passage Kant cites, “*sic volo, sic jubeo*” (“What I command, I will”), was a popular phrase in the late Medieval and Early Modern period to describe the arbitrariness of political or religious authority (as when Luther refers to the authority of the Pope). There are many passages in the second *Critique* and elsewhere where Kant speaks of the law “giving” and “forcing” its requirements on the human will, so one might think the allusion to Satire 6 is fitting after all. The moral law commands unconditional obedience, and its command *stands for* its reason.²⁰

When we look at the allusion to Juvenal more closely, we can see that Kant may have been onto a structural similarity. The wife’s willful assertion is in response to a *series of questions* the husband puts to her. She demands the execution of a slave, and the husband asks for a reason. Why? “What crime has he done to deserve it?” the husband asks. “What witnesses are there? Who’s his accuser? Give him a hearing.” The wife ends these questions, not by giving the husband a reason—for that would presumably attract another question, *ad infinitum*—but by asserting her will. In the second *Critique* Kant cites only the first part of her answer, “*sic volo,*

¹⁹“However, in order to avoid misinterpretation in regarding this law as given, it must be noted carefully that it is not an empirical fact but the sole fact of pure reason which, by it, announces itself as originally lawgiving (*sic volo, sic jubeo*)” (KpV 5:31).

²⁰Proops (2003), and Kain (2010) have—unfairly, in my view—attributed this reading to Franks (1997). It seems clear to me that Franks is thematizing the aspect of violence in *Juvenal’s* Satire, not in Kant’s doctrine of the fact of reason.

sic jubeo” (“What I will, I command”), but following this Juvenal has her say: “sit pro ratione voluntas” (“My will is reason enough”). Of course, the context makes it obvious why her will is enough: she issues her desires as if they were unconditional laws. For Kant, the form of pure practical reason is enough because it is a source of unconditional laws. It is the categorical imperative.

There is a historical point worth exploring here. The “pro ratio voluntas” aphorism came to have a pejorative meaning by the Early Modern period, one often used to characterize the selfish absolutism of papal or royal authority. Nevertheless, scholars have observed that Late Medieval legalists and canonists attached a positive meaning to the phrase, according to which the will of the prince or pope was not above *ratio*, but in line with it.²¹ Jurists of the time also used the phrase in ordinary contexts to describe situations of unclear or vague legal action, whereby an individual’s *voluntas* could *step in* for a justifying reason. Gaines Post (1973) gives an example of a son refusing to accept his father’s inheritance because of the undue complications it would cause him. His displeasure at bearing the burden of the legal procedures following the inheritance would be enough reason for rejecting it (p. 162). Such cases did not violate the law but made it a matter of personal preference for how the outcome of the situation would unfold. However, there was one context in which sovereign authority could in fact violate the letter of the law, although its practice was beyond dispute. I am referring to the king’s practice of extending written acts pertaining to the dispensation of royal gifts, the royal appointments of justices and officials, etc., called “facta.” These *facta* were valid solely as a result of the king’s *unlimited voluntas*.²²

Whether or not Kant had this juridical history in mind, the philosoph-

²¹Post (1973) speculates that it was the early thirteenth-century decretalist, Vincentius Hispanus, who first applied “pro ratio voluntas” to papal authority in his glosses on Pope Innocent III’s letter to Germany. After Vincentius, the distinction between the reason of the law (*ratio iuris*) and the will (*voluntas*) became more commonplace. However, Post doubts that Vincentius knew the passage from Satire 6.219-224. He argues that the aphorism itself would have been commonly known and that it was Vincentius who “had the wit to apply it to the words of Pope Innocent III” (p. 165). However, I see no reason to doubt Vincentius’ knowledge of Juvenal’s work, since medieval glossators and commentators of Roman and canon law were well-studied in the classical authors and often turned to their texts for the purposes of interpreting law (Kantorwicz 1961). It is also true that canonists reacted to Innocent’s rhetoric by using hyperbole in their prose (Pennington 1984), and on this point Vincentius was not alone.

²²The specific case I have in mind involved the conditions under which a will was valid according to the requirements of Roman law. Normally, seven persons were required as witnesses. Yet the king’s *factum* could suspend this requirement. For if he was the sole witness, his testimony alone was sufficient to validate the will in court. Of course, matters of this sort did not reduce to the king’s *voluntas*, but they paralleled the substitution-of-reasons characteristic of the “pro ratio voluntas” aphorism.

ical issue remains a delicate one. The wife of Satire 6 treats her arbitrary impulses as if they were laws, and she takes them to replace any reason to convince her husband. She cuts off the series of questions for why her command to execute the slave is justified with her arbitrary wish. There is no further condition to back it up—to explain why anyone should obey her. Similarly, there is no further condition to back up the moral law’s authority—to explain why anyone should give it preference. But there is a difference. Nothing could justify the wife’s command because it is, after all, a *groundless* whim. When Kant says that reason is “constrained to assume” the moral law, however, he means that reason must take itself, *its own form*, as a sufficient ground of action. That is how reason *grounds itself* as practical. There is no further condition we can discover to explain this. Kant admits it is “strange enough, and has nothing like it in all the rest of our practical cognition” (KpV 5:31). Yet he says we find “confirmation” of it in our ordinary, pre-philosophical ways of thinking. “One need only analyze the judgment that people pass on the lawfulness of their actions in order to find that, whatever inclination may say to the contrary, their reason, incorruptible and self-constrained, always holds the maxim of the will in an action up to the pure will, that is, *to itself inasmuch as it regards itself as a priori practical*” (KpV 5:32; my emphasis). Just as the wife of Satire 6 issues her command in place of any further reason, or as the king issues a royal dispensation as he pleases, reason takes itself as the sufficient condition of its maxims.

2.8 Looking Ahead

At the beginning of this chapter I said that the sense of disappointment most readers of *Groundwork* III feel is due to a false expectation. If we go into the text thinking that Kant will provide the skeptic (or the skeptic *within us*) with an incentive to be moral, then we will surely feel let down. I have said that Kant’s aim is more modest than we think, despite its rhetoric of establishing the “supreme principle of morality.” Looking back, I may want to qualify that claim. Kant’s approach is modest only to the extent that he denies we can answer the question “Why be moral?” for someone apparently indifferent to morality. Only someone who actively adopts the moral law as a sufficient ground of action can, for Kant, experience a categorical “ought.” And that is why the central moral question of *Groundwork* III is not “Why be moral?” but rather “How is the moral ‘ought’ possible?” This is a modest question, in the strict sense. But Kant’s strategy for answering it is, perhaps, one of the most ambitious in the history of ethics. For as we have seen, Kant also thinks we can use a deductive strategy for securing our status as rational agents (the Question of Identity), and that we can defend

both the force of morality (the Question of Motivation) and its legislative standing (the Question of Authority). So the moral question is, for Kant, a question with four parts.

One of my general claims in this thesis is that the preoccupation in the secondary literature over the direction of Kant's argument (from morality to freedom? or freedom to morality?) obscures a more significant change in Kant's thinking—a change that occurs, not with his theory of practical reason, but with his theory of practical sensibility. Having clarified in this chapter the four-part structure of Kant's moral question, I now have the resources to develop this claim in more detail. I will argue in Chapter 4 that Kant's reason for placing the Question of Motivation on the “outermost boundary” (that is, *beyond deduction*) is symptomatic of his impoverished conception of sensibility in *Groundwork* III, a conception that will undergo a radical development three years later. As a way of leading up to this claim, I first want to examine Kant's answers to the initial two questions. In the next chapter, I will present Kant's argument for why we should think of ourselves as rational agents, and why we risk arguing in a circle if we fail to distinguish *two* ways of thinking about our agency. This will help explain Kant's reference to a “critique of pure practical reason,” and his worry about a “hidden circle” in his argument. With this in place, Chapter 4 will move to the centre of *Groundwork* III, the fourth subsection, where we find Kant's explanation of how our experience of constraint is possible.

Freedom and Rational Agency

3.1 Preliminaries

KANT left us little room to doubt the importance of freedom in his critical system. At one point he called it the “keystone [Schlußstein] of the whole structure of a system of pure reason, even of speculative reason” (KpV 5:4). Kant was also unambiguous in his claim that the concept of freedom connects *analytically* to the moral law—what Henry Allison has dubbed the “reciprocity thesis.”¹ The basic thought common to the *Groundwork* and the second *Critique* is that a will that determines itself to action by the mere lawgiving form of its maxims is, by definition, independent from sensible affection; independent, that is, “in the transcendental sense” (KpV 5:29).²

As I mentioned in the last chapter, there is a good deal of controversy over how Kant’s project of moral justification developed from *Groundwork* III to the second *Critique*. Most commentators agree that, in the earlier work, Kant was still attracted to what Karl Ameriks has called a “direct argument.”³ On this reading, his basic strategy is to show, first, that rational agency involves freedom, and second, that we have non-moral grounds to attribute rational agency to ourselves: namely, on the basis of the spontaneity we display in our theoretical capacities. Looking over the literary remains from the “silent decade,” we can see that Kant toyed with a number of such claims; for example, that transcendental freedom follows simply

¹ See his (1990, ch. 11).

²In the first *Critique*, Kant defines transcendental freedom as a “power of beginning a state,” i.e., of initiating a new series of conditions without itself being conditioned (KrV A445/B473).

³See his (2003, ch. 9)

from the fact that we use the first-person singular pronoun spontaneously, i.e., that we say “I think,” “I judge,” etc., in a way that exhibits our independence from sensible affection: “When I say: I think, I act, etc., then either the word I is applied falsely, or I am free. Were I not free, then I could not say: *I* do it, but rather I would have to say: I feel in me a desire to do it. . . . But when I say: *I* do it, that means spontaneity in the transcendental sense (*in sensu transcendentali*)” (LM 28:269).

However we read these early attempts at a deduction of morality, there is no doubt that by 1788, and apparently to the end of his career, Kant believed the only grounds we have to attribute freedom of will to ourselves come from morality, but that morality itself (and our cognition of its authority) is an irreducible “fact”:

Lest anyone suppose that he finds an *inconsistency* when I now call freedom the condition of the moral law and afterwards, in the treatise [i.e., KpV 5:29], maintain that the moral law is the condition under which we can first *become aware* of freedom, I want only to remark that whereas freedom is indeed the *ratio essendi* of the moral law, the moral law is the *ratio cognoscendi* of freedom. For, had not the moral law already been distinctly thought in our reason, we should never consider ourselves justified in *assuming* such a thing as freedom (even though it is not self-contradictory). (KpV 5:4n)

And later, in the *Religion*:

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 over to it. Yet this law is the only law that makes us conscious of the independence of our power of choice [i.e., freedom] from determination by other incentives. (Rel 6:26n)

It seems, then, that Kant’s project of moral justification underwent a “great reversal”⁴ between 1785 and 1788—that he gave up trying to establish the validity of the moral law from non-moral premises, and instead declared the moral law a “fact of reason” whose authority we cognize in an immediate and underived manner. Many have since accused Kant of abandoning philosophical argumentation all-together, and have considered the “fact of reason” to be what Hegel called the last “undigested lump” of revelation in the stomach of reason; a final but perhaps inevitable appeal to *dogmatism* that some believe Kant would come to regret late in life.⁵

This accusation raises more questions that for lack of space I cannot go into. As I’ve said, I believe the current debate over the direction of

⁴This well-known expression is Ameriks’s, from his (2003, ch.6).

⁵See Ameriks (2003, p.184).

Kant's argument obscures a more fundamental development in Kant's ethical thought, one concerning his theory of moral sensibility. While I do not want to get caught up in the intricacies of this debate, I think it is important to understand that Kant's argument in *Groundwork* III is mostly consistent with the second *Critique*. This will also support my claim that Kant's theory of moral sensibility provides a final shift of frameworks missing in 1785.

In this chapter, then, I want to keep my focus on *Groundwork* III. In the first part I want to show that Kant's suspicion of a "hidden circle" arises because he anticipates the reader will attribute rational agency to himself without first venturing on the difficult path of a "critique of pure practical reason." I will also argue that this "critique" involves a distinction between two ways of thinking about our agency, a distinction from "two worlds," and that Kant maintains we only have reason to think of ourselves as practical reasoners when we actively *adopt* the form of universal law (a form given to us in the idea of an intelligible world). Later in this chapter, I will explain why Kant's "two world" distinction does not refer to two ontological "realms," but rather to two "models" or "archetypes" of the will. This will, I hope, shed light on Kant's answer to the Question of Identity.

Not many interpreters have taken up this reading of Kant's development. Those who do often point out that it would have been odd for Kant to change his strategy of argument so radically within the space of just two or three years. After all, the second *Critique* was published in 1788, and Kant must have started working on it soon after *Groundwork* III. For those who remain unimpressed by this timeline, I would like to add that it would have been even more perplexing for Kant to change his strategy within the space of the *Groundwork* itself. For, in Section II, at the beginning, Kant says quite clearly that reason "first becomes aware that it can of itself also be practical . . . [through] the pure thought of duty and in general of the moral law" (G 4:410). This may not be decisive evidence to suggest Kant was arguing from morality to freedom in *Groundwork* III, but it should at least give us pause.

3.2 Acting Under the Idea of Freedom

The first place where Kant seems to propose a "direct argument" is on page 447 of the Academy Edition, where he says that every rational being with will must act "under the idea of freedom." As I read it, Kant is discussing a merely conceptual relation between rational agency and the idea of freedom, and nothing he says on page 447 is meant to supply an argument for *our* status as rational agents. That's why, in the next paragraph, he goes on to write that he has yet to prove freedom "is something real in ourselves and in human nature"; and that he has only shown "we must presuppose

it [freedom] if we want to think of a being as rational and endowed with consciousness of his causality with respect to actions, that is, with a will” (G 4:449).

Now, if Kant believed anything he had previously said shows why we must think of ourselves in this way, under the idea of freedom, I doubt he would have asked why beings like us, “who are also affected by sensibility,” must renounce their interest in happiness in favour of the standard of autonomy. Nor would he have said that if someone pressured us with the question of why we should adopt the latter standard “we could give him no satisfactory answer” (G 4:450). It is conceivable at this stage of the argument (G 4:449-450) that we possess genuine theoretical spontaneity, even that we are practically free to “step back” from our desires, but that we still lack the kind of motivational independence required for moral agency. For all we know, we might be rational but *heteronomous* beings—that is, beings who depend on sensible influences for making choices. In Kant’s view, then, any argument for morality will beg the question if it fails to establish our status as practical reasoners.

The problem is that, for Kant, *agency* does not follow from the concept of a rational being. We do not have grounds to think we possess a will just because we can engage in deliberative capacities normally associated with free agency, such as our capacity to think about the “sum” total satisfaction of our desires. As Kant puts it in a later text:

From the fact that a being has reason it does not at all follow that, simply by virtue of representing its maxims as suited for 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. 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 absolute imperative of the moral law.⁶ (Rel 6:26n)⁷

⁶“Denn es folgt daraus, daß ein Wesen Vernunft hat, gar nicht, daß diese ein Vermögen enthalte, die Willkür unbedingt durch die bloße Vorstellung der Qualification ihrer Maximen zur allgemeinen Gesetzgebung zu bestimmen und also für sich selbst praktisch zu sein: wenigstens so viel wir einsehen können. Das allvernünftigste Weltwesen könnte doch immer gewisser Triebfedern, die ihm von Objecten der Neigung herkommen, bedürfen, um seine Willkür zu bestimmen; hiezu aber die vernünftigste überlegung, sowohl was die größte Summe der Triebfedern, als auch die Mittel, den dadurch bestimmten Zweck zu erreichen, betrifft, anwenden: ohne auch nur die Möglichkeit von so etwas, als das moralische, schlechthin gebietende Gesetz ist, welches sich als selbst und zwar höchste Triebfeder ankündigt, zu ahnen.”

⁷A similar possibility comes up in Kant’s discussion of the “favoured creature” from

The point is not that there may be such a being in the world, a being with the power of rational reflection, but lacking the power of self-determination. What Kant is concerned with is only the *possibility* of such a being—and more crucially, with the possibility that *we* may be like that. This is just another way of putting the claim that our assumption of free rational agency is only secured on the basis of our cognition of the moral law (as I quoted Kant above, “had not the moral law already been distinctly thought in our reason, we should never consider ourselves justified in *assuming* such a thing as freedom” (KpV 5:4n)). There is no form of practical deliberation, instrumental or otherwise, that would provide us with the “*ratio cognoscendi*” of freedom. The example of a rational but heteronomous being, therefore, is just a way of personifying the logical gap between rationality and agency.⁸

With this in mind, we can understand why Kant’s opening claims about rational agency in *Groundwork* III are merely analytic. It remains to be seen whether we have reason to think of ourselves in this way. As Kant says: “every being that cannot act except under the Idea of freedom is because of that—in a practical respect—really free” (G 4:448; modified).⁹ On first glance, it seems we can draw a substantial conclusion from this, given the material Kant has presented so far. If we must represent a rational agent under the idea of freedom, and if a free will and a will under the moral law are “reciprocal concepts,” then we can say a rational agent is *bound* to the moral law.¹⁰ Notice, though, that Kant is speaking about a rational agent

Groundwork I. There Kant tells us that if a rational being with will were designed for the purpose of attaining happiness, nature would have let *instinct* govern its will, and it would have assigned reason a passive role: namely, to self-consciously admire its pursuit of inclinations, “to contemplate the fortunate constitution of its nature” (G 4:395). For such a creature, Kant says, “nature would have taken care that reason should not break forth into *practical use*” (G 4:395).

⁸This should be enough to show why Kant is not a constitutivist, contra Burge (1998), Velleman (1996b; 1997; 1999), and Korsgaard (1996b). While the standards of pure practical reason may be “constitutive” of rational agency and underlie non-moral forms of practical deliberation, Kant is explicitly denying that we can derive those standards *outside* of our immediate cognition of the moral law. For Kant, the logical space between rational reflection and moral agency allows us to conceive of a being that *lacks* a commitment to autonomy or the moral law; it allows us to conceive of a “martian” agent (Lavin’s (2004)), or a “shmagent” (Enoch 2006)). Somewhat ironically, then, I find much to recommend in Lavin’s and Enoch’s arguments against “Kantian constitutivism,” arguments I think we find prefigured in Kant himself.

⁹“Ich sage nun: Ein jedes Wesen, das nicht anders als *unter der Idee der Freiheit* handeln kann, ist eben darum in praktischer Rücksicht wirklich frei.”

¹⁰It is tempting to hear in Kant’s qualification, “in a practical respect,” a psychological claim about the “sense of freedom” we unavoidably have, or a pragmatic assumption that we must act “as if” we were free. But “in a practical respect” refers to the particular function Kant has assigned to practical *Vernunft*: namely, the function of deriving actions from objective laws. Because *Vernunft* is higher than sensibility and understanding (we know this from the first *Critique*, but Kant repeats it on page 452), we can say that motivational

in the third person. We must presuppose freedom, he says, “if we want to think [*denken*] of a being as rational” (G 4:449). But what about beings like us? Do *we* have the capacity to take an interest in morality? Do we have the capacity to determine ourselves independently of sensible inclination? Kant raises the question in the following paragraph—and for the first time in *Groundwork* III he speaks in *first-person singular*:¹¹

But why, then, ought I to subject myself to this principle and do so simply as a rational being, thus also subjecting to it all other beings endowed with reason? I am willing to admit that no interest *impels* me to do so, for that would not give a categorical imperative; but I must still necessarily *take* an interest in it and have insight into how this comes about. (G 4:449)¹²

What is Kant requesting here? In a footnote, he defines an interest as a “cause that determines the will” (G 4:460n).¹³ He also distinguishes between “pure” and “sense-based” interests, saying that our interest is pure only when the mere form of lawfulness suffices to move us to action (cf., G 4:413n). Understood in this way, we can see that Kant is requesting insight into how we can determine ourselves by the moral law. The “insight” is what he will later call the “*ratio cognoscendi*” of freedom, i.e., the *reason for thinking* of ourselves as beings possessing spontaneity of will. This reason would explain how we *take* an interest in the principle of autonomy—that is, *determine ourselves* by the principle—rather than feel *compelled* to act in a certain way.

The argument so far has moved from the concept of a rational being with *Wille*, and Kant has shown that we must think of such a being under the idea of freedom. But he has yet to say why *we* are justified in thinking of ourselves in this way. The problem is that the concept of a human

independence is part of the concept of practical reason. The maximum of this concept is thus the “Idea,” in Kant’s technical sense, of complete motivational independence. This is just what Kant means by transcendental freedom.

¹¹Zinkin (2006) observes that the first place in the *Groundwork* where Kant speaks in the first person occurs in Section I, on page 400 of the Academy Edition. What’s noteworthy about this section is that Kant shifts to the first person when he introduces the concept of “respect” (*Achtung*). As Zinkin writes: “Here, Kant explains that the feeling of respect is the subjective determination of my will by the moral law. It is what motivates *me* as the particular subject that I am, to follow the moral law. In this way, the moral law, which is the law of reason in general, also becomes the law of my own subjective will” (p. 35).

¹²“Warum aber soll ich mich denn diesem Princip unterwerfen und zwar als vernünftiges Wesen überhaupt, mithin auch dadurch alle andere mit Vernunft begabte Wesen? Ich will einräumen, daß mich hiezu kein Interesse *treibt*, denn daswürde keinen kategorischen Imperativ geben; aber ich muß doch hierannothenwendig ein Interesse *nehmen* und einsehen, wie das zugeht.”

¹³Kant seems to repeat this in the third *Critique*, that “to will something and to have satisfaction in its existence, i.e., to take an interest in it, are identical” (KU 5:209).

will contains, in addition to practical reason, practical sensibility. Human agents are rational but also affected by representations of pleasure. So the skeptical thought arises here that maybe we lack the capacity to determine ourselves to action by objective laws—maybe our will is, rather than practical reason, a mere faculty of desire.¹⁴ Of course, if we had insight into how we can act on principles of reason, then we would have grounds to think of ourselves under the idea of freedom, and morality would be valid for us. But *where* is this reason to be found?

3.3 The Standpoint of Theoretical Reason

I should pause here. One might complain I am making things more difficult than they really are. Kant appears to argue that we do have supporting evidence of our rational agency, and this evidence comes from the spontaneity we exercise in our theoretical activities. He says:

Reason must regard itself as the author of its own principles independently of alien influences. It follows (*folglich*) that reason, as practical reason, or as the will of a rational being, must regard itself as free. (G 4:448)¹⁵

How are we to understand this? Is Kant appealing to our theoretical capacity to support his claim that we must act under the idea of freedom? Some readers sympathetic to Kant believe it does. Recently Allen Wood (1999) has suggested that the evidence we have of our capacity to reason by norms of logic secures our status as free agents. The thought is that when we guide our reasoning by such norms, we presuppose our capacity to make inferences and draw conclusions about what to believe in the absence of non-rational influences. According to Wood's reading, we must think, or judge, or reason under the idea of freedom, and this is evidence of our capacity to act and make choices by a principle of rational choice. Because "we cannot intelligibly doubt that we have such a capacity in one case [the theoretical case], we have no good ground for doubting that we have it in the other [practical case]" (p. 176).

The following reconstruction of Kant's argument now seems available. If every rational agent must act under the idea of freedom, and if freedom and morality are reciprocal concepts, then every rational agent is bound to

¹⁴ For the sake of convenience, I will mostly follow the *Groundwork* distinction where "practical reason" is equivalent to "higher faculty of desire" and "pure practical reason," and "faculty of desire" is equivalent to "lower faculty of desire" and "empirical practical reason."

¹⁵ "Sie muß sich selbst als Urheberin ihrer Principien ansehen Unabhängig von fremden Einflüssen, folglich muß sie als praktische Vernunft, oder als Wille eines vernünftigen Wesens von ihr selbst als frei angesehen werden."

morality. It appears that we are rational agents because we possess spontaneity in guiding our inferences by logical norms, and this removes any doubt that we also possess spontaneity in forming our maxims. Therefore, one could conclude, we are bound to the moral law. The problem, though, is that nothing about the spontaneity we exercise in theoretical reason justifies thinking of ourselves under the idea of freedom. For one thing, nothing in the activity of reasoning by logical norms requires our motivational independence in the practical domain. Even if I can successfully guide by inferences by *modus ponens*, it may not be the case that I can form my maxims out of sole consideration for their universal validity. I may still lack practical reason.

Unfortunately, Kant obscures the point in saying it “follows” (*folglich*) that a rational agent must regard itself as free, as this suggests we must act under the idea of freedom because we form our theoretical judgments spontaneously. The misleading implication here is that we can establish our status as self-determining beings by appealing to our capacity to judge. However, we must bear in mind that when Kant says reason must regard itself as the author of its principles, he is making a claim that follows from the concept of *Vernunft*, and this is easily lost in translation. By *Vernunft* Kant is talking about a capacity independent from sensibility. This means *Vernunft* is inseparable from a general concept of autonomy. The use of “follows” could then be understood—as I want to suggest—as a claim of comparison, not of inference. In its theoretical capacity, *Vernunft* is autonomous because it can think Ideas independently from the sensible world. Similarly, *Vernunft* is autonomous in its practical capacity because it can determine the will to action independently from inclination. This follows from the concept. But we have yet to see if the concept is real for beings like us.¹⁶

3.4 Happiness and Rational Heteronomy

If I am right, and Kant’s appeal to our capacity to judge is meant as a comparison, not a supporting inference, then it turns out we still lack grounds to think of ourselves as rational agents. The threat of skepticism arises here because, as sensibly affected beings, we may only act out of a concern for our happiness. For Kant, interestingly, happiness is also an “idea” in his technical sense. As the concept of a “sum,” happiness presents us with a

¹⁶ I believe this comes out more clearly in his “Review of Schulz,” written two years before the publication of the *Groundwork*. Kant says that although Schulz’s fatalistic metaphysics leaves no room for freedom of will, Schulz himself “always admits freedom to think, without which there is no reason.” “In the same way,” Kant continues, “he must also assume freedom of the will in acting.” As I am reading it, “in the same way” is a claim of comparison, and it is possible Kant meant the same thing in the passage from *Groundwork* III. A similar suggestion is made in Schönecker (2005).

standard for deliberation. When we act on this standard, then, we are not led from one impulse to the next: rather, we decide what to do by considering how we would act if we were to successfully harmonize all our desires, in a way similar to when we assess the options available to us from standpoint of freedom. But we cannot say we act under the idea of freedom when we pursue our happiness, as the former requires complete independence from non-rational influences. I may be able to exercise control over my desires, at least to resist acting on them without further evaluation; but that does not show how I can act on principles of reasons.

So it turns out that the concept of “the most rational” but heteronomous agent is consistent with Kant’s definition of happiness. And this is where the threat of agency skepticism gets its foothold. While such an agent would not be determined to act on its representations of pleasure, it would still be dependent on its sensible faculty, and so it would lack the kind “transcendental” freedom Kant thinks is necessary for moral agency. As I said, the issue here is of course not whether there is such an agent in the world, but only about the possibility that we might be like that—undetermined but still in need of incentives coming to us from our inclinations. As long as this possibility hangs over us, we cannot secure our status as the kind of beings who must act under the idea of freedom.

Before moving on, I want to point out that Kant hints at another definition of happiness in *Groundwork* III, one that relates specifically to moral agency. He says:

We do indeed find that we can take an interest in a personal characteristic that brings with it no interest at all in a condition, if only the former makes us fit to participate in the latter in case reason were to effect the distribution, that is, that mere worthiness to be happy, even without the motive of participating in this happiness, can interest us of itself. (G 4:450)¹⁷

Kant introduces this line of thinking to illustrate a hopelessly circular strategy for establishing our rational status. One might be tempted, he thinks, to argue that we can take an interest in morality because it makes us fit to participate in happiness, and this shows we can determine ourselves by principles of reason (e.g., the equivalence of happiness according to merit), which is independent from sensible inclination. But this line of thinking fails to explain *moral* interest, because such an evaluation is the result of the overriding authority we have already attributed to the moral law “when by

¹⁷“Zwar finden wir wohl, daß wir an einer persönlichen Beschaffenheit ein *Interesse* nehmen können, die gar kein *Interesse* des Zustandes bei sich führt, wenn jene uns nur fähig macht, des letzteren theilhaftig zu werden, im Falle die *Vernunft* die Austheilung desselben bewirken sollte, d. i. daß die bloß e Würdigkeit, glücklich zu sein, auch ohne den Bewegungsgrund, dieser Glückseligkeit theilhaftig zu werden, für sich interessiren könne.”

the idea of freedom we detach ourselves from all empirical interest” (ibid). And so the argument presupposes the very capacity it is trying to prove. The problem, as Kant puts it, is that “we cannot yet see [*nicht einsehen*], in this way, that we ought to detach ourselves from such interest, that is, to regard ourselves as free in acting and so hold ourselves yet subject to certain laws” (ibid). We cannot see why we should think of ourselves under the idea of freedom, and so we cannot see on what grounds morality presents us with an obligation.

If Kant is saying only an autonomous agent could evaluate its fitness to receive happiness (and this evaluation depends on the principle of morality), then only an agent committed to that principle could form a conception of its happiness as a rational good. Yet, this is just another way of saying that the idea of happiness does not supply the reason (or “insight”) for thinking of ourselves as autonomous agents. There are two dead-ends here. On the one hand, if we define happiness in a broad sense, as a normatively vague idea of an agent’s “sum-total” well-being, then we can imagine a heteronomous being acting under this idea. On the other hand, if we define happiness in a narrow sense, as the agent’s evaluation of desires according to their agreement with morality, then we are presupposing an agent *who is already autonomous*. Either way, we lack conceptual resources to explain why we should think of ourselves as having the capacity of self-determination in the first place.¹⁸

This brings us back to the skeptical possibility I outlined above. The problem Kant is gesturing at is that we might only ever form our maxims on the basis of desires, which are “foreign” to us, so that our reason would not be practical in the full (Kant would say “pure”) sense. All that reflective agency shows us is that we are not determined by inclinations, as an *arbitrium brutum* or nonagent animal is. It does not show that we have the capacity to act on principle, that we have a rational will in Kant’s technical sense of *Wille*. Our ability to step back gives us independence, but for all we know this independence might be relative: our spontaneity may itself derive from a hidden cause, which would make our acts of volition no different than the machine-like spontaneity of a watch or a turnspit (two images Kant uses with reference to Leibniz’s “spiritual machine”). Understood in this way, the question of why we should give priority to the standard of autonomy is simply a different way of asking whether we possess a

¹⁸I am not intending to settle the question of whether a heteronomous agent is subject to hypothetical imperatives. Engstrom (1993) has suggested this may not be the case, and that only an agent committed to morality could be subject to imperatives stemming from its conception of happiness. Here I am simply following what Kant says: that a being could submit its desires to rational scrutiny without having the least sense of the unconditional imperative of morality.

will governed by that standard, a rational will. And behind this question is the thought that we might be, not agents, but automata—“thinking things” not “persons.”¹⁹

3.5 The “Hidden Circle”

We can now begin to appreciate why Kant thinks arguing directly from freedom to autonomy is question-begging. Remember, we cannot derive an obligation to the principle of autonomy from the concept of freedom because our status remains uncertain at this point. As we have seen, neither reflective distance nor theoretical spontaneity legitimates our claim to rational agency, because even a heteronomous being could step back and reflect on its desires in forming its maxims (and, presumably, it could also guide its inferences by *modus ponens*). This being, as we are conceiving it, would still be dependent on sensibility for making its choices, and so it would lack real spontaneity from the standpoint of deliberation. It would lack practical reason.

Kant formulates the problem in terms of a “hidden circle.”

It must be freely admitted that a kind of circle comes to light here from which, as it seems, there is no way to escape. We take ourselves as free in the order of efficient causes in order to think ourselves under moral laws in the order of ends; and we afterwards think ourselves as subject to these laws because we have ascribed to ourselves freedom of will. (G 4:450)²⁰

This passage in fact summarizes what we have been discussing so far.²¹ Kant is anticipating the following set of moves. We take ourselves as free because we have already attached a value to acting on the moral law, and then we

¹⁹Sellars (1970) provides one of the most acute discussions of this problem in Kant’s theoretical and practical philosophy: “That practical reason is autonomous means that a choice is possible in which practical reason itself affirms the antecedent. A rational being which had no inclinations (desires and aversions) would always act in accordance with the moral law, and hence in the way of autonomy. A being which has inclinations, but not the possibility of acting for the sake of principle, would always act in the way of heteronomy. It would be, an ‘it (the thing) which thinks’” (p. 30). However, I think we can refine Sellars’s point. Even if we are able to reject skepticism of absolute spontaneity in theoretical cognition (as Pippen (1987) argues, I think convincingly), the problem of skepticism about rational agency remains in the practical domain.

²⁰“Es zeigt sich hier, man muß es frei gestehen, eine Art von Cirkel, aus dem, wie es scheint, nicht heraus zu kommen ist. Wir nehmen dem, wie es scheint, nicht heraus zu kommen ist. Wir nehmen uns in der Ordnung der wirkenden Ursachen als frei an, um uns in der Ordnung der Zwecke unter sittlichen Gesetzen zu denken, und wir denken uns nachher als diesen Gesetzen unterworfen, weil wir uns die Freiheit des Willens beigelegt haben.”

²¹For a different reading, see Sussman (2008).

claim we are subject to the moral law—as a binding norm for us—*because* we are free. But we have yet to establish whether we are the kind of beings for whom morality would be binding, i.e., beings who must act under the idea of freedom. The problem is that we have assumed our status as rational beings with wills, the kind of beings who must act under the idea of freedom. So the origin of the circle refers to an illicit assumption of agency Kant fears the reader will make.

Kant thinks the reader will be tempted to do this because he has assumed the importance of morality. The problem is not with the starting point, morality, but with the assumption. Kant will go on to argue that only our practical cognition of morality (presented to us in the idea of a “world of understanding”) gives us reason to think of ourselves as agents. What Kant is worried about, then, is that we will accept the value of acting on the law without providing an account of how we attain insight into it, an account that will require a critical distinction between “two worlds.” In other words, we will argue in a circle if we attribute freedom of will to ourselves because of the importance we have uncritically attached to morality, only to turn around and say that we are bound to morality because we are free. As Kant summarizes it, after the two-world distinction:

The suspicion we raised above is now removed, the suspicion that a hidden circle was contained in our inference from freedom to autonomy and from the latter to the moral law—namely that we perhaps took as a ground the idea of freedom only for the sake of the moral law, so that we could afterwards infer the latter in turn from freedom, and that we were thus unable to furnish any ground at all for the moral law. (G 4:453)²²

Because we didn’t establish our status as rational agents in the first place, this conclusion moves in a circle. But who would argue so carelessly? I take it Kant’s warning is directed to the reader, the “well-meaning soul” willing to look past the question of why we should think of ourselves as agents (G 4:453). Remember, in *Groundwork* III Kant is not speaking directly to a skeptic as someone who denies we possess practical reason. He is speaking to “anyone” who does not think morality is an illusion and who has, with Kant, made the transition from popular philosophy to a metaphysics of morals. So in the opening pages of *Groundwork* III, *before* Kant has even begun his deduction, he is trying to show why we cannot rest content here,

²²“Nun ist der Verdacht, den wir oben rege machten, gehoben, als wäre ein geheimer Cirkel in unserem Schlusse aus der Freiheit auf die Autonomie und aus dieser aufs sittliche Gesetz enthalten, daß wir nämlich vielleicht die Idee der Freiheit nur um des sittlichen Gesetzes Willen zum Grunde legten, um dieses nachher aus der Freiheit wiederum zu schließen, mithin von jenem gar keinen Grund angeben könnten”

why we need a special “synthetic” method of justification, and why we need to make the difficult transition to a “critique of pure practical reason.”²³

What, then, does this “critique” involve? I will explain this in more detail shortly, but here is an outline we can start with. When we think of ourselves as members in the world of sense, we can only explain our actions in terms of what we find subjectively agreeable. From the standpoint of this world our will is nothing more than a faculty of desire. For Kant, this is why we require a different way of looking at ourselves, a “different standpoint” (*anderen Standpunkt*)²⁴ in a world unaffected by sensibility. This brings us to the world of understanding.²⁵ At first it is just the negative or limiting concept of a non-sensible world. (The world of understanding is on par with the negative definition of freedom as “independence from external causes.”) By thinking our way into this world, however, Kant says we are conscious of a different kind of nature, a nature “under rational laws.” Somehow this stage of the argument (pages 451-453) is meant to secure *our* status as beings with practical reason. But it remains unclear how the argument is supposed to work.

This will be my focus for the remainder of this chapter. To begin, I want to show why, for Kant, membership in the world of understanding secures our possession of practical reason. Recent commentators have argued

²³Kant’s warning of a hidden circle is perhaps less strange when we look back to the opening pages of the Deduction of the Categories from the first *Critique*. He says: “the reader must be convinced of the unavoidable necessity of such a transcendental deduction before he has taken a single step in the field of pure reason; for he would otherwise proceed blindly, and after much wandering around would still have to return to the ignorance from which he had begun. But he must also clearly understand from the outset its inevitable difficulty, so that he will not complain of obscurity where the subject-matter itself is deeply veiled or become annoyed too soon over the removal of hindrances” (KrV A88/B121).

²⁴Kant says this “different standpoint” is available to us “when by means of freedom we think of ourselves as causes efficient a priori” (G 4:450). This is potentially confusing. We cannot use a positive concept of freedom as a springboard for our argument, because we have yet to see whether we really are self-determining beings. The sense of freedom Kant is referring to here is the “negative definition” he offered in the opening paragraphs, i.e., freedom from natural determining causes. As we will see below, only our practical cognition of the world of understanding supplies the “positive definition” of freedom Kant says we need to escape the hidden circle.

²⁵Kant’s use of “world” (*Welt*) is the subject of much controversy. Contemporary philosophers sympathetic to Kant’s ethics often shy away from “two world” readings of *Groundwork* III because of its commitment to ontological dualism. My reason for rejecting it is that—as far as I can tell—it leaves Kant without an argument for why we should think of ourselves as rational agents. The metaphysics of transcendental idealism can, at best, secure the logical possibility of freedom, along with our status as rational beings with self-consciousness. But I don’t see how it can secure our status as rational beings with *Wille*, which would make freedom a real possibility for us. On my reading, the world of understanding only provides us with the idea of universal lawfulness, an idea which Kant says we must adopt as the determining principle of our own will. I will develop this point below.

that the metaphysics of transcendental idealism provides the required link between the two, but I will argue this creates unnecessary difficulties for Kant's position. The clue to a successful reconstruction of Kant's argument is available, I think, in his later remark that the world of understanding is simply a "model" or "archetype" of the will.²⁶ One advantage of my reading is that it avoids the charge, raised by Henry Allison (1990) and Paul Guyer (2007), that Kant is guilty of equivocating between weak and strong notions of rational agency.

3.6 "Raw" Transcendental Idealism

One of the most striking features of *Groundwork* III is the assumption that our ordinary understanding of the world contains the very distinction Kant had struggled to justify on philosophical grounds in the first *Critique*. Within the space of a few paragraphs, Kant sets up the distinction between the world of noumena and the world of appearances, and he purports to draw this distinction from sources available to common reason. It is, he says, a distinction common reason may stumble upon, not by "subtle reflection," but "by an obscure discrimination of judgment which it calls feeling" (G 4:451). Feeling lets us form a "raw" (*Rohe*) separation between phenomena and noumena. As Kant puts it, common reason feels itself "passive" to appearances that come to it involuntarily, which lead it to assume something behind such appearances, some "active" essence which it does not cognize or see. The same "raw distinction" must hold, he continues, when a person of common understanding regards his own existence. That is, beyond how he appears to himself, in his own consciousness, such a person must assume the existence of his "ego" in its essential and active power. From this we can distinguish two "standpoints" within the self. When I regard myself as "passive" to sensations and impulses I think of myself as having membership in a "world of sense." Yet when I regard myself as active in my own nature, I think of myself as having membership in an "intellectual world" (G 4:451).

Kant introduces these distinctions starting from the passive side (the "thing as it appears" or the "world of sense") in order to establish the assumption of a hidden, active side (the "thing in itself" or the "world of understanding"). Now he changes directions and asserts that the hidden, active side within us is confirmed by our capacity of reason (G 4:452). Reason displays our "self-activity" (*Selbstthätigkeit*), Kant says, because only reason

²⁶In the second *Critique*, Kant refers to the world of understanding as "supersensible nature." While it has eluded generations of commentators, he tells us "the most ordinary attention to oneself confirms that this idea [of supersensible nature] is really, as it were, the model [*Muster*] for the determination of our will" (KpV 5:43).

can afford us to think of pure “Ideas” which bear no relation to the sensible world.

Now, a human being really finds in himself a capacity by which he distinguishes from all other things, even from himself insofar as he is affected by objects, and that is *reason*. This as pure self-activity, is raised even above the *understanding* ... [It] shows in what we call “ideas” a spontaneity so pure that it thereby goes far beyond anything that sensibility can ever afford it. (G 4:452)²⁷

Kant is no longer speaking of the concept of a rational being in general. He is speaking about beings like us, about *Menschen*. Beings like us must count themselves as members of the world of understanding, a world unaffected by sensibility, because we find a capacity within ourselves that bears no relation to the sensible. It is the capacity to think cosmological concepts—freedom, the soul, or God. Only with reason are we capable of this. The spontaneity of the understanding is limited. For example, I must order my intuition of a chair under the concept of quantity, but my ordering—my activity in bringing unity to the sensible data—still depends on the sensible data. Only with reason and my capacity to think Ideas is my spontaneity “pure” (*reine Spontaneität*), and so removed from anything sensible.

What conclusion can we draw from this? Kant’s wording suggests that we are now entitled to think of ourselves as rational *agents*. As I argued in section 3.4, it is not clear at this stage in the discussion whether we are agents, whether we possess *Wille*. The threat of a circle arose because Kant anticipated the reader would ascribe *Wille* to himself only because of the importance he already attaches to the moral law. Now Kant seems to be saying that we find non-moral evidence of our agency in our capacity to think Ideas, a capacity “a human being really finds in himself.” So it seems the following “direct argument” is at our disposal. If the capacity to think Ideas reveals a *pure* spontaneity in us, we cannot reasonably doubt our independence from sensibility in the theoretical domain. And if the capacity of *Wille* requires nothing more than the spontaneity of self-determination, i.e., the capacity to act on principle, then we cannot reasonably doubt our independence from sensibility in the practical domain. Somehow the possession of reason secures our membership in the world of understanding, and this membership justifies our claim to rational agency. It seems, then, that we have the “one resource” (*eine Auskunft*)²⁸ Kant said would provide

²⁷“Nun findet der Mensch in sich wirklich ein Vermögen, dadurch er sich von allen andern Dingen, ja von sich selbst, so fern er durch Gegenstände afficirt wird, unterscheidet, und das ist die *Vernunft*... da hingegen die Vernunft unter dem Namen der Ideen eine so reine Spontaneität zeigt, daß sie dadurch weit über alles, was ihr Sinnlichkeit nur liefern kann.”

²⁸The German “*eine Auskunft*” has two senses: (1) “information” and (2) “route.” Gre-

our escape from the hidden circle. From the non-moral capacity to think cosmological concepts, we display a spontaneity so pure that we are justified in thinking of ourselves as members of the world of understanding, and as members of this world we must think of ourselves acting under the idea of freedom. The rest of the argument falls into place, as Kant writes: “With the idea of freedom the concept of autonomy is now inseparably combined, and with the concept of autonomy the universal principle of morality” (G 4:452).

As natural as this reconstruction of Kant’s argument may be, there are good reasons to reject it. First, it would be odd if the self-activity of reason on page 452 were meant to carry argumentative weight, since Kant had already introduced the capacity of reason on page 448. He did not explicitly mention the status of reason as the highest capacity of the human mind, distinct from understanding in its capacity to think the unconditioned; but the basic point is the same. As Kant put it, “Reason must regard itself as the author of its principles independently of alien influences” (G 4:448), which is just to say that reason is spontaneous. In section 3.3 I argued that Kant’s further statement about practical reason does not add a substantial conclusion (namely, that *we* are rational agents because theoretical reason must regard itself as free from sensibility), but merely adds a point of comparison. It follows from the definition of *Vernunft* that we must represent a rational being under the idea of its unlimited independence from sensibility, i.e., under the idea of freedom. The threat of a circle arose because Kant anticipated the reader would attribute rational agency to himself without a critical argument. If Kant is just repeating the same point about the spontaneity of reason on page 452, it is not clear how we are any closer to a solution.

One obvious difference is that on page 452 Kant links possession of reason to membership in the world of understanding, and the latter secures our status as “intelligences” (G 4:452). This is what we were missing on page 448. As Henry Allison notes (1990, p.222), we find a parallel move in the first *Critique*, where Kant claims that we know ourselves through “pure apperception, and indeed in actions and inner determinations which cannot be accounted at all among impressions of sense” (KrV A547/575). We thus

gor and Abbott emphasize the former in their translation, “one resource.” Jens Timmerman (2007, p. 133, note 30) makes a convincing case for why this misses Kant’s eighteenth-century German, and he prefers “one route” (Zweig) and “one way out” (Wood). I have no particular commitment here, but I think both represent what Kant is up to. As I will argue, the world of understanding contains the idea of universal lawfulness, and this provides us with the necessary “resource” to determine the concept of freedom. Since this concept-determination gives us reason to think of ourselves as agents, i.e., as beings possessing *Wille*, the world of understanding is also an “escape route” from the threat of circularity Kant wants us to avoid.

have, in addition to an empirical character under empirical laws, an intelligible character under laws of reason. Kant uses the same line of argument in *Groundwork* III, where he claims that “with regard to what there may be of pure activity in him (what reaches consciousness immediately and not through affection of the senses) he must count himself as belonging to the intellectual world” (G 4:451). A few paragraphs later he concludes: “Because of this a rational being must regard himself as intelligence (hence not from the side of his lower powers) as belonging not to the world of sense but to the world of understanding” (G 4:452).

According to Allison, the argument proceeds in two steps. The first is that pure apperception lets us discover our capacity of reason. Since reason is raised above sensibility and understanding, we are justified in thinking of ourselves as intelligible beings. The first step, then, links possession of reason to membership in an intelligible world. Secondly, and more crucially, Kant has to link our membership in an intelligible world to our possession of practical reason, or *Wille*. And this is what we seem to find him doing: “As a rational being, and thus as a being belonging to the intelligible world, the human being can never think of the causality of his own will otherwise than under the idea of freedom” (G 4:452). Unfortunately, Kant does not say exactly how membership in the world of understanding is supposed to secure our possession of *Wille*. While admitting it is “somewhat speculative,” Allison thinks we can fill in the missing details of Kant’s argument along the following lines:

The essential point is that the connection of this consciousness with an activity that is “merely intelligible” and therefore located in the intelligible world serves to obviate any potential suspicions regarding its illusory character. It does so by undermining the assumption on which such suspicion could alone be based, namely, that our apparent practical rationality is really tropistic, the product of some “hidden mechanism of nature.” Given transcendental idealism, the location of this activity in the intelligible world rules out the latter possibility and therefore any reductive, epiphenomenalistic treatment of the will because it exempts it, *ex hypothesi*, from this mechanism. (1990, p. 224)

Now, it is true that Kant’s appeal to the world of understanding is meant to provide a different way of thinking about our agency, one that does not reduce human action to the causal mechanism of nature. The difficulty we mentioned above is that the world of understanding restricts us to the thought of our independence from sensibility, which is Kant’s negative definition of freedom. This falls short of a claim to agency because we still lack insight into our capacity of self-determination. On Allison’s reconstruction, locating the spontaneity of reason in a non-mechanistic world responds to

the threat of agency skepticism by eliminating its underlying assumption: namely, that our spontaneity may be illusory, the result of natural causes we are simply unaware of. But again, it is difficult to see how this shows agency is anything “more” than a logical possibility, which is as far as Kant’s resolution to the Third Antinomy gets us (KrV A558/586).²⁹

3.7 The *Verstandeswelt* as an Archetype

If this really is Kant’s strategy of argument, then we can accuse him for failing to establish our status as rational agents, and this would put us right back into the circle he wished to preemptively avoid (Allison 1990, p. 228). Before drawing this conclusion, however, we should see if there is another way to reconstruct Kant’s argument that doesn’t leave him in such a hopeless position. Part of the difficulty is that Kant gives almost no indication of how membership in the world of understanding secures our possession of practical reason.³⁰ There are, I think, enough clues to rebuild a coherent line of argument.³¹

Right after Kant says a rational being must regard himself “as intelligence,” he says: “hence he has two standpoints from which he can regard himself and cognize laws for the use of his powers *and consequently for all of his actions*” (G 4:452; my emphasis). How should we read this? On the one hand, it would be odd if Kant were saying the spontaneity we exercise in our capacity to think cosmological concepts allows us to cognize laws for practical reason, i.e., laws of *action*. For what could I possibly gain from the Ideas of freedom, the soul, or God when I am faced with a decision? How could these Ideas give me guidance from the standpoint of deliberation? On the other hand, it would be equally odd if Kant were saying the world of

²⁹“It should be noted that here we have not been trying to establish the **reality** of freedom, as a faculty that contains the causes of the appearances in our world of sense... Further, we have not even tried to prove the **possibility** of freedom; for this would not have succeeded either, because from mere concepts a priori we cannot cognize anything about the possibility of any real ground or any causality.”

³⁰Henrich (1975) writes: “Diese Deduktion [des Freiheitsbewußtseins] unterschiedet somit zwar notwendig die beiden Welten, sie setzt die aber nicht in Beziehung zueinander” (p. 98). In one way this is true: Kant’s argument for our status as rational agents does not require the distinction between two worlds. Later, his argument for the possibility of moral obligation will involve a claim about the relation between the world of understanding and the world of sense. In another way, however, Henrich may be overstating his case, since Kant obviously needs to move past the view of human agency that the standpoint of the world of sense limits us to, i.e., our agency as a mere faculty of desire. Thus Kant needs to distinguish between two worlds to ground his deduction of freedom from the moral law, but Henrich is right to point out that the deduction itself does not depend on the two-world relationship.

³¹My discussion here is indebted to Tenenbaum’s insightful analysis of *Groundwork* III. See his (Forthcoming).

the understanding in itself allows us to cognize practical laws. For Kant is careful to add that we have “no further cognizance” of this world (G 4:451), and that “it is only a negative thought with respect to the world of sense: it gives reason no laws for determining the will” (G 4:458).

So how do we cognize laws for the use of practical reason? The answer is partially contained in the second half of the sentence I cited above. After Kant says a rational being has two standpoints from which he can view himself, he adds: “first, insofar as he belongs to the world of sense, under laws of nature (heteronomy); second, as belonging to the intelligible world, under laws which, being independent of nature, are not empirical but grounded in reason” (G 4:452). What can the latter sort be? Practical laws do not follow from cosmological concepts; and we have no further insight into the mechanism of the world of understanding. So the only laws “grounded in reason” that pertain to action are the laws of morality. What I take Kant to be saying, then, is that only our cognition of morality provides us with laws for the practical use of reason, *and so only morality provides the reason for thinking of ourselves as rational agents.*

We now have the materials to provide an alternative reconstruction of Kant’s argument. Starting from the beginning: We can form a “raw” distinction between the world of understanding and the world of sense just by feeling our passivity to sensible appearances, a feeling that leads us to assume something active, but hidden, behind appearances: noumena. Through pure apperception we are aware of our own activity, revealed to us in our capacity to think Ideas, and this grants us membership in the world of noumena. But the world remains a negative concept for us; it is whatever is “left over” from the conditions of sensibility. We lack insight into its mechanism. The question we’re after, however, does not concern this. We are only after the question, “How should I act?” The worry is that by abstracting from the conditions of sensibility we have cut off the very source that makes human action meaningful: our inclinations. By thinking our way into the world of noumena, but without the capacity to know anything about its laws, it seems we have left ourselves without any ground for determining the will.

But here a new prospect suddenly opens up. Without claiming any kind of special insight into the noumenal world, we can ask ourselves how we would act if we were intelligible beings belonging to it. We can thus employ the idea of a noumenal world as an “archetype” for reason’s practical use. This is the key step in Kant’s argument. The idea of a noumenal world abstracts from the conditions of sensibility, conditions that reduce our will to a faculty of desire. It leaves us with the *mere form of universal lawfulness*, and that form provides us with a model for the will. With the mere form of universal lawfulness, we have insight into our capacity for self-determination,

i.e., our capacity to act independently of “alien” influences. So we can now speak of this capacity in positive terms, as a capacity of self-determination.

Kant makes this point clearly later in *Groundwork* III. The world of understanding, he says,

is only a negative thought with respect to the world of sense: it gives reason no laws for determining the will and is positive only in this single point: that freedom as a negative determination is combined with a (positive) capacity as well, and indeed with a causality of reason that we call a will, a capacity so to act that the principle of actions conforms with the essential constitution of a rational cause, that is, with the condition of the universal validity of a maxim as a law [i.e., the moral law]... The concept of the world of understanding is thus only a *standpoint* that reason sees itself constrained to take outside appearances *in order to think of itself as practical*. (G 4:458; Kant’s emphasis).³²

As I understand it, Kant is saying that the only way we can model our will onto the world of understanding is to think of ourselves forming maxims out of consideration for their universal validity, i.e., out of consideration for the moral law. This is the only way we can cognize a law for reason’s practical use. Contra Allison, the idea of a world unaffected by sensibility does not provide an ontological realm for “locating” the will, such that we can obviate reductive or epiphenomenalistic accounts of agency. Rather, the idea functions as an *archetype for agency*. It enables us to “fill in” the thought of our independence from inclination (freedom in the negative sense) in terms of our capacity to act on principle (freedom in the positive sense). On my reading, then, Kant is not saying we are entitled to call our independence “will” when we *think* the problematic concept of a world of understanding (or when we are conscious of the “pure self-activity” of reason within us). Rather, we are entitled to regard our independence as “will” only when we *adopt* the form of the world of understanding—i.e., the form of universal lawfulness—as a sufficient ground of action. This answers the Question of Identity.

³²“Jenes ist nur ein negativer Gedanke in Ansehung der Sinnenwelt, die der Vernunft in Bestimmung des Willens keine Gesetze giebt, und nur in diesem einzigen Punkte positiv, daß jene Freiheit als negative Bestimmung zugleich mit einem (positiven) Vermögen und sogar mit einer Causalität der Vernunft verbunden sei, welche wir einen Willen nennen, so zu handeln, daß das Princip der Handlungen der wesentlichen Beschaffenheit einer Vernunftursache, d. i. der Bedingung der Allgemeingültigkeit der Maxime als Eines Gesetzes, gemäß sei. Würde sie aber noch ein *Object des Willens*, d. i. Eine Bewegursache, aus der Verstandeswelt herholen, so überschritte sie ihre Grenzen und maße sich an, etwas zu kennen, wovon sie nichts weiß. Der Begriff einer Verstandeswelt ist also nur ein *Standpunkt*, den die Vernunft sich genöthigt sieht, außer den Erscheinungen zu nehmen, *um sich selbst als praktisch zu denken*.”

One advantage of this reconstruction is that it avoids the charge of equivocation raised by Allison (1990) and Guyer (2007).³³ As I discussed above, if Kant's appeal to transcendental idealism is meant to overcome agency skepticism, then his argument is "doomed to failure" and the threat remains intact. Nothing about membership in the world of understanding shows that reason is practical. Yet without this, without practical reason, we lose the condition we need to show why the moral law is a valid principle of choice for us, and so we risk falling back into the circle Kant wanted to avoid. If we cannot establish our status as beings with the spontaneity of *Wille*, i.e., the capacity of self-determination, we cannot say we are the kind of beings for whom the moral law is a binding norm. According to Allison and Guyer, Kant is guilty of sliding from a weak notion of agency as independence from sensible affection, to a strong notion of agency as self-determination. The problem is that membership in the world of understanding, which Kant secures through our capacity to think Ideas, fails to establish our possession of the latter. So we are left without any basis for attributing rational agency to ourselves, and so without any basis for thinking of ourselves under the idea of freedom. As we can now see, the charge of equivocation rests on a misunderstanding of Kant's basic strategy of argument: it assumes that Kant is trying to argue to morality from a non-moral account of freedom. Once we reject this assumption, we can locate the key step of the deduction, not in the *thought* of our membership in the world of understanding, but in our *act* of adopting the form of lawfulness presented to us in the idea of such a world. As I have reconstructed it, then, Kant's point is that we can only secure our possession of practical reason on the basis of our cognition of the moral law, without which it is conceivable that we may be, like "the most rational being of this world," dependent on incentives coming from our inclinations.

³³Allison: "given the identification of will and practical reason the claim that rational beings possess a will can mean either merely that reason is practical or that pure reason is practical. The former suffices to show that we are genuine rational agents rather than automata; but the latter is required to establish our autonomy... Once again, then, it is clear that Kant needs the latter and stronger claim but membership in the *Verstandeswelt* provides, at best, support for the weaker. Consequently, the attempt to establish the necessity of presupposing the kind of freedom that is both necessary and sufficient for morality (transcendental freedom) on the basis of a non-moral premise about our rationality is doomed to failure even when buttressed by transcendental idealism" (p. 228). Guyer: "The problem with the argument is ... that Kant has no real basis for his positive assertion that our real selves are genuinely rational except for his illegitimate inference from the phenomenal rationality that distinguishes us from other things in nature to the genuine rationality of our real selves. By his own insistence on the distinction between appearances and things in themselves, he should not be able to infer from appearances to a similar feature of things in themselves, thus he should not be able to infer from phenomenal rationality to noumenal rationality" (p. 160).

The Moral “Ought”

4.1 Preliminaries

WHY doesn't Kant stop here? It appears he has accomplished the main task of *Groundwork* III. By showing that we are justified in thinking of ourselves under the idea of freedom, he has answered the question of why the moral “ought” is an expression of our “will.” From the standpoint of the intelligible world, we can say the moral law is the law under which we act as self-determining beings (pages 452-453 of the Academy Edition). However, Kant cannot stop here because he has yet to explain how our experience of “constraint” (*Zwang*) or “necessitation” (*Nötigung*) is possible. This is the Question of Obligation.

Let me put the entire argument so far in review. In pages 446-447 Kant outlines a merely analytic connection between the idea of a free will and a will under the moral law. This forms the basis of his reciprocity thesis. He then says we must represent a rational being with a will under the idea of freedom. This connection is also analytic, but it deepens the reciprocity thesis to include rational agents in general. In pages 449-450 Kant warns against drawing any substantial conclusions here. If we ascribe freedom to ourselves because of the importance we already attach to the moral law, we risk begging the question. The only way of avoiding this, Kant argues, is to distinguish between two perspectives on the will, one from the “world of understanding,” the other from the “world of sense.” This is the the “critique” of pure practical reason (pages 451-453). From the world of understanding, we can determine the concept of our independence from sensibility in positive terms. This allows us to discover our will in its pure capacity.

From the world of sense, however, we can only view our will under the principle of heteronomy, because the only laws that apply to us from this perspective come from our impulses and inclinations.

With the critique, we can now overcome the threat of agency skepticism that first lead Kant to voice his suspicion of a hidden circle. The suspicion arose because Kant anticipated the reader might extend his argument from the analytic claim on page 447 that every rational agent must act under the idea of freedom to the synthetic conclusion that we are subject to the moral law. Because we still lack grounds to think of ourselves under the idea of freedom, we cannot say on page 447 that we are bound to the moral law by virtue of the reciprocity thesis. The “one route” that remains to us, Kant says, is to place ourselves in a different order of nature when we think of our independence from sensibility, an order that provides us with a model of rational causality. This completes the first stage of Kant’s deduction.

We are now in a position to understand the second stage. In subsection four (pages 453-455), Kant argues that we are bound to the moral law as human beings because the world of understanding (to which our pure will belongs) contains the “ground” (*Grund*) of the world of sense (to which our sensibly affected will belongs).¹ This is what I am calling the Grounding Thesis. Moreover, because we are conscious of belonging to these worlds “at the same time” (*zugleich*), we can see how our experience of moral obligation is possible. This is the Simultaneity Thesis. My first task in this chapter is to separate these claims and show how they function in the final stage of Kant’s deduction. Next I will work toward identifying and diagnosing the problem with this stage, which points more generally to the failure of *Groundwork* III. As I will argue, Kant is unable to explain how the imperatival character of morality also functions as an incentive for us, i.e., how we are responsive to the moral “ought” at the level of our sensibility. Because the world of sense is only a system of heteronomy in *Groundwork* III, we can only theorize about human sensibility in terms of pathological compulsion. In the end I will suggest this limitation explains why Kant pushes the Question of Motivation to the “outermost limit” of moral philosophy.

4.2 The Intelligible Ground of the World of Sense

It is helpful to keep in mind what Kant is presupposing in subsection four (pages 453-455). We touched on these in the last chapter: (1) that reason is the highest faculty of the human mind; (2) that every rational being is conscious of this faculty in herself through pure apperception; and (3) that practical and theoretical reason are grounded in a single principle. In subsection three Kant defines reason as the faculty that distinguishes between

¹“Belonging” here means our capacity to think of ourselves from different standpoints.

understanding and sensibility and marks out their respective boundaries. He argues that every rational human being is conscious of this faculty in herself through the “I think” accompanying all cognitive activity. From this Kant says we are led to believe that one part of us—the thinking part—is an active “spontaneity,” which we distinguish from the part of us that receives impressions, our passive “sensibility.” Here Kant is more or less repeating what he said in the first *Critique*: we discover our “intelligible character” through the activity of pure apperception, and our “sensible character” through the receptivity of sense (KrV A547/A575). Without going into the details of this claim, I want to highlight its relevance for the argument in *Groundwork* III. This is a claim for the unity of consciousness we have of ourselves as intelligent beings who also belong to the world of sense. Kant hints at this in the Preface when he says that theoretical and practical reason are unified by a single principle and differ only in their application (G 4:391).

All of this is very rough, but it helps us understand Kant’s claim on page 453 that the intelligible world contains the “ground” (*Grund*) of the world of sense, or what I am calling the Grounding Thesis. Kant introduces the thesis in two different ways, which is somewhat confusing. First, he says the intelligible *world* contains the ground of the world of sense:

... because the world of understanding contains the ground of the world of sense and so too of its laws [it] is therefore immediately lawgiving with respect to my will (which belongs wholly to the world of understanding). (G 4:453)²

Then, in the next paragraph, he says my intelligible *will* contains the ground of my sensibly affected will:

... to my will affected by sensible desires there is added the idea of the same will but belonging to the world of understanding—a will pure and practical of itself, which contains the supreme condition, in accordance with reason, of the former will. (G 4:454)³

There is little agreement among commentators how Kant wants us to understand this relationship, between “world” (*Welt*) and “will” (*Wille*).⁴ I

²Weil aber die Verstandeswelt den Grund der Sinnenwelt, mithin auch der Gesetze derselben enthält, also in Ansehung meines Willens (der ganz zur Verstandeswelt gehört) unmittelbar gesetzgebend ist.”

³dadurch daß über meinen durch sinnliche Begierden afficirten Willen noch die Idee ebendesselben, aber zur Verstandeswelt gehörigen reinen, für sich selbst praktischen Willens hinzukommt, welcher die oberste Bedingung des ersteren nach der Vernunft enthält.”

⁴Henrich complains that Kant gives us “no trace of a suggestion” for how to understand the subordination of the sensible to the intelligible world as the subordination of the sensibly affected will to the intelligible will (1975, p. 97). I agree that Kant does not make this relationship clear, but I think he leaves enough clues to fill in the missing details.

want to extend the archetypal or two-model reading from the last chapter and argue for its advantages over the more common ontological reading. In section 3.7, I argued that the world of understanding provides us with a “model” of rational causality, which Kant says frames the will as practical reason. As expected, the claim also holds in the reverse direction. The world of sense provides us with a model of natural causality, which in turn frames the will as a mere faculty of desire. On my reading, then, the two models provide two ways of knowing ourselves as agents. Only from the standpoint of the world of understanding do we know our will as it is “in itself,” i.e., as practical reason, in contrast to what our will “appears” to be from the standpoint of the world of sense, i.e., as a faculty of desire.

Kant is not supporting his argument on realist claims about independent ontological realms, although he helps himself to the “raw” transcendental distinction we discussed earlier. The claim is that we only attain a full understanding of ourselves as agents when we model the will on the rational laws of the world of understanding. That is why we only know the will as practical reason when we transfer ourselves in thought to an order unaffected by sensibility. The idea of such an order, although incomprehensible to theoretical cognition, gives us a *pattern for forming our maxims*. It gives us the form of universal lawfulness, which is left over when we abstract from every condition of the world of sense. So when we speak of a heteronomous will, we are not speaking of a will belonging to a separate ontological realm. We are speaking of *Wille* under the conditions of sensibility. Kant’s point, I think, is that the idea of rational causality underlies and informs every account of the will, even naturalistic accounts that reduce the will to a faculty of desire.⁵ The world of understanding contains the “ground” of the world of sense because it contains the highest principle by which we can think of ourselves as self-efficient causes.⁶ Naturalistic accounts of the will are not wrong, then, but incomplete. They limit us to a view of ourselves as partial causes.

We may think Kant has now explained the possibility of moral obligation. If the moral law is valid for my will as I belong to the world of understanding, and if the world of understanding contains the “ground” of the

⁵Onora O’Neill (1989) emphasizes this point: “Any explanations offered in terms of events and their effects is incomplete because it presupposes an account of the form of certain principles. Putting this in an old-fashioned way we might say that explanations under the heading of efficient causality presuppose explanations under the heading of formal causality. In Kant’s terminology empirical relations presuppose intelligible relations” (p. 68). I am indebted to O’Neill’s discussion in this section. For a similar and equally helpful account, see Piper (2001).

⁶Kant clarifies this when he says the idea of the will belonging to the world of understanding contains the “supreme condition” (*oberste Bedingung*) of the same will belonging to the world of sense (G 4:454).

world of sense, then the moral law is also valid for me as a being affected by sensible inclinations. But there is a problem here. On closer scrutiny, it seems that Kant’s appeal to the metaphysics of transcendental idealism has created an unbridgeable rift in the heart of the moral subject. For one could point out that the primacy of the intelligible world only shows why the moral law is valid *for intelligible beings*. Kant reinforces this worry on page 454, where he writes:

All my actions as only a member of the world of understanding would therefore conform perfectly with the principle of autonomy of the pure will; as only a part of the world of sense they would have to be taken to conform wholly to the natural law of desires and inclinations, hence to the heteronomy of nature. (G 4:453)⁷

How should we read this? If we can only derive the validity of the moral law as a principle of action for intelligible beings—and for ourselves insofar as we consider our will under a model of rational causality—we are no closer to explaining the phenomenon of obligatoriness. We know that the moral law would not be necessitating for perfectly rational or holy wills: there would be no dissimilarity in what they want and what the moral law requires. But if the moral law is only valid for someone who wants to perform moral actions anyway, we have failed to show why beings like us experience a moral “ought.” We have failed to establish the law’s synthetic character.

The dualism of Kant’s “critique” is, I think, only apparent. After the above-cited passage, Kant says we can consider ourselves as members of the world of understanding, and when we do we recognize the moral law is valid for us. But we are not members *only* of this world. We are *also* members of the world of sense, and when we consider ourselves from the standpoint of this world we recognize the presence of our own desires and inclinations. Kant sets up the standpoint of the pure will, then, as a point of contrast to the human standpoint. The experience of constraint peculiar to beings like us is possible, for Kant, only because we can think of ourselves *simultaneously from two frameworks*. This is what I am calling the Simultaneity Thesis.⁸ Kant says that because the world of understanding contains the “ground” of the world of sense, “it follows that I shall cognize myself as intelligence, *though on the other side* as a being belonging to the world of sense, as nevertheless subject to the law of the world of understanding. . . consequently the

⁷“Als bloßen Gliedes der Verstandeswelt würden also alle meine Handlungen dem Princip der Autonomie des reinen Willens vollkommen gemäß sein; als bloßen Stücks der Sinnenwelt würden sie gänzlich dem Naturgesetz der Begierden und Neigungen, mithin der Heteronomie der Natur gemäß genommen werden müssen.”

⁸Kant first states the thesis at the end of subsection three: “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 [*zugleich*] to the world of understanding” (G 4:453).

laws of the world of understanding must be regarded as imperatives for me, and actions in conformity with these as duties” (G 4:454; my emphasis).

What Kant may be saying is that I experience morality in the form of an imperative because I am conscious of myself as an intelligible will (for whom morality is “immediately lawgiving”) at the same time as I am aware of myself as a being affected by sensibility (for whom morality is “necessitating”). Previously, when Kant asked us to transfer ourselves in thought to the world of understanding, he was setting up the first condition of his argument. He was asking us to think of ourselves as pure or holy wills. Now, in pages 454-455 Kant is asking us to do something more difficult, and more interesting—namely, to think about our pure will *under human conditions*. If I were only a member of the world of understanding I would, like a holy will, only do what morality requires. But I am not only a member of the world of understanding. If I were only a member of the world of sense I would, like an animal will, only do what my happiness requires. But I am not only a member of the world of sense. *I belong to both worlds at the same time.*⁹ My experience of constraint is possible, then, because I recognize the validity of the moral law “at the same time” as I am aware that I may resist it because of my inclinations.

This completes the second stage of the deduction. Kant can finally explain why I experience what I *would* do as an intelligible being as what I *should* do as a being affected by non-rational inclinations. Here are the three steps:

1. The world of understanding provides the condition for understanding myself as an agent, and this condition in turn “grounds” every empirical conception of the will available to me from the world of sense.
2. When I view myself from the standpoint of the world of understanding, I recognize the validity of the moral law for my pure will “at the same time” as I consider myself a being affected by sensible impulses and inclinations.
3. Therefore, the same actions I “would” do as a pure will (for whom the moral law is descriptive) appear to me as what I “should” do as a sensibly affected will (for whom the moral law is prescriptive).

By showing the possibility of a moral “ought,” Kant has established the

⁹It is precisely because the two worlds are two different ways of thinking about ourselves as agents that Kant can speak of the simultaneity of moral experience. He makes this fairly clear in saying that “to my will affected by sensibly desires there is added the idea of the same will but belonging to the world of understanding” (G 4:454). I do not see how we can explain the Simultaneity Thesis on the ontological reading.

synthetic character of the law. Notice, however, that he has not said anything that would answer the question, “Why be moral?” When I transfer myself in thought to the world of understanding, I am left with the mere form of lawfulness, without insight into the noumenal world or its laws. My act of adopting this form as a sufficient determining ground of the will—my act of imposing the moral law on myself—is one of pure spontaneity. It is not preceded by any cause (how else could it be spontaneous?), and so there is no prior condition that would explain why I have—or should have—done this. The act of adopting the form of lawfulness (the “fact of reason”) is inexplicable in this sense. And that’s why Kant has no ambition to answer the Question of Authority, the question of how the form of reason is immediately lawgiving with respect to our maxims. The scope of Kant’s argument in subsection four is the phenomenon of obligatoriness. So when he says the world of understanding is binding for us because it contains the “ground” of the world of sense, he is not saying why we should prefer morality to self-love. He is only saying why we experience the actions we “would” do as intelligible beings as actions we “should” do under the conditions of sensibility. As beings who also belong to the world of sense, we do not in theory always do (or want to do) what reason prescribes to us with objective necessity, and that is why we experience the law in the form of an imperative.

4.3 Intuition and Practical Sensibility

My aim in the last two chapters has been to clarify the argument structure of *Groundwork* III and to address a number of exegetical questions that inevitably arise for the reader. I have remained neutral so far about whether Kant’s argument succeeds, but I am now in a better position to elaborate on the “missing step” I spoke of in the Introduction. We have seen that Kant uses a transcendental worry to spur his argument forward, from a metaphysics of morals to a “critique of pure practical reason.” His basic point is that without two different ways of thinking about our will, we will not be able to explain the *autonomous* character of moral obligation. This is what the second stage of the deduction comes down to in *Groundwork* III. Kant puts the worry somewhat misleadingly in terms of the potential illusoriness of moral obligation. But what’s at stake is whether our experience of constraint is an expression of practical reason, i.e., whether we are *constraining ourselves*. And the worry is that perhaps the categorical “ought” is, in some sense, external to our will—a covert hypothetical imperative based on our fear of transgressing God’s rules, or society’s. Kant obscures the point, I think, in speaking of our membership in two “worlds,” but I have argued that his claim is that we can only understand ourselves, as agents, when we take ourselves to act on laws of reason, and that the idea of a “world

of understanding” provides us with a model for thinking about this. With this critical distinction, then, we are in a position to see how the actions we *would* do from the standpoint of the world of understanding are what we *should* do from the standpoint of the world of sense. This is how our experience of constraint is possible.

As we have seen, Kant denies that we can gain theoretical insight into the world of understanding, since it presents nothing for our intuition. Rather, the idea of such a world gives us a model of nature under rational laws, the form of which we cognize in the unconditional imperative of morality. The form of a world of understanding is simply the form of pure practical reason, whose sole requirement is that we develop our maxims out of consideration for their universality. As Kant puts it, after I abstract from every condition of the world of sense, “nothing is left for me but the form of it—namely the practical law of the universal validity of maxims,” and all I can do, Kant says, is to think of my reason conformable to this form “as a possible efficient cause, that is, a cause determining the will” (G 4:462). We cannot say why I should adopt this form as the ground of my will, as opposed to some material principle like self-love. All we can say is that the world of understanding gives me a pattern for thinking of myself as a rational will. It supplies the form of lawfulness under which I can determine myself to action, independently from desire. We experience this act of self-determination in the form of constraint, even though it issues from our own will, because we are also affected by sensible impulses and inclinations. The experience of constraint is possible, then, because we stand in a relation of dissimilarity with ourselves. It is only because we can think of ourselves from two standpoints “at the same time” that we experience practical reason within us the activity of practical necessitation.

Kant summarizes the second stage of his deduction with the example of a “hardened villain.” The example helps bring out what I think is missing from Kant’s argument. He writes:

There is no one—not even the most hardened villain, if only he is accustomed to use reason in other ways—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. He cannot indeed bring this about in himself, though only because of his inclinations and impulses; yet at the same time he wishes to be free from such inclinations, which are burdensome to himself. Hence he proves, by this, that with a will free from impulses of sensibility he transfers himself in thought into an order of things altogether different from that of his desires in the field of sensibility. (G 4:454; modified)¹⁰

¹⁰“Es ist niemand, selbst der ärgste Bösewicht, wenn er nur sonst Vernunft zu brauchen

Kant chooses someone whose physical actions are opposed to morality, whose choices are directed by a principle of self-love, because this will prevent us from evaluating the villain in terms of his life-conduct. The villain is someone who is simply unable to act on his representation of the law because he remains tied down by his inclinations. He still expresses a "wish" that he might become virtuous when presented with examples of virtue. So his "wish" displays his *capacity* to give deliberative priority to moral considerations. Kant's point is that if the pursuit of happiness constituted the sole law of the villain's will, he would have no basis, no vantage point, to be moved by the examples of moral action. His "wish" to be virtuously disposed, therefore, can only issue from the standpoint he implicitly takes when he puts himself, in thought, to an order of nature different than the order of his inclinations.¹¹

But there is a problem here. When the villain adopts the standpoint of the world of understanding, he suspends preference to his sensible commitments, and by doing so he takes himself to act out of consideration for the moral law. From this standpoint, the villain is on equal footing with a holy will, for whom morality is self-descriptive. And yet, the villain is only conscious of himself "under obligation" when he views himself from the world of sense. Kant is clear about this: "from this standpoint he is conscious of a good will that, by his own acknowledgments, constitutes the law for his evil

gewohnt ist, der nicht, wenn man ihm Beispiele der Redlichkeit in Absichten, der Standhaftigkeit in Befolgung guter Maximen, der Theilnehmung und des allgemeinen Wohlwollens (und noch dazu mit großen Aufopferungen von Vortheilen und Gemächlichkeit verbunden) vorlegt, nicht wünsche, daß er auch so gesinnt sein möchte. Er kann es aber nur wegen seiner Neigungen und Antriebe nicht wohl in sich zu Stande bringen, wobei er dennoch zugleich wünscht, von solchen ihm selbst lästigen Neigungen frei zu sein. Er beweiset hiedurch also, daß er mit einem Willen, der von Antrieben der Sinnlichkeit frei ist, sich in Gedanken in eine ganz andere Ordnung der Dinge versetze, als die seiner Begierden im Felde der Sinnlichkeit."

¹¹Kant foreshadows the case of the villain's "wish" early on in *Groundwork* II, when he speaks of the powerful influence examples of moral conduct can have on us: "For the most ordinary observation shows that if we represent, on the one hand, an 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, it leaves far behind and eclipses any similar act that was affected in the least by an extraneous incentive [*fremde Triebfeder*]; it elevates the soul, and awakens a wish [*die Seele erhebe und den Wunsch erzeuge*] to be able to act in like manner oneself" (G 4:410n). In each case, then, Kant is saying that the purity of a moral example only strengthens its motive force. Presumably, this is the basis of the "wish" Kant is speaking of here. The "wish" expresses an attitude of "respect." As Kant says: "What I cognize immediately as a law for me I cognize with respect [*Achtung*], which signifies merely consciousness of the *subordination* of my will to a law without the mediation of other influences on my senses" (G 4:401n). We feel respect for examples of moral action because they reveal to us the law of our own will. Thus the villain expresses a "wish" to be virtuous when presented with moral examples because they make the law of his own will explicit, a law whose authority he recognizes—as Kant says—"even while transgressing it."

will as a member of the world of sense... The moral ‘*ought*’ is then his own necessary ‘*will*’ as a member of an intelligible world, and is thought by him as ‘*ought*’ *only insofar as he regards himself at the same time [zugleich] as a member of the world of sense*” (G 4:455; last emphasis mine). The problem is that, from the standpoint of the world of sense, our will is a mere faculty of desire. As Kant says, if we were only a part of the world of sense, our actions “would have to be taken to conform wholly to the natural law of desires and inclinations, hence to the heteronomy of nature” (G 4:453). How, then, can the villain express a “wish” to be virtuous?

As I see it, there is a further question here about whether the moral law is compatible with the psychological conditions of a human will. Even after we recognize the authority of the moral law as the highest principle of practical deliberation, we are faced with the question of how this principle applies to us subjectively, i.e., how it can function as an interest in the same way (and along the same pathways) that empirical interests do. From what Kant has said, the villain not only acknowledges the authority of the moral law; he *feels respect* for it, and that feeling must be how the moral law influences him. But how is *that* possible? How are we to understand the subjective force of morality? For, as soon we theorize about our psychological capacities, we are immediately brought back to the world of sense, and yet this world restricts us to a view of ourselves as compelled by our inclinations and impulses.

4.4 The Case of the Mafioso

We find a similar problem arise with the character of the Mafioso, introduced in G.A. Cohen’s commentary on Korsgaard’s (1996b) Tanner Lectures. Like Kant’s hardened villain, Cohen’s Mafioso does not live according to moral claims. His commitments extend to a code of “strength and honour” and it does not matter to him if this code requires that he lie, steal, or murder. Cohen raises the example as a challenge to Korsgaard’s view that the normativity of moral obligation depends on an act of “reflective endorsement,” her claim that a principle of action holds for us only if we adopt it into a conception of our identity as practical agents. Reflective endorsement, Korsgaard believes, is how we bind ourselves to norms of choice and action. Cohen’s complaint is that reflective endorsement does not allow us to distinguish the Mafioso’s commitment to a code of honour—which he no doubt feels bound to uphold—and our obligation to morality. The act of endorsement cannot be what generates the normativity of the principle, Cohen argues, for we can endorse whatever principle we want.

What I find relevant here is not the point of Cohen’s criticism; I am more interested in Korsgaard’s response. For Korsgaard takes it upon her-

self to explain how the Mafioso could bring about a *change of practical identity*, how he could come to see his existing set of commitments under the lights of morality. In outline, she says we would have to get the Mafioso in a "place" where he could recognize his humanity, a place that would allow him to see a deep inconsistency with his life of crime. Here, in order to preserve consistency, the Mafioso would have to abandon his code of honour and begin a process of moral self-reform. Korsgaard is confident he would do this if he recognizes his own humanity, since that will make him aware of a commitment more valuable, and hence more authoritative, to any attachment he has qua Mafioso. Presumably he will, like Kant's villain, first express a wish that he could practice the ideal of humanity he discovers in this place we guide him to. And presumably he will feel regret and disappointment for having led a life opposed to it, a life of treating persons (including himself) only as means and never as intrinsic sources of value.

Regardless of how the Mafioso gets to this special place, we are faced with the same question. How can the Mafioso take an interest in the value of humanity? Suppose Korsgaard is right and we can present the Mafioso with a sound argument for why he should treat others (including himself) as intrinsic sources of value, as ends and never as means. And suppose he accepts our argument and identifies with what Korsgaard calls his "rational and active side," or what Kant calls his autonomy. What then? Does any of this explain how the Mafioso could express a wish to change his ways? We are not asking how he could acquire the strength of will to abandon his life of crime under the direction of morality. We are asking about the source of the change itself, the possibility that the Mafioso's recognition of his rational and active side could itself function as an incentive. The mystery is that such an incentive could only arise in what Kant calls the "world of sense," but its underlying cause could only be in the special place Korsgaard leads him to, in what Kant calls the "world of understanding." Again, the question is not whether he has sufficient energy to conform his actions to the moral "ought," but rather how he can be *responsive* to this "ought" in the first place.

Let's step back for a moment. Kant chooses someone whose actions are opposed to morality, and he does so in order to direct our attention away from empirical concerns. Since the beginning of the *Groundwork* Kant has argued that proof of our subjection to morality cannot be decided on the basis of how we act. Thus the character of the villain serves to correct any temptation we might have to search for evidence—for example, to examine the villain's history of conduct, anything that would prove (or disprove) his commitment to morality. Kant is asking us to imagine someone who has never displayed qualities of virtue. But he is not asking us to imagine someone without a capacity to respond to morality. The villain is hardened,

not amoralistic or sociopathic. It may be that he’s only put his reason in the service of selfish ends, but he still has the capacity to use his reason “in other ways”—that is, in purely practical ways. By directing our attention away from what the villain has accomplished in his life, Kant wants to re-orient the question of morality’s application in terms of our capacities. The error of the motivation skeptic from *Groundwork* II is that he oriented the question in terms of what we’ve accomplished (or can accomplish) with our actions. That is why Kant believes the villain “tests the correctness” of his argument. If he can show that someone like the villain is still responsive to morality despite his constitution, he can justify the law’s application to him and, by extension, the law’s application to us.

But Kant’s argument in *Groundwork* III—like Korsgaard’s—moves us in the wrong direction. In each case, the very fact of motivation remains mysterious: namely, the fact that our interest in morality takes effect, not from our “active side,” but from our “passive side,” the side of our sensibility. If we can show that the villain or the Mafioso has given priority to the law, we can say the law applies to him qua rational being. But this does not extend to the further claim, that the law applies to him qua sensible being; for it doesn’t explain how the law can provide him with an incentive or thus function as a “subjective determining ground of the will.”

Let me frame the problem more clearly. It starts with Kant’s claim that we experience morality in the guise of an imperative because we have two standpoints on our will “at the same time.” This is what I have called the Simultaneity Thesis. Here Kant is forced to limit the access we have to ourselves as sensibly affected beings to the mode of intuition. As he puts it, I am conscious of a moral “ought” only when “I intuit [*Anschau*] myself as a member of the world of sense” (G 4:454). In the mode of intuition, however, I can only think of myself under the heteronomy of nature, whereby my will is passive to sensible influences. The standpoint of the world of sense does not let us think of the will as anything “more” than a faculty of desire (as Kant says: if I were only a member of this world, I would act in perfect conformity with the principle of heteronomy). We know ourselves “in the mode of sense” as an appearance, whereby all of our actions fit within a determined series of causes. This claim comes from the first *Critique*: “The human being is one of the appearances in the world of sense, and to that extent also one of the natural causes whose causality must stand under empirical laws. As such he must accordingly also have an empirical character, just like all other natural things” (KrV A546/B574).¹² We find the

¹²“The human being himself is an appearance. His powers of choice has an empirical character, which is the (empirical) cause of all his actions. There is not one of these conditions determining human beings according to this character which is not contained in the series of natural effects and does not obey the laws of nature according to which no em-

same point in *Groundwork* III when Kant says that we can only think of our actions in the world of sense “as determined by *other appearances*, namely, desires and inclinations,” which are part of the “heteronomy of nature” (G 4:453; my emphasis). But if the world of sense limits us to the principle of heteronomy, we are prevented from thinking about how a practical determination of our sensibility is possible.¹³ The problem is that our interest must take effect at the level of our sensibility, but according to Kant’s argument we can only think about our sensibility in terms of heteronomous compulsion.

4.5 Actually or Possibly Motivating?

Kant is sensitive to the problem I have been pressing on, but he thinks he can sidestep it. Later in *Groundwork* III he says that the question of how the moral law can motivate us is incomprehensible, that it would amount to asking how freedom is possible for beings like us (G 4:460). His point is that we can only take an interest in morality by determining ourselves to act independently of sensible influences—and that requires transcendental freedom. The ground of moral interest is incomprehensible, then, because there is no further condition we can cite in a series of conditions that would explain why we determine ourselves. Moreover, causal explanation requires the interaction of two objects of experience, a cause and an effect. But the principle of morality is not an object of experience, so it is impossible to explain the causal force the law exerts on us. The most we can say without overstepping the limits of human cognition is that moral motivation is “not inconsistent” with the idea of freedom. It is not inconsistent with the premises of the deduction that our reason could “instil” (*einzuflößen*) a feeling in us, a so-called “moral feeling.” But the question of how this feeling is possible admits no answer, because it presupposes a mysterious causal relationship between reason and sensibility. That is why the Question of Motivation is beyond deduction.

But are the two questions the same? According to the first stage of Kant’s argument, we can only secure our status as rational agents from the stand-

 pirically unconditioned causality is present among the things that happen in time” (KrV A552/B580).

¹³As Allison puts it, “qua presented to itself in empirical intuition, the rational being is a causally conditioned object in the phenomenal world and therefore incapable of being motivated to act in accordance with the principle of autonomy” (p. 281, note 17). Allison takes this to be reason for rejecting the view that Kant is trying to establish the validity of the categorical imperative for imperfectly rational beings like us. His point is that it would place an impossible standard of success on Kant’s argument, since we can never intuit ourselves as anything more than causally conditioned agents (i.e., agents who can only act on inclination). As I see it, this highlights the fatal shortcoming of *Groundwork* III—that Kant lacked the resources for thinking about the connection between practical reason and sensibility.

point of the world of understanding. This introduces a distinct presupposition, that from this standpoint we adopt the mere form of universal law as a sufficient ground of action. It is on the basis of this presupposition, this “act” of reason, that Kant is able to fill in the concept of freedom with positive content. He can specify the activity of our will through the principle of acting on universally valid principles. But nothing about this activity explains the effect that our consciousness of the law has on sensibility. For beings like us to act out of respect for the law, our reason must have the additional power to induce a feeling of pleasure in the thought of fulfilling our duty. This is the only way our sensibility can conform to principles of reason. As Kant puts it in an earlier lecture: the human being “has no such secret organization, that he can be moved by objective grounds” (LE 27:1429). Human beings are limited in this way: Just as we require *intuitions* coming to us through our sensible faculty in order to cognize objects, we also require *incentives* in order to form maxims of choice. So Kant was not quite right about this: The question of freedom is not identical to the question of moral motivation. If anything, Kant should have said the problem of freedom is identical to the problem of moral authority—how we adopt and impose the moral law on ourselves—because that is what allows us to specify the concept of freedom (cf. KpV 5:29). It is a slightly different question to ask how this act of reason influences us, for this introduces a question about our motivational capacities and the laws that govern those capacities in the world of sense.

The deeper problem, though, is that Kant only raises the question of reason’s capacity to influence feeling from a causal perspective. This comes out clearly in the following passage, which I will quote in full:

In order for a sensibly affected rational being to will that for which reason alone prescribes the “ought,” it is admittedly required that his reason have the capacity to *instil [einzufloßen] a feeling of pleasure* or of delight in the fulfillment of duty, and thus there is required a causality of reason to determine sensibility in conformity with its principles. But it is quite impossible to see, that is, to make comprehensible a priori, how a mere thought which itself contains nothing sensible produces a feeling of pleasure or displeasure; for that is a special kind of causality about which, as about any causality, we can determine nothing whatever a priori but must for this consult experience alone. But since this cannot provide us with any relation of cause to effect except between two objects of experience—whereas pure reason, by means of mere ideas (which yield no object at all for experience), is to be the cause of an effect that admittedly lies in experience—it follows that for us human beings it is quite impossible explain how and why the *universality of a maxim as law* and hence

morality interests us. (G 4:460)¹⁴

As far as this question goes, Kant is right. We cannot explain the metaphysics of moral motivation, for that would require us to have insight into a noumenal cause. We can only address the question of how we actually take an interest in morality with “defence”: we can deny theoretical knowledge of moral feeling but affirm its consistency with the presupposition of freedom. What I want to suggest, though, is that there is an alternative way of raising the Question of Motivation. We can ask how the moral law *possibly* instils a feeling of pleasure in us. And we can raise this question without committing ourselves to the metaphysics of moral motivation. By neglecting this alternative, however, Kant’s argument in *Groundwork* III remains vulnerable to a skeptical threat. For the possibility remains that a human will, in its subjective constitution, is unfit to act out of respect for the law. Kant is right to maintain that the question of rational causality is a puzzle we cannot comprehend. But this does not figure into our experience of morality. Our experience reveals a different problem. We are struck by the gap between our sensibility and the demands that morality places on us, and our worry is that this gap is a real limit. It may be—and this is the skeptical thought—that our sensibility is undeterminable, that we are, at the level of our desires, unresponsive to morality and its claims.

Voicing our worries about this capacity needn’t take the form of a demand for causal explanation. On this point I think Kant is justified in rejecting motivational skepticism from *Groundwork* II. The empirical skeptic doubts our capacity to act out of respect for the law because he cannot observe a single case where he is not driven by self-interest. He expects proof of our interest in morality—lacking this, he denies we possess a moral disposition. Without demanding empirical proof, however, we can still raise the Question of Motivation with a concern for understanding (as much as we can) the connection between reason and sensibility. The problem with

¹⁴I prefer Paton’s translation of “einzufloßen” as “instil” over “infuse” (Abbot/Ellington) and “induce” (Gregor). “Um das zu wollen, wozu die Vernunft allein dem sinnlich-afficirten vernünftigen Wesen das Sollen vorschreibt, dazu gehört freilich ein Vermögen der Vernunft, ein *Gefühl der Lust* oder des Wohlgefallens an der Erfüllung der Pflicht *einzufloßen*, mithin eine Causalität derselben, die Sinnlichkeit ihren Principien gemäß zu bestimmen. Es ist aber gänzlich unmöglich, einzusehen, d. i. a priori begreiflich zu machen, wie ein bloßer Gedanke, der selbst nichts Sinnliches in sich enthält, eine Empfindung der Lust oder Unlust hervorbringe; denn das ist eine besondere Art von Causalität, von der wie von aller Causalität wir gar nichts a priori bestimmen können, sondern darum allein die Erfahrung befragen müssen. Da diese aber kein Verhältniß der Ursache zur Wirkung, als zwischen zwei Gegenständen der Erfahrung an die Hand geben kann, hier aber reine Vernunft durch bloße Ideen (die gar keinen Gegenstand für Erfahrung abgeben) die Ursache von einer Wirkung, die freilich in der Erfahrung liegt, sein soll, so ist die Erklärung, wie und warum uns die Allgemeinheit der Maxime als Gesetzes, mithin die Sittlichkeit interessire, uns Menschen gänzlich unmöglich.”

Groundwork III, however, is that Kant fails to provide an account of this connection that would explain how principles of reason can appear to us in the world of sense, as incentives, without reducing to the heteronomy of nature. He fails to account for a practical (that is, non-pathological) determination of our capacity to feel, such that we could take an interest in the moral law. As a result, he leaves enough space left for empirical theories of motivation to remain as attractive alternatives to those who, out of despair, conclude that the human will is unfit for morality.

Let's step back for a moment. For Kant, the transition to a “critique of pure practical reason” is the final *Übergang* of *Groundwork* III. Beyond this, he says we are confronted with moral presuppositions we cannot justify by way of deduction. The authority of the moral law, and its force as an incentive, remain essential but inexplicable commitments which we can, at best, *defend* against those who presume to know more than us. However, I think Kant is in error to limit the Question of Motivation to the “outermost limit” of moral inquiry. As we have seen, the problem is that he restricts the question of our interest in morality to a speculative issue of how the mere thought of duty can affect our sensibility. This is a question of how morality is *actually* motivating, and to that extent his answer is right: we cannot explain a causal relation between reason and sensibility. But Kant's mistake is that he assumes we can *only* raise the question of moral motivation from a causal perspective. We can still ask ourselves—and Kant will later come to appreciate this—how the mere thought of duty *must affect* us and what the activity of practical reason *must feel like*. When we bracket the causal question, then, we are free to account for how the moral law can enter into the motivational pathways of our will, i.e., how it can appear to us subjectively, and what feelings (negative and positive) must arise as a result. My claim, then, is that with Kant's theory of moral sensibility from the second *Critique*, we have a final *Übergang* missing from *Groundwork* III.¹⁵

¹⁵During the writing of the *Groundwork* Kant was still of the opinion that the concepts of morality do not completely belong to transcendental philosophy, which deals exclusively with a priori cognition. The concepts of pleasure, desire, and life that pertain to the human will are all of empirical origin. For this reason Kant concluded in the first *Critique* that the concepts of morality are not the proper subjects of transcendental inquiry, for “everything practical, insofar as it contains motives, is related to feelings, which belong among empirical sources of cognition” (KrV A15). Kant modified this claim in the B-Introduction by replacing “motives” (*Bewegungsgründe*) with “incentives” (*Triebfedern*), presumably to avoid closing off the possibility of having rational motives to morality (KrV B29). But this is puzzling in light of the second *Critique*, which takes up an account of the *Triebfedern* of pure practical reason. When did Kant arrive at the view that moral feeling is a proper object of transcendental inquiry? I cannot go into this here; but for a related discussion, see Forster (2008). Under this restriction, then, Kant would not have seen human sensibility as offering a space for theorizing about moral consciousness.

4.6 Looking Ahead

We can at least appreciate why Kant appealed to the “outermost limit” of moral inquiry in *Groundwork* III. As soon as we speak about an idea’s influence on our capacity to feel we are talking about a non-natural causality—what Kant calls a “causality of reason”—and we would need an intellectual intuition to grasp this. Of course, our knowledge of objects is limited to natural causes, so any theory that tried to explain our interest in morality would rest on illicit metaphysical claims. Alternatively, we could connect the moral law to our will by way of some material object, like happiness. But in that case the principle of morality would be heteronomous. Given these alternatives, Kant rejects any attempt to explain our interest in morality in speculative or psychological terms. The attempt runs a double risk: either it appeals to metaphysical claims that fall outside the bounds of human cognition (and “impotently flaps its wings” among pure concepts), or it appeals to empirical claims that fall within the bounds of theoretical cognition (and endlessly “searches among the world of sense”). Under these constraints, Kant understandably turns to “defence” in *Groundwork* III.

Interestingly, Kant does not overstep this limit in Chapter III of the *Critique of Practical Reason* (hereafter, Ch.III). He writes: “For how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem and identical with that of how a free will is possible” (KpV 5:72). On this point, the second *Critique* is consistent with the “outermost limit” marked out in *Groundwork* III. However, I will argue in Part II that the second *Critique* goes further. It goes further by accounting for the *effects* the moral law has on our sensibility, effects that allow us to understand the experience of practical reason peculiar to beings like us. In what follows, then, I will argue that in the second *Critique* we find Kant re-orienting the Question of Motivation from a causal question (“How can morality immediately determine the will?”) to a phenomenological question (“What effects must this determination have on our sensibility?”). In both *Groundwork* III and the second *Critique* Kant maintains that the former question is unanswerable, and that previous systems of morality were mistaken in their attempts to answer it, either by claiming direct insight into the noumenal world (rationalism) or by identifying moral interest with a material motive (empiricism). This puts us against a new set of questions. What lead Kant to reorient the Question of Motivation to a question of the *effects* of practical reason on our sensibility? What is Kant’s method of argument in Ch.III, and why does it focus on our capacity to feel? What are moral feelings anyway, and how do they differ from empirical feelings? And how does all of this figure into Kant’s project of moral justification? How does Kant’s theory of moral sensibility

4. THE MORAL "OUGHT"

address a skeptical threat? Is it successful?

My task in Part II will be to answer these questions.

Part II

The Solution

CHAPTER 5

Moral Feeling

OVER the last twenty years, a small but exciting literature has emerged on the topic of Kant's theory of moral motivation.¹ This is where one will find the most focused debate over Kant's account of moral feeling, specifically in Ch.III of the second *Critique*. I have learned a great deal from these studies, but I also want to distance myself from their underlying concerns. In particular, I want to distance myself from a reading of Ch.III that frames the concept of moral feeling in terms of his doctrine of moral worth. I shall suggest that what we find in Ch.III—and what is new to Kant's ethics in 1788—is a way of thinking about human sensibility that does not reduce to heteronomous compulsion.

The current interpretive dispute misses this. If we think Kant is trying to show how the feeling of respect functions in generating moral action, then the scope of the chapter quickly narrows down to two possibilities. Either our intellectual recognition of the law's authority is the ground of respect, the active motivating factor, in which case any feeling we have (say, our feeling of "humiliation") will be a phenomenal side-effect of this experience. Or, the phenomenal side-effect, the feeling we have in response to our recognition of the law, will be how this recognition becomes effective in human conduct, and that will be the active motivating factor. This is an oversimplification of course, but it outlines the interpretive dispute between

¹In the Anglophone literature, see Allison (1990, ch.6); Ameriks (2006, ch.4); Guevara (2000); Herrera (2000); McCarty (1993; 1994; 2009, ch.6); Reath (2006, ch.1); Zinkin (2006); Morrisson (2008). For earlier studies, see Paton (1947), Beck (1960), Wolff (1973), and Broadie and Pybus (1975).

intellectualist and affectivist readings of Ch.III.² My worry is that both interpretations, while insightful in different ways, fail to provide a genuine “aesthetic” of pure practical reason. On the one hand, by locating the active motivating factor of respect in our recognition of the law, the intellectualist is forced to view the feeling of respect, i.e., the painful “checking” of the inclinations, as a secondary feature of moral experience. On the other hand, by reducing this factor to an empirical feeling, the affectivist comes dangerously close to framing moral motivation as a vector of non-cognitive forces, which may reduce the moral law to a heteronomous principle. In each case, we are left without a full account of the *connection* between reason and sensibility, which is what Kant’s theory of moral sensibility is meant to provide.

Rather than engage in this debate directly, then, I want to propose an alternative reading of Kant’s theory. This will be my first task in this chapter. In what follows, I will outline Kant’s broader argumentative strategy in the Analytic of the second *Critique*, paying attention to the concept of “ectypal” nature that arises in Ch.I and Ch.II. In the previous chapter, I argued that Kant’s mistake in *Groundwork* III was to conflate two ways of raising the Question of Motivation—how the moral law *actually* motivates us and how it *possibly* motivates us. With the concept of ectypal nature, I believe that he is making a conscious effort to overcome this problem. My next task will be to explain Kant’s account of the self-reflexive character of feeling, his claim that one’s feeling for an object is, at the same time, a feeling of oneself. This reflexivity underlies the feelings we have for empirical incentives, but also for what Kant calls the “incentives of pure practical reason,” which is the title of Ch.III.³ Of course, not everyone has been convinced that this chapter achieves what it sets out to accomplish. Kant’s immediate successors—notably Schiller and Hegel—believed that Kant lacked the resources to establish a *positive* connection between reason and sensibility.

²As McCarty (1993) nicely summarizes it: “Intellectualists hold that respect for the moral law is, or arises from, a purely intellectual recognition of the supreme authority of the moral law, *and* that this intellectual recognition is sufficient to generate moral action independently of any special motivating feelings or affections. Opposed to the intellectualist interpretation is what I shall call the *affectivist* view. Affectivists need not deny that Kantian moral motivation initially arises from an intellectual recognition of the moral law. Contrary to intellectualists, however, they maintain that it also depends on a peculiar moral feeling of respect for law, one consequent to the initial recognition or moral judgment the intellectualists emphasize exclusively” (p. 423).

³“*Triebfeder*” is a term of art for Kant, and there is no satisfying English translation. Literally, “*Triebfeder*” means something like “driving spring.” Abbot translates it simply as “spring.” I will follow Gregor’s use of “incentive,” although this is not without its difficulties. “Incentive,” in English, is close to “motive”—although Kant wants to keep the two separate. Motives for Kant are objective grounds of action valid for all rational beings, including divine or holy wills. Incentives, however, refer specifically to beings who, like us, are also affected by sensible influences. For an illuminating discussion of these translation issues, as well as the motive/incentive distinction, see Herrera (2000).

In the final section of this chapter, I will provide a sketch of this challenge and show why, on closer scrutiny, it rests on an uncritical conception of the “harmony” between duty and inclination.

5.1 The “Aesthetic” of Pure Practical Reason

The second *Critique* is divided into an Analytic and a Dialectic (Book One), with a concluding Methodology (Book Two). The Analytic falls into three chapters, each under the heading of “pure practical reason”: “principles” (Ch. I), “objects,” (Ch. II), and “incentives” (Ch. III). At one point Kant refers to Ch.III as an account of the relationship between reason and sensibility, or what he calls an “aesthetic” of pure practical reason (KpV 5:90). There is a parallel here, as Kant points out. The Aesthetic in the second *Critique* is the final chapter of the Analytic because it addresses the question of how practical principles can appear to us subjectively, as incentives, and thereby apply to the formation of maxims. This reverses the order of the Analytic in the *Critique of Pure Reason*, which began with our capacity to be affected by objects and then proceeded to ask how objects of experience relate to pure concepts of the understanding. The parallel can be misleading, however, and Kant admits the division of transcendental Logic and Aesthetic is “not altogether suitable” in the practical context. Somewhat tentatively, he asks that he “may be allowed, merely by an analogy, to use these terms” (KpV 5:90).

Although Kant does not speak in this way, what we might call “theoretical sensibility” is our capacity to be affected by objects of experience, and “practical sensibility” our capacity to be affected by incentives (including, for Kant, the incentives of pure practical reason). In this respect, the term “aesthetic” is appropriate; but there are still differences between the two, and Kant’s remark can be misleading. For one thing, practical sensibility does not have a two-fold division into forms of intuition, space and time. Indeed, practical sensibility lacks a cognitive function: it is “non-intuitive,” meaning that it does not pertain to the reception of objects in the formation of synthetic judgments. Kant recognizes this when he says it “is not regarded as a capacity for intuition at all but only as feeling (which can be a subjective ground of desire)” (KpV 5:90). The point holds for all kinds of feeling, not just to our feeling for the moral law. When I have an empirical incentive, say to eat a strawberry, I must intuit the strawberry in space and time. But the feeling of pleasure the representation elicits from me is subjective: it relates only to my constitution.⁴ The feeling-of-pleasure for the

⁴This claim is not unique to the second *Critique*: “But what is a feeling? That is something hard to determine. We sense ourselves... The subjective representation of the entire power of life for receiving or excluding objects is the relation of satisfaction or dissatisfaction. Thus

strawberry is what Kant calls a feeling-of-life, life in our appetitive capacity as physical beings.⁵ Whatever elicits the feeling of life is practical—a basis for making choices and taking action. So while the strawberry appears to me in a manifold of appearances (i.e., an object with particular shape, texture, and colour), *its desirability is not a feature of the manifold*.⁶ Kant believes this is also true of moral feeling, although with one crucial difference: we can know *a priori* how our representation of the law will elicit feelings of pleasure and displeasure in us—a point I will return to in section 5.4.

I suspect Kant neglected these differences in the section under discussion, the “Critical Elucidation of the Analytic,” because his aim there is to justify the reverse-order of the second *Critique* (i.e., its progression from principles, to concepts, to incentives), by comparison to the first *Critique*. If we had to identify the theoretical parallel to Kant’s “aesthetic” of pure practical reason, I think our best bet would be §24 and §26 of the B-Deduction, where we find Kant struggling to show how the categories apply to specific objects of experience. For lack of space, I cannot pursue this comparison in any detail, but I think Lewis White Beck is on the right track in relating Kant’s theory of moral sensibility to the role of pure imagination. As he writes: “We are concerned not with the effect of objects on the sensibility, since we are not concerned with relating our concepts to given objects as objects of knowledge, but with the effect of reason itself and its principles and concepts upon our sensibility. The problem is that of the subjective determination of the will, the subjective factor being provided by sensibility or feeling and the determining factor by the principles of practical reason. The analogous problem in the *Critique of Pure Reason* is perhaps that of the pure imagination, which represents in the sensibility the synthetic operations of the understanding” (1960, p. 67). Regrettably, Beck does not develop this suggestion further, so it is up to us to provide some of the missing links.

Kant’s transcendental deduction of the categories is notoriously obscure and has generated some of the most heated debates in Kant scholarship over the past century. For the most part, commentators agree that Kant revised the argument strategy of the Deduction in the B Edition of the first *Cri-*

feeling is the relation of objects not to the representation, but rather to the entire power of the mind” (LM 28:247). “Feeling consists in the relation of a representation not to the object, but to the entire subject. Pleasure and displeasure are not cognitions at all” (LM 28:586). “The feeling of pleasure is the ability of my power of representation to become determined by a given representation to its maintenance or avoidance. With displeasure we summon up our entire faculty to prevent a representation from penetrating further into the mind” (LM 29:890-891). Also, the concept of “self-feeling” (*Selbstgefühl*) is not unique to Kant. For an illumination discussion of pre-Kantian views, focusing on Christoph Meiners and Michael Hissmann, see Theil (1997).

⁵I will come back to this in section 5.3.

⁶Speaking of a “manifold of desires,” as Kant does in the second *Critique*, is a just a way of highlighting the subjective character of the inclinations.

tique, published in 1787. They also agree that the B-Deduction contains two steps, or what Henrich (1968) has famously called a “two-steps-in-one proof.” There is, however, a good deal of controversy over what these “two steps” consist of and how they are supposed to work in Kant’s argument. More recently, Henry Allison (2004) has made a good case for thinking of the first step as an attempt to establish the objective validity of the categories for thinking objects in general, and the second step as an attempt to establish the objective reality of the categories for objects of experience in particular. On this reading, then, after Kant has shown that the categories are necessary for a discursive intellect *in general* (i.e., an intellect that requires a manifold to cognize objects), the question remains whether the categories connect to *our* form of intuition, which is spatio-temporal.

Recall what I said earlier—in section 2.4—about the skeptical worry Kant used to motivate the transcendental deduction in the first *Critique*. He said it is conceivable at this stage of the discussion that appearances may be constituted in such a way that they would not conform to the pure concepts of the understanding. This is a problem of application. For example, my perception of successive states may be nothing more than a “bundle” of sense impressions (to use Hume’s expression), i.e., a series of appearances that do not reflect a rule of synthesis according to the concept of cause and effect. The concept of cause, then, may be nothing more than a function of unity for thinking, but not of experience—“so that this concept would thereby be entirely empty, nugatory, and without significance” (KrV B123). Kant reminded us, however, that we would still *intuit* appearances, because appearances must be given in space and time.

While this reminder did not seem like much at the time, it proves to be essential to Kant’s completion of the deduction in §24 and §26.⁷ In a recent study, Béatrice Longuenesse has suggested that Kant’s task is to show that space and time, as a priori intuitions, are themselves generated by a synthetic activity, and that this activity links the categories to the manifold of empirical intuitions.⁸ What’s remarkable is that Kant assigns this synthetic activity to the “pure” (or what he also calls the “productive”) imagination. The problem of application first raised at (B123)—that the categories may be functions of thinking, but not of experience—is thus solved by an appeal to the pure imagination, whose synthesis gives unity to the very forms of intuition in which objects are given to us. The pure imagination, then,

⁷For a rigorous interpretation of these sections, see Longuenesse (1998, ch. 8).

⁸As Longuenesse summarizes the second step: “if we accept the argument of section 26, the very fact that appearances are given in space and time is sufficient ground for their being in conformity with the categories, even though it remains true that they are not *in* a category (as “in” an intuition) or even cognized *under* a category until the relevant operations of comparison/reflection/abstraction, together with a a priori construction, have generated such cognition” (p. 226).

serves to mediate between sensibility and the understanding.

Without going into the details, I think we can begin to see a relevant parallel. The task of the “aesthetic” of pure practical reason—like the task of §24 and §26 in the B-Deduction—is to overcome the apparent gap between reason and sensibility. In the B-Deduction, the productive imagination serves to overcome this gap, as it gives unity to the forms of intuition, space and time, which underpin every empirical intuition of a manifold. In Ch.III of the second *Critique*, our *capacity to feel* serves to overcome the gap—since Kant will argue that we can analyze the effects the moral law must have on our faculty of desire, effects that manifest as feelings of pleasure or displeasure. In this respect, the concept of feeling clears room for thinking about a subjective determination of our will, even though Kant admits that it will forever be a mystery how this subjective determination is possible. We cannot comprehend how the moral law functions as an incentive for us; but we can comprehend what effects on our sensibility must arise as a result. For Kant, this is sufficient to overcome the skeptical thought that a human will is unfit for morality. But rather than pursue these parallels further, I want to find further support for my interpretation by situating Kant’s theory of moral sensibility in relation to the Analytic of the second *Critique* itself. While the Analytic is complex territory, I want to focus on a set of issues that, I hope, will clarify the aims and ambitions of Kant’s “aesthetic” of pure practical reason. These include the concepts of *ectypal* nature, the *typic* of pure practical judgment, and the *causa noumenon*. In discussing the “typic” I will also outline the Schematism chapter from the first *Critique*, and here again the concept of the “pure” imagination will prove to be essential for Kant’s argument.

5.1.1 ECTYPAL NATURE

Kant introduces the concept of “ectypal” nature in Ch.I and returns to it in his discussion of the “typic” in Ch.II. At (KpV 5:43) he says our practical cognition of morality directs us to the idea of an intelligible world and gives it positive content, namely, a law. We heard a similar claim in *Groundwork* III, but now Kant adds an important point. He says that the intelligible world in turn gives form to the world of sense “without infringing upon its mechanism.” A few lines later he says the “counterpart” of this idea “is to exist in the sensible world but without infringing upon its laws” (KpV 5:43). Without explaining this, he goes on to distinguish between an archetypal and ectypal world, and says the latter “contains the possible effect of the idea of the former as the determining ground of the will” (KpV 5:43). While this may not seem obvious to us, Kant adds that “the most ordinary attention to oneself confirms that this idea is really, as it were, the model [*Muster*] for

the determination of our will” (KpV 5:43). The example Kant goes on to discuss is how we can apply the categorical imperative to specific cases of judgment through the Formula of the Universal Law of Nature (FULN). I think this sheds some light on what Kant is saying here.

Kant uses FULN to distinguish two ways of thinking about “an enduring natural order.” On the one hand, we can think of an order of causality in terms of “actual nature,” i.e., the nature studied by the scientist or empirical psychologist, in which case the only model for the determination of our actions will be a principle of heteronomy. As I understand it, Kant’s claim is that when we think about what enduring order will manifest from “actual nature,” we will help ourselves to pathological laws (presumably those from our own psychology) and imagine forming maxims out of consideration for our “sum-total” well-being. Because pathological laws are, for Kant, a species of physical laws, we will ultimately be thinking of a deterministic order of nature along with ourselves as causally conditioned objects. On the other hand, with FULN we can think of a different order of causality in terms of “ectypal nature,” which Kant describes as the possible effect of the intelligible world. In this case the model for the determination of our actions will come from the principle of autonomy. Kant’s claim is that when we think about what enduring order will manifest from “ectypal nature,” we will help ourselves to the rational laws of the intelligible world and thereby imagine forming maxims out of consideration for their validity. Because rational laws are, for Kant, the same as moral laws, we will be thinking of a moral order within the basic framework of causality found in the world of sense. So we will not be “infringing” upon the mechanism of the latter.⁹

5.1.2 THE TYPIC

Kant returns to the concept of ectypal nature in Ch.II, entitled “On the Objects of Pure Practical Reason.” The objects in question refer to the morally good and evil, which Kant identifies with the maxim of one’s choice rather than the results or effects of one’s actions.¹⁰ In this way, the objects of prac-

⁹This is Kant’s statement of FULN from the *Groundwork*: “Since the universality of law in accordance with which effects take place constitutes what is properly called nature in the most general sense (as regard its form)—that is, the existence of things as it is determined in accordance with universal laws—the universal imperative of duty can also go as follows: act as if the maxim of your action were to become by your will a universal law of nature” (G 4:421).

¹⁰The first part of Ch.II concerns the distinction between the morally good or bad and the agreeable or disagreeable. Kant gives the example of the Stoic “who in the most intense pains of gout cried out: *Pain*, however you torment me I will still never admit that you are something evil.” We can laugh at him, Kant adds, but what he said is correct. “He felt that the pain was an ill, and his cry betrayed that; but he had no cause whatever to grant that any attached to him because of it” (KpV 5:60). Evil only refers to the maxim of the will, which

tical reason hold for “every reasonable human being” (KpV 5:61), independently of what we find privately agreeable. Since they concern the maxim of the will, “good” and “evil” presuppose a rule of rational choice. In the case of pure practical reason, the rule is the condition of bringing one’s maxim into conformity with universal law, and the action that satisfies this condition is “good in every respect” (KpV 5:62). As we know, a rule of pure reason sets down an unconditional law, a law for how we should act in all cases and across situations. Kant calls this a “law of freedom” because it presupposes our capacity to act independently of external influences (KpV 5:67).

There is one problem here. A law of freedom, as a universal rule of action, must still be possible in the world of sense, because “all cases of possible actions that occur can only be empirical, that is, belong to experience and nature” (KpV 5:68). A law of freedom under which we determine ourselves must occur *in* the sensible world, although it cannot arise *from* that world. How, then, can we judge according to the concepts of “good” and “evil”? What we need is a case that would allow us to evaluate the application of a universal rule to a particular action—and yet no empirical action can provide this. So, Kant writes, “it seems absurd to want to find in the sensible world a case which, though as such it stands only under the law of nature, yet admits of the application to the supersensible idea of the morally good, which is to be exhibited in it *in concreto*” (KpV 5:68).

We have a familiar problem on our hands. Kant reminds us of the difficulty he faced in the first *Critique*, where the question arose how the concepts of pure understanding—the rules for the synthesis of objects of experience—apply to specific intuitions. A concept must be homogenous with the object it represents, which is just what it means to say “an object is contained under a concept” (KrV A137/B176). It is easy to see how empirical concepts fit with their objects; for example, the concept of dog as a four-footed animal enjoys homogeneity with a particular Golden Retriever. Without being limited to any particular intuition of a dog, the *schema* of a dog allows me to imaginatively sketch out the shape of a four-footed animal. This is what Kant calls “schematism,” which he describes as “a hidden art in the depths of the human soul” (KrV B181).

The problem is that categories, as pure concepts, cannot be encountered in experience; they are un-homogenous with everything empirical. The question arises, then, how the application of categories to appearances is possible, and this is where the difficulty lies. We can speak of objects “under” empirical concepts because the two always have some degree of homogeneity. But how can we speak of objects “under” pure concepts? We

must always arise from freedom. The pain could not lower his evaluation of self-worth, because he knew he was not responsible for it. He had done nothing morally wrong.

cannot experience categories in any appearance; for example, I do not experience “a quantity” or “a causality.” What we need, then, is a “third thing” mediating between appearances and the categories. This is what Kant labels the “transcendental schema,” which he says “must be pure (without anything empirical) and yet intellectual on the one hand and sensible on the other”(KrV A138/B177).

Unfortunately, Kant says the solution he found in the first *Critique* does not generalize to the practical context. In the first *Critique* he could show how pure concepts and intuitions apply through the mediation of transcendental schemata, products of the pure imagination which exhibit an activity of synthesis in conformity with the categories. The schemata provide cases for judging the particular application of the categories according to their function in time. Kant does not go into any detail here, but his thought in the second *Critique* seems to be this.¹¹ While we can find a corresponding intuition (in conformity with the categories) in the activity of pure imagination, we cannot find a corresponding intuition of the idea of moral freedom. The idea of moral freedom is wholly intelligible, so there is nothing for the imagination to trace, no schema in intuition, in conformity with the laws of freedom.

Fortunately, though, the question of how an action in the world of sense is possible under laws of freedom does not concern its *physical possibility*, and this is a distinction we must keep in mind. “Here,” Kant writes,

we have to do not with the schema of a case in accordance with laws but with the schema of the law itself (if the word schema is appropriate here), since the *determination of the will* (not the action with reference to its result) through the law alone without any other determining ground connects the concept of causality to conditions quite other than those which constitute natural connection. (KpV 5:68-69)

For readers of the first *Critique*, this may seem to violate Kant’s epistemology. How can we apply categories *beyond possible experience*? We know that, for Kant, we can think whatever we like so long as we do not contradict ourselves. Thinking requires the use of concepts, but concepts become objects of cognition only with corresponding intuitive data. As much as we contribute to the order of experience by bringing intuitions under rules of synthesis—the categories—we cannot make objects real just by thinking them. The human mind is limited in this way: it requires intuitions, which are passively received through sensibility, to cognize objects of possible experience. As Kant puts it: “it is possible experience alone that can give our

¹¹There is a second reason: judging the concrete application of moral principles does not depend on the form of inner sense. This is of course central to Kant’s solution in the Schematism. See (KrV A 139/B178).

concepts reality; without it, every concept is only an idea, without truth and reference to an object” (KrV A489/B517). It seems, then, that the concept of freedom remains empty, and that connecting the category of causality to an intelligible ground is, at best, without contradiction—a logical possibility as opposed to a real possibility. How does Kant propose to get around this?

5.1.3 THE CAUSA NOUMENON

In the first *Critique* Kant makes it clear that the categories are only restricted to sensibility when the question is how our experience of objects is possible. Without sensible intuition, the categories remain “empty.” But the categories are not restricted in this way when the question is how agency is possible, for then we are talking about the subject in action, not about the object in appearance (KrV B 166n). In the former case the categories may have application without sensibility, i.e., their content may derive, not from intuition, *but from the activity of reason in its practical capacity*.¹²

I do not want to complicate the matter with unnecessary textual detail, but I think we can clarify what Kant is saying here by reading this passage in light of his account of the *causa noumenon* from (KpV 5:49-50). One thing to keep in mind is that, for Kant, the categories fall into two general classes, the mathematical (quantity, quality) and the dynamical (causality, necessity). Of the former, Kant says we can never expect to find the unconditioned, since quantity and quality always belong to the sensible manifold, hence to a series of conditions. The dynamical categories are not restricted in this way; for example, causality as a rule of synthesis does not require homogeneity between the condition and the conditioned (KpV 5:104). This lack of homogeneity is what enabled Kant to resolve the Third Antinomy.

The topic of the Third Antinomy deserves a separate study, and I am only skimming the surface here. But let me try to narrow in on what Kant thinks is relevant for his argument in the second *Critique*. He says that with the category of causality we can put the unconditioned of a causal series in the intelligible world, which is compatible with the strictly determined series of conditions that follow as appearances in the sensible world. As Kant

¹²After Kant writes, “*no a priori* cognition is possible for us except solely of objects of possible experience” (KrV B166), he adds: “So that one may not prematurely take issue with the worrisome and disadvantageous consequences of this proposition, I will only mention that the categories are not restricted in **thinking** by the conditions of our sensible intuition, but have an unbounded field and only the **cognition** of objects that we think, the determination of the object, requires intuition; in the absence of the latter, the thought of the object can still have its true and useful consequences for the **use** of the subject’s **reason**, which, however, cannot be expounded here, for it is not always directed to the determination of the object, thus to cognition, but rather also to that of the subject and its willing” (KrV B166n).

puts it: “I grant the mechanism of natural necessity the justice of going back from the conditioned to the condition ad infinitum, but on the other side I keep open for speculative reason the place which for it is vacant, namely the intelligible, in order to transfer the unconditioned to it. But I could not realize this *thought*, that is, could not convert it into *cognition* of a being acting in this way, not even of its mere possibility” (KpV 5:49). Now, however, Kant says that if we can find a principle that determines the causality of a freely acting being, we can “convert” the category of an intelligible cause from a mere thought to a (practical) cognition. What we need is a principle that “does not call upon something else as the determining ground with respect to its causality but already itself contains this determining ground,” and that is the principle of morality (KpV 5:105). The *act* by which reason defines its practical capacity—the act of adopting the mere form of lawfulness as a sufficient ground of action—is what supplies (practical) content for the thought of an intelligible cause.¹³ This is how the concept of causality, which “is always found a priori in the understanding, even independently of any intuition” (KpV 5:49), acquires significance apart from intuition.¹⁴ But what does this have to do with the concept of ectypal nature?

Recall what Kant says, that “the *determination of the will* (not the action with reference to its result) through the law alone without any other determining ground connects the concept of causality to conditions quite other than those which constitute natural connection” (KpV 5:68-69). The conditions are different from any natural causal connection because they derive from an intelligible ground, i.e., our will. In Ch.II, then, Kant is saying we can place the concept of an intelligible cause as the ground of a possible enduring order of nature, and this is how we can think of the world of sense as the possible effect, or *ectype*, of freedom. For reasons that should now be clear, the word “schema” is not appropriate here, and that is why Kant calls the sensible application of rational laws a “typic” (KpV 5:69). A schema refers to a sensible intuition, traced by pure imagination, that already displays synthetic conformity to the pure concepts of the understanding. A typic refers to a law of nature that could arise from the free exercise of our will according to the concepts of the morally good or evil. In the second *Critique*, Kant is clear that the typic arises from the understanding, not the imagination (KpV 5:69). A typic is not an idealized sensible intuition, but

¹³“The moral law is, in fact, the law of causality through freedom and hence a law of the possibility of a supersensible nature” (KpV 5:48).

¹⁴Recall the opening sentence of *Groundwork* III: “*Will* is a kind of causality of living beings insofar as they are rational, and freedom would be that property of such causality that it can be efficient independently of alien causes determining it” (G 4:446). The causality of a free will must be understood here as an a priori concept of the understanding, prior to its schematization, which acquires positive content only through our cognition of the moral law.

a form of sensibility determined by an intelligible ground. In this respect, the function of the *typic* in the second *Critique* is significantly different from the schema in the first *Critique*. The latter involves the conformity of specific intuitions to the categories (by way of their transcendental function in time), whereas the former involves the conformity of the moral law to the world of sense.

With a *typic* in hand, then, we do not appraise the possibility of actions according to physical-pathological laws, i.e., according to *actual nature*. We appraise the possibility of actions as the effects of *rational* laws, effects that imprint in nature, as it were, the idea of ourselves as intelligible beings. We can thus appraise our maxims in view of whatever enduring order of nature would follow as the possible effect of *our autonomy*. Of course, we needn't be conscious of this framework in the everyday practice of judgment. Kant says we use a *typic* of nature whenever we evaluate maxims according to the concepts of good or evil. It is as simple as asking ourselves if we could participate in a natural order that would in fact spring from our own volition (KpV 5:69). Can I will to participate in an order of nature where I can end my life because I'm sick of it; where I can borrow money knowing I can't pay it back; where I can leave my talents uncultivated; or where I can refuse helping others in need? (G 4:442). In asking ourselves such questions, we do not need anything like a schema, because the concepts of the morally good or evil are not conditioned by anything sensible. All we need as a standard of judgment is a type of natural law arising from the idea of our own freedom.

At the risk of cutting my discussion short, I want to put things in review. Once Kant has shown that reason is objectively practical, as part of his doctrine of the "fact of reason" in Ch.I of the *Analytic*, he still needs to situate the moral law in its application to human judgment. This is his task in Ch.II. As I have argued, the concept of the *typic* plays a key role here because it shows how the moral law connects to our faculty of judging particular cases of action that conform to universal rules. While Kant does not make this explicit, I believe the concept of *ectypal nature* also plays a key role in his account of moral feeling from Ch.III. For, if we can think about our sensibility in a way that does not reduce to heteronomous compulsion, a space opens up for thinking about what the activity of practical reason must feel like. I have only been able to touch the surface here, but we can begin to see that what's at stake in each of these chapters is a question of how the moral law connects to us—namely, to our faculty of judgment and to our faculty of desire. My task in the next section, then, is to explain what Kant means by desire, and how it relates to our capacity to feel. In the final section I will briefly consider an objection raised by Kant's successors that, if correct, would show that his theory of moral sensibility is empty.

5.2 Life, Desire, Feeling

In the Introduction to the second *Critique*, after dismissing the critical reviews of the *Groundwork*, Kant says “a further objection could have been raised” (KpV 5:9n). Here, in a footnote, he tells us one could have criticized the *Groundwork* for not defining the concepts of life, desire, or feeling. It is hard to tell if this is Kant’s way of passing criticism on his earlier account. He does not say anything to indicate a fundamental change of thinking from 1785. Instead, Kant tells us the complaint would be “unfair” because his account of these concepts in psychology “could be reasonably presupposed” (KpV 5:9n).¹⁵ What’s important, though, is that for the remainder of the footnote Kant provides three “transcendental definitions” which he did not provide in the *Groundwork*. This is an important but often overlooked development in Kant’s ethical thought.

For Kant, a definition is “transcendental” when it explains an empirical concept, like pleasure, in a priori terms—not in order to fix its status, cognitive or affective, but in order to leave its status open.¹⁶ Kant offers us three in the above-mentioned footnote (I will distinguish them with the subscript “T”): Life_T is “the faculty of a being to act in accordance with laws of the faculty of desire.” The Faculty of Desire_T is “a being’s faculty to be by means of its representations the cause of the reality of the objects of its representations.” And Pleasure_T is “the representation of agreement of an object or of an action with the subjective conditions of life” (i.e., with the faculty of desire).¹⁷ Because these definitions do not settle the question of whether we can be affected by principles of reason, Kant says they allow us to “proceed like the mathematician,” who leaves the data of his problem undetermined “in order to bring its synthesis under pure arithmetic.”¹⁸ Kant’s point is that

¹⁵Presumably Kant is referring to the chapter on psychology offered in his *Lectures on Metaphysics*. See, for example, the *Metaphysik Mrongovious* (LM 29:891); *Metaphysik L1* (LM 28:246); *Metaphysik L2* (LM 28:586); *Vigilantis K3* (LM 29:1013).

¹⁶Kant only gives it this particular name in the First Introduction to the *Critique of the Power of Judgment* (KU 20:230). He had long since warned against the dangers of providing definitions or hypothesis prematurely. See Beck (1956) for a fuller discussion of Kant’s theory of definition.

¹⁷Cf., (LM 29:891): “Whatever excites the feeling of the promotion of life, arouses pleasure.” And later: “Pleasure is the representation of the agreement of an object with the productive power of the soul, and displeasure the opposite. The faculty of desire is the causality of the object which is produced. Accordingly, pleasure is the agreement and displeasure the conflict with our faculty of desire” (LM 29:893).

¹⁸The full passage occurs in the third *Critique*. Kant writes: “It is useful to attempt a transcendental definition of concepts which are used as empirical principles, if one has cause to suspect that they have kinship with the pure faculty of cognition a priori. One then proceeds like the mathematician, who makes it much easier to solve his problem by leaving its empirical data undetermined and bringing the mere synthesis of them under the expressions of pure arithmetic” (KU 20:230). He also says we can think of empirical concepts in this way

we can theorize about the psychological presuppositions of a human will (our faculty of desire, and our capacity to feel) without prematurely settling the question of whether the moral law is compatible with these presuppositions.

Taking the footnote at (KpV 5:9n) as a clue, I want to suggest that Kant's transcendental definition of feeling allows him to account for the "effects" of practical reason on our sensibility. Thus the concept of moral feeling allows us to understand a practical determination of our faculty of desire, and this bridges the gap between reason and sensibility, in the same way that the typic from Ch.II bridges the gap between universal rules and particular actions. I think this explains Kant's otherwise puzzling use of causal terms in Ch.III. He says: "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" (KpV 5:72). The "effects" Kant is interested in are not phenomenal effects, i.e., causal relations between two objects of experience. Our practical cognition of the law as a "fact of reason" is not the phenomenal cause of our feeling of respect (with respect as the "effect" of this consciousness). Moral motivation is not like one moving billiard ball striking another and causing it to move, with the motion of the second ball as the "effect" of the first hitting it. For one thing, our consciousness of the law is the result of our own freedom, the result of our taking up the law as a sufficient determining ground. We cannot intuit spontaneity of will, so the moral law cannot be a phenomenal cause in this sense. But remember, Kant thinks our cognition of the law gives content to the idea of freedom, and this allows us to connect the concept of causality (as it resides in the understanding prior to experience) to an intelligible ground: namely, our will. In section 5.1.1, I argued that the idea of noumenal causation is central to Kant's discussion of ectypal nature as a possible "effect," in the world of sense, of our own moral freedom. I now want to suggest that it is central to Ch.III—that moral feeling is the possible "effect" that our taking up the law has on our faculty of desire.

Glancing back for a moment, we can see the failure of *Groundwork* III more clearly. Kant lacked the conceptual resources to think about sensibility that did not reduce to heteronomous compulsion. His mistake was to think of causal explanation as *phenomenal* explanation. And that's why he placed the concept of moral feeling on the "outermost boundary" of moral inquiry (as he put it: "for that is a special kind of causality about which, as about any causality, we can determine nothing whatever a priori but must for this consult experience alone" (G 4:460)). In the outline I have just provided, it should be clear that Kant is not *overstepping* this limit in the second *Critique*.

if we recognize their "kinship" or "affinity" with reason (KU 2:230n; cf., KU 5:177n).

Indeed, he restates it at the beginning of Ch.III: “For, how a law can be of itself and immediately a determining ground of the will (though this is what is essential in all morality) is for human reason an insoluble problem” (KpV 5:72). With the concept of “ectypal” nature, however, I think Kant has made a breakthrough, because he is now in a position to analyze the feelings that must arise in our sensibility when we add the moral law, as an incentive, to the ground of our faculty of desire.

5.3 Heidegger’s Mistake

Why does Kant think we have access to ourselves in the mode of feeling that is distinct from that of apperception and intuition? I want to answer this by turning, briefly, to Heidegger. On first glance, my interpretation resembles Heidegger’s from his *Basic Problems in Phenomenology*, originally a lecture delivered in the summer of 1927. To his credit, Heidegger is one of the few readers to understand Kant’s theory of moral sensibility in terms of the *access* we have to ourselves as agents, and the account of feeling Heidegger develops is, I believe, an illuminating starting point. As we will see, though, Heidegger’s mistake is that he confuses how the moral law becomes accessible to us objectively (through the “fact of reason”) and how it becomes accessible to us subjectively (through the “feeling of respect”). The mistake is nevertheless instructive, as we will see.

As a preliminary, let me begin by saying a bit more about Kant’s conception of desire. When Kant says that life is “the faculty of a being to act in accordance with laws of the faculty of desire,” he means that living beings act in accordance with representations of objects or actions that promote their self-activity. *Life* is the principle of that self-activity (LM 28:247). So as living beings we seek to be the cause of objects that, in our view, promote life: we want to make them real through our activity. This is, for Kant, the essence of desire. It is an activity of striving, with the aim of realizing those objects agreeable to our constitution. What Kant means by “pleasure” is this relation of agreement.¹⁹ As living beings, then, we seek to make our relations of agreement, our pleasures, real. Starting with a simple example, we can say an apple agrees with me because food promotes physical life. This grounds a representation of pleasure. When I seek to be the cause of this representation, to make it real through my activity, then I have a desire: namely, to eat an apple. The apple must be an object of experience, I must represent it as an object of sensible intuition; but my feeling of pleasure for an apple pertains to me subjectively. “Nothing at all in the object is designated,” Kant says; in feeling “the subject feels itself as it is affected by the representation” (KU 5:204).

¹⁹Compare Baumgartens’ (1739) *Metaphysik* (“Der vierte Abschnitt,” especially §396)

We can then say feeling has a reflexive structure: my feeling is both *for* something, x , but also *of* something, my-relation-to- x . In every feeling I am disclosed to myself in a relation of pleasure or displeasure. To feel pleasure in the representation of x is to be conscious of my agreement with it: it is a self-relation or a relation-to-my-constitution. This is why we must distinguish feelings from “sensations” and “intuitions.” Sensations refer to the matter of an object; intuitions to the object’s (spatio-temporal) form. Feelings are representations that pertain only to the subject, and nothing about them can enter into my cognition of an object (MS 6:212n). On this point, then, Heidegger is right:

What is phenomenologically decisive in the phenomenon of feeling is that it directly uncovers and makes accessible that which is felt, and it does this not, to be sure, in the manner of intuition [or sensation] but in the sense of a directly having-of-oneself. Both moments of the structure of feeling must be kept in mind: feeling as feeling-for and simultaneously the self-feeling of this having-feeling-for. (p. 133)

Feeling, then, is always self-feeling.

This definition is formal (or “transcendental,” as Kant puts it) because it leaves the source of x unspecified. X may be empirical, like my representation of an apple, or it may be intellectual, like my representation of the moral law. Defined in this way, we do not prematurely settle the question of whether pure reason is subjectively practical, i.e., whether the moral law can serve as an incentive for us, or whether we can only take an interest in empirical objects. What’s important, though, is that the feeling-relation is still sensible in structure. All feelings pertain to the “subjective aspect of our representations in general” (MS 6:212n), which is sensibility. But that is not to say sensibility is the only *source* of feelings. The source of x may be non-sensible, but my relation to x can only *occur* in sensibility. For Kant, this is what opens room for thinking about moral feelings. My representation of the law has its origin in pure reason, in my consciousness of the law’s authority, but the subjective aspect of this representation—that is, its relation to me—occurs in sensibility. The subjective aspect of this relation is what Kant calls “respect.”²⁰ There are both positive and negative aspects of this relation, as we will see in the next chapter.

Heidegger points out that respect for the law does not arise after moral action has occurred, but is a condition of its possibility. I think he is right about this: “Respect for the law . . . is the way in which the law first becomes accessible to me as a law” and “the only way in which the moral law as such is able to approach me” (p. 135). But Heidegger goes on to claim

²⁰This explains Kant’s apparently contradictory remark that respect is non-sensible but “presupposes sensibility” (KpV 5:76), a point I will return to in section 5.4.

that through respect the law becomes accessible to me as “pure reason,” a “free, self-determining being” (p. 135). How could this be? How could feeling give me access to myself as an intelligible being? For Kant, our practical cognition of the law provides grounds for knowing ourselves as self-determining beings. It is the thought of *duty*, then, which first leads us to think of ourselves as free from the influence of inclination. This is part of the objective determination of the will Kant ties to the “fact of reason.” But we are now concerned with the subjective determination of the will and the access we have to our sensibility through the feeling of respect. Heidegger was onto something deep, then, but I think he should have said that respect for the law is the way in which the law first becomes accessible to me as the subjective determining ground of the will, that is, as an incentive.

5.4 Hegel’s Challenge

In defending this reading of Ch.III, I find myself up against a challenge raised by Kant’s immediate successors. Hegel (1807) and Schiller (1793), for example, were aware that Kant needs to explain the rational force of the moral law at the level of human sensibility, and they understood why Kant requires moral feeling in the manner I am proposing. In their view, the concept of respect arises out of a need to explain the connection between reason and sensibility, which Kant’s ethical project makes explicit. But it is a need which we are unable to explain on Kantian premises.²¹ As Dieter Henrich puts it, Kant needs a theory of moral sensibility, but his moral philosophy prevents him from offering one. This is what I will call, for short, *Hegel’s Challenge*. By way of introduction, it will be helpful to look at Henrich’s (1963) formulation of it.

Henrich says that Kant has to assign the positive and negative moments of respect to separate faculties of the human soul. He maintains that Kant can only establish a negative connection between consciousness of the law and sensibility (through the feeling of humiliation or restraint). Second, Henrich argues that the humiliation or constraint of one’s sensibility runs counter to the original movement of the moral law, which is to provide us with unconditional reasons for acting. In effect, Kant is unable to give a positive account of how the rationality of the law manifests in the world of sense. He can only represent that manifestation in terms of an external authority, Henrich argues, making the law no different—at the phenomenal

²¹As Henrich puts it: “One must distinguish between what Kant claimed to achieve with his talk of respect for the law and what his theory is really able to achieve. If one does this, it becomes evident that here too his theory cannot explain the intention of the moral consciousness” (1963, p. 109). See Henrich (1963). Citations will come from *The Unity of Reason: Essays on Kant’s Philosophy*, ed. Richard Velkley (1994).

level—from the coercive forces of legal or religious power.

This is how Henrich presents the problem:

The limitation of the inclinations occurs in “sensibility” as an incomprehensible effect of the intelligible freedom of the moral being. One cannot, however, say of sensibility itself that it is “elevated.” For it is not possible to see on the basis of Kantian presuppositions how sensibility can acquire a positive connection to reason solely because it claims were rebuffed. For this reason Kant understands this elevation merely as a relation of practical reason to itself. The removal of the obstacles to practical reason promotes its activity. Such promotion is approved by practical reason and equated with an increase in its energy. This takes place in a “judgment” of reason, thus wholly “on the intellectual side” of moral life. (p. 110)

On Henrich’s reading, only the feeling of humiliation before the law manifests itself in sensibility. So our practical cognition of the law influences our sensibility only negatively: our inclinations are held “in check,” and our pretensions to self-conceit are “struck down.” This is why Kant is unable to explain the phenomenon of moral consciousness. He is unable to explain the unity of the moment of humiliation, which occurs on the side of sensibility, and the moment of elevation, which occurs on the side of rationality. We are only elevated as rational beings once our sensible impulses are removed from the pathways of practical reason. We are elevated because we are free from the opposing weight that such impulses put on the will. But this means the moment of respect has no positive relation to our capacity to feel, and so no positive relation to us as sensibly affected beings.

It is not difficult to see where Kant’s successors get this reading. It is a misreading, I believe, but one for which Kant is to be held responsible. At (KpV 5:79) Kant uses the metaphor of a balance-beam scale to illustrate the law’s effects on the faculty of desire, and there he speaks of a “sensible side” and an “intellectual side” in the exact manner Henrich outlines. As Kant puts it: “the lowering of pretensions to moral self-esteem, that is, humiliation on the sensible side—is an elevation of the moral, that is, practical, esteem for the law itself on the intellectual side; in a word, it is respect for the law, and so also a feeling that is positive in its intellectual cause, which is known a priori.” This implies we only have a feeling caused by our awareness of the law (a “moral feeling”) when our inclinations are thwarted. Kant suggests as much earlier—at (KpV 5:73)—when he writes that “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 a feeling. 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.” It seems that we only have a negative feeling

for the law at the level of our sensibility. And that is Henrich's point: the moment of elevation bears no positive relation to our inclinations; it cannot be called a "feeling" at all. The positive moment of our relation to the law is a relation of reason to itself.

One problem with Henrich's reading is that he does not say what a "positive" connection between reason and sensibility would amount to. Remember that Kant's aim in Ch.III is to provide an account of how practical reason can pose as an incentive for us, i.e., how it can enter into the same motivational pathways that empirical incentives do. The account assumes that practical reason can appear to us in this way, as a subjective ground of action. Kant does not think we can comprehend this. Yet he does think we can comprehend the effects that must arise in our faculty of desire as a result of this subjective determination. This is why I am suggesting that Kant is re-framing the Question of Motivation from a causal perspective ("How can morality immediately determine the will?") to a phenomenological perspective ("What effects must this determination have on our sensibility?"). The question at stake is how morality can get a foothold in our sensibility—how it can appear to us as an incentive for action, and how that incentive can be incorporated into our maxims.

So there are some basic points left out of Henrich's account. First, Kant says that respect for the law is not the incentive *to* morality, which suggests a prior feeling for the law, but is rather "morality itself subjectively considered as an incentive" (KpV 5:76). By this, he means to say that we do not have a capacity for moral feeling that attunes us to the law, a so-called "moral sense" that would determine our judgements of the good, say, according to an occurrent feeling of pleasure, or delight, or satisfaction we have for virtuous actions. But Kant is also denying that what we call "respect" is a pre-existing sensible feeling that is somehow called to morality's allegiance. We may have proto-moral feelings like sympathy, benevolence, and honour that can be effective for promoting morality once we bring our maxims into conformity with the law (this is Kant's understanding of "rational self-love"); but this cannot be how the moral incentive is established. The moral incentive cannot have its basis, its determining ground, in anything sensible. And that is why Kant says that "respect" is just the moral law itself, i.e., the moral law in the guise of an incentive.

Secondly, and more importantly, Kant is clear that the *effect* of the moral incentive on us is an effect *on our sensibility*. This explains his apparently contradictory remark that sensibility is "the *condition* of that feeling we call respect" (KpV 5:75; my emphasis). A non-sensible being (Kant mentions the "supreme being" but we could also add a "holy will") would not *feel respect* for the law. "With regard to this it should be noted that, since respect is an effect on feeling and hence on the sensibility of a rational being, it pre-

supposes this sensibility and so too the finitude of such beings on whom the moral law imposes respect” (KpV 5:76). The way morality effects our sensibility, then, will be no different than the way empirical incentives effect us, and this is why Kant’s transcendental definition of feeling is so important. We can define feeling in such a way that we clear enough room for thinking about a practical determination of our faculty of desire.²²

5.5 The “Harmony” of Duty and Inclination

With these points in mind, I think we can isolate a difference between Kant and his immediate successors. For Kant, the “harmony” of duty and inclination involves the feeling of pleasure we have for our rational nature, what Kant calls the feeling of “elevation” (*Erhebung*). It does not involve an “elevation of sensibility” if that means the inclinations come to possess deliberative authority. Yet this is what Kant’s successors seemed to want. For Schiller, someone’s inclinations are in “harmony” with duty when they replace any need to test his or her actions against principles of reason. The Beautiful Soul, for example, is someone with pro-moral impulses, someone who spontaneously desires to do what morality requires of him. “One refers to a beautiful soul,” Schiller writes, “when the ethical sense has at last so taken control of all a person’s feelings that it can leave affect to guide the will without hesitation and is never in danger of standing in contradiction of its decisions” (1793, p. 152).

Curiously, Schiller thought his departure from Kant was one of emphasis. He thought Kant’s stress on the imperative form of morality was a burden put on him by the “lax character” of the times, that Kant put on the role of a Draco “because he did not regard his time as worthy of a *Solon*, or yet able to receive one” (1793, p. 151). Kant was also unwilling to recognize a deep rift between his position and Schiller’s. In the *Religion* he mentions Schiller’s disapproval of representing morality in terms of constraint and necessitation, but he thinks this is superficial: “we are however at one upon the most important principles” (Rel 6:22n). What we are to make of this?

My view is that Kant and Schiller failed to clarify the very issue under discussion. There are at least two questions at stake here. First, there is a question about the *cultivation* of moral feeling. Schiller seems to be saying that we should cultivate our sensibility so that our inclinations come to possess pro-moral content. Kant denies that the inclinations are capable of this role. The matter of inclinations for Kant always comes from our sensibility: this can be *in accord* with morality, but not moral *in content*. That being

²²Broadie and Pybus (1975) argue—correctly, I believe—that Kant is providing a synthesis of sentimental and rationalist theories of moral motivation, and that the incentive to morality is, for Kant, “rationality embodied in feeling” (p. 63).

said, Kant does believe we are able to cultivate our capacity to respond to the thought of duty, that through education we can make the satisfaction we feel in maintaining our dignity *stronger* than any counter-moral incentive (KpV 5:38). But this does not involve the cultivation of a distinct sentiment that somehow *replaces* practical reason as a determining ground of the will.

Kant and Schiller may have been of “one mind” about a second question, of what *constitution* is best suited for promoting the intensity of moral feeling. Schiller asks: “And now the question arises as to what kind of *personal constitution* allows greater freedom to the sensuous instruments of the will”? His answer: “When the mind expresses itself in the sensuous nature that depends on it in such a way that nature faithfully carries out the will of the mind and expresses its sentiments clear, without contravening the demands that the senses make upon them as upon appearances, then there will arise what we call grace . . . Only with grace associated with it can it establish beyond doubt whether what we take to be control is not actually a dullness (hardening) of sensibility” (pp. 147-163). In rejoinder, Kant writes: “Now, if we ask, ‘What *aesthetic* constitution, the temperament so to speak of *virtue*: is it courageous and hence *joyous*, or weighed down by fear and dejected?’ an answer is hardly necessary. The latter slavish frame of mind can never be found without an hidden *hatred* of the law, whereas a heart joyous in the *compliance* with its duty (not just complacency in the *recognition* of it) is the sign of genuineness in virtuous disposition” (Rel 6:22n).

Even here, though, Kant and Schiller may be saying different things. The “joyous frame of mind” Kant speaks of refers to the self-contentment we feel in bringing our maxims into conformity with the law, a feeling he thinks we can make more powerful through habitual practice. This kind of temperament may be different than what Schiller means by an individual whose actions are attended by “gracefulness,” since gracefulness involves a natural outpouring of feeling attuned to—but not restrained by—dutifulness. It seems, then, that Schiller was operating under a different, and perhaps uncritical, conception of “moral sensibility”; that he thought of the harmonization of duty and inclinations in terms of an “elevated sensibility”; and that his goal of moral cultivation was to acquire pro-moral impulses and inclinations. As a result, the problem he has with Kant is rather superficial, since he thinks Kant’s emphasis on the imperative form of morality leaves the demands of taste unsatisfied, that our natural approval of beauty (especially the beauty Schiller associates with virtuous action) will be neglected. Kant’s draconian approach to morals is seen as unjustified on aesthetic grounds. But this falls short of a *skeptical* problem. At no point does Schiller raise the more troubling question of whether our sensibility is fit for morality. Schiller is worried that Kantian morality may leave us with an

uncultivated connection between reason and sensibility, that its emphasis on duty may frighten the senses and leave our aesthetic demand for beauty unsatisfied. The skeptical threat is that Kantian morality may leave us *without* a connection all-together, that we may be incapable of responding to the moral “ought,” and the morality itself may lack access to our faculty of desire.

This is of course Hegel’s Challenge. Hegel takes issue with the apparently contradictory claims I cited above: that respect is just morality “subjectively considered,” but that it also “presupposes” human sensibility.²³ As Hegel sees it, there is a dissemblance at work here. On the one hand, Kant is saying that the aims of morality are pure, and that action motivated by the thought of duty must not have any admixture with the sensible. But then he claims moral action is only actualized as a feeling, and that the feeling we call “respect” is an effect on our sensibility. (So sensibility becomes what Hegel calls the “organ” for the realization of pure practical reason: the “middle term between pure consciousness and actuality.”) As a result, moral consciousness must somehow bring itself into conformity with sensibility, and it does this by “giving itself the shape of an *impulse* [*Triebes*], which is to say, it is immediately the present harmony of impulse and morality” (§622). But there is a problem here. Our inclinations are part of the world of sense, which has its own “laws and springs of action” (*Gesetze und Springfedern*), so the shape of moral incentive is *empty*. It is not really an impulse at all, since it does not figure into the driving mechanism of human sensibility.²⁴

For lack of space, I am unable to offer a full discussion of Hegel’s criticism. But there are two points I want to raise before closing. First, Hegel thinks that moral consciousness on Kant’s account *eliminates* the “purposes of sensibility” (*Zwecke der Sinnlichkeit*), along with the direction and driving mechanism of our inclinations. But Kant’s claim is that only an excessive kind of self-love must be “struck down” by the moral law, the kind of self-love he thinks is at the root of our desire to elevate the demands of sensibility, to make our wants and wishes *law-giving*. The inclinations are not inherently counter-moral for Kant, and so the fact that we have them does

²³“Moral self-consciousness puts forward the view that its purpose is pure purpose, that is, that is independent of inclinations and impulses in such a way that the pure purpose has eliminated within itself the purposes of sensibility.—Yet this proposed sublation of the sensuous essence is once again made into a matter of dissemblance. Moral consciousness acts, that is, it brings its purpose into actuality” (1807, §622).

²⁴“But impulse is not in fact merely this empty shape which could have within it a spring of action other than the one it is, and be impelled by it. For sense-nature is one which contains within itself its own laws and springs of action; consequently, morality cannot therefore be in earnest about being itself the mainspring of the impulses [*Triebfeder der Triebe*], the angle of inclination for inclinations [*der Neigungswinkel der Neigungen*] ... The harmony of the two is thus merely *implicit*, merely *postulated*” (§622).

not reveal a necessary connection between the moral law and a painful feeling of humiliation. Kant even makes fun of the Stoics for mistaking their enemy, “who is not to be sought in the natural inclinations, which merely lack discipline and openly display themselves unconcealed to everyone’s consciousness, but is rather as it were an invisible enemy, one who hides behind reason and hence all the more dangerous” (Rel 6:57).

Secondly, Hegel seems to buy into Schiller’s expectation that the harmony between duty and inclination is one in which our sensibility somehow becomes pro-moral in its driving mechanism. The moral incentive is an “empty form,” on Hegel’s view, because it does not replace this mechanism, or because it does not let human sensibility become the mainspring of moral action. And so he concludes that, instead of a positive harmony between duty and inclination, Kantian morality leaves us with “merely *implicit*” harmony, “merely *postulated*” (§622). From Kant’s perspective, however, it is hard to see this as a criticism, since Kant denies that we can comprehend how the moral law serves as the mainspring or incentive for the will (which is why his account in Ch.III only concerns the “effects” of the moral law on our sensibility, not its determining ground). In this respect, Hegel is right: Kant *is* postulating the harmony between moral consciousness and sensibility, but that is only because Kant denies we can have insight into the causal relation between the two. What we *can* have insight into, for Kant, is the effects the moral incentive must have on us.

My aim in the next chapter is to explain what these effects are.

Humiliation, Elevation

6.1 Preliminaries

HERE it would be a good idea to clarify what I mean by Kant’s “moral phenomenology,” as this is potentially misleading. To begin with, we can distinguish between the “discipline” of phenomenology and its “method.” As David Woodruff Smith (2008) puts it: the former refers loosely to the “study of structures of experience or consciousness,” and the latter, “the historical movement of phenomenology,” which is “the philosophical tradition launched in the first half of the twentieth century by Edmund Husserl, Martin Heidegger, Maurice Merleau-Ponty, Jean-Paul Sartre, *et al.*” When commentators speak of Kant’s moral phenomenology in the second *Critique*,¹ they are referring to the *method* of phenomenology (although Heidegger clearly wanted us to think of Kant as a *proto-phenomenologist* in the history of philosophy, whose shortcomings Heidegger wants to overcome).

Within the past few years there has been a resurgence of interest in phenomenological arguments in ethical theory. A common starting point in this growing literature is that phenomenology concerns the what-it-is-like features of concrete moral experience,² features directly accessible through introspection.³ One question that arises here is whether robust phenomenological claims can support an ethical theory or reveal its superiority over others—for example, whether the what-it-is-like features of moral deliberation are better captured by sentimentalism, or egoism, or rationalism. Fran-

¹Heidegger (1927); Beck (1960); Allison (1990); Stratton-Lake (2000); Grenberg (2009).

²Horgan and Timmons (2005; 2008); Gill (2007; 2009).

³Mandelbaum (1955).

cis Hutcheson develops an argument along these lines, without of course using this terminology. He argued that we see through introspection that every moral judgment has an accompanying phenomenal state—*pleasure* with a judgment of approval, *pain* with a judgment of disapproval. These are introspectively accessible “facts” of moral experience that philosophers like Clarke and Balguy are unable to account for—because of their commitments to rationalism. As a result, their theory leaves an important aspect of moral life unexplained; or so Hutcheson claimed.

Kant never used the term “phenomenology” in this sense, yet commentators believe he developed arguments like Hutcheson’s. A common worry is that the *Groundwork* left us with a distorted picture of moral agency—as if the virtuous person is someone who consciously tests each of her maxims against the categorical imperative, who reflects on abstract moral principles *before* acting. This doesn’t match our moral experience, let alone the image we have of a virtuous person. As Philip Stratton-Lake puts it: “Nobody deliberates from anything like the moral law to particular moral laws and then to some specific moral verdict in their everyday deliberation. In concrete situations, say, where someone is in need and I can help him, or where I have made a promise, or where I could show gratitude, I neither need, nor typically do deliberate from the moral law to the particular moral verdict” (2000, p. 127).

Many writers on Kant turn to Ch.III of the *Critique of Practical Reason* for a solution, and there we seem to find all the phenomenological details missing from the *Groundwork* (details that explain, for example, the painful self-reproach we feel in breaking a promise, or the pleasurable self-approval we feel in helping someone in need, etc). Part of what makes Kant’s moral phenomenology attractive, on this reading, is that it can be detached from his broader meta-ethical project: the project of securing our status as transcendently free agents and explaining why we experience morality as an imperative. We can *keep* the phenomenological richness, so the thought goes, and *abandon* the metaphysics.

I am edging close to a big question, whether or not we can reconcile Kant’s moral phenomenology with a more naturalistic account of human agency. I want to back away from this—not because I think it unworthy of discussion, but because my concern here is more specific. My view is that the interpretation of Kant’s moral phenomenology I have outlined is mistaken, and that what Kant is up to in Ch.III does not resemble anything like a phenomenological argument in the sense used by recent authors. For one thing, Kant does not think his account of moral feeling follows by observing the occurrent phenomenal states of moral experience; that would require introspection, and Kant is pessimistic about the kind of knowledge we can attain of ourselves through self-observation. But if Kant’s method is

not introspective, what is it? In the second *Critique*, he gives us the following, if obscure, clue. He says: “nothing further remains than to determine carefully in what way the moral law becomes an incentive and, inasmuch as it is, what happens to the human faculty of desire as an effect of that determining ground upon it.” And later: “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” (KpV 5:72; my emphasis).

What does he mean by this? When Hutcheson claims to discover a distinct feeling of pleasure attending his positive moral judgments, he is offering us a description. He is telling us how the facts of his moral experience have revealed themselves through introspection. Kant is saying something different, however. He is saying we need to determine what the activity of pure practical reason *must* feel like, given what we know a priori about finite rational beings like us. He does not think we need to “look within” and describe the features of moral experience as they arise in concrete situations. Rather, he thinks we need to discover the *necessary connections* between the moral law, as an incentive, and our sensibility.⁴

In theory, then, there should be no variability in how these connections manifest from one agent to the next and across situations. For example, Kant argues that the first effect of the moral incentive on our sensibility will be negative, a painful feeling of humiliation. We know this a priori because, according to Kant, we know that beings like us have a propensity to resist the moral law. This is not an introspectively accessible feature of experience, but rather a claim about our propensity to turn the principle of happiness into an objective determining ground of the will, which Kant thinks is an ineliminable feature of beings who, like us, have a natural concern for their happiness. Our propensity to self-conceit makes us *necessarily resistant* to the claims of pure practical reason, and so we can say the first effect of the moral law on our faculty of desire will be pain-like. By calling Kant’s moral phenomenology “a priori,” then, I want to suggest that he is concerned with what-it-*must*-be-like for sensible rational beings to experience morality.

It may seem that in stressing the a priori character of Kant’s method in Ch.III, I have ridden it of anything phenomenological. Speaking of Kant’s moral phenomenology may now seem redundant, or misleading. But I think there are good reasons to continue using this label, even though it takes on an idiosyncratic meaning in the context of Kant’s ethics. One reason is that Kant likens his general strategy of argument in the second *Critique* to a chemistry procedure.⁵ He says we can set up an experiment within

⁴This point is brought out clearly by Herman (2006, p. 20). Also, see Beck (1960).

⁵“We have at hand examples of reason judging morally. We can analyze them into their elementary concepts and, in default of *mathematics*, adopt a procedure similar to that of

every human practical reason that displays the priority we give, in thought, to the moral law over self-love. Kant says the philosopher has this advantage, that he can elicit this response from a common person, and so display (without explaining) the act of giving preference to the pure component of practical reason. “[A]lmost like a chemist,” Kant writes, “he can at any time set up an experiment with every human practical reason in order to distinguish the moral (pure) determining ground from the empirical, namely, by adding the moral law (as a determining ground) to the empirically affected will” (KpV 5:92). In this way, Kant believes we can elicit the *immediate preference* reason gives to its own form in place of any further condition that would justify its legislative standing.⁶

I want to suggest we can extend Kant’s chemistry metaphor to the account of moral feeling from Ch.III. By setting up a thought-experiment within common practical reason, we display the immediate preference we give to the moral law (for example, in judging that we should tell the truth). By doing so, we can determine the *effects* this act must have on a being also affected by non-rational incentives. We can say, for example, that someone confronted with a judgment of what she “ought” to do (“don’t lie!”) will feel humiliated in comparing this judgment to her propensity to make self-love law-giving (that she will feel pained by viewing her desire to lie for self-interest under the lights of pure practical reason). While this account does not rest on introspective claims, it is still meant to clarify the a priori features that precede and condition our everyday, concrete moral encounters.

With these points in mind, we are in a better position to look at Ch.III more closely. In what follows, I will argue that Kant’s account consists of a dialectic between our intellectual recognition of the moral law (expressed in the “fact of reason”) and the feelings that arise from this recognition: humiliation, reverence, and elevation. The *feeling of humiliation* arises when we compare the moral law to our propensity of self-conceit, which Kant defines as a tendency to elevate our natural desire for happiness into an unconditional practical law. The *feeling of reverence* pertains to our sense of awe before the categorical imperative as a causal law distinct from, and

chemistry—the *separation*, by repeated experiments on common human understanding, of the empirical from the rational may be found in them—and come to know both of them *pure* and what each can accomplish of itself” (KpV 5:163).

⁶“When an analyst adds alkali to a solution of calcareous earth in hydrochloric acid, the acid at once releases the lime and unites with the alkali, and the lime is precipitated. In just the same way, if a man who is otherwise honest (or who just this once puts himself only in thought in place of an honest man) is confronted with the moral law in which he cognizes the worthlessness of a liar, his practical reason (in its judgment of what he ought to do) at once abandons the advantage, unites with what maintains in him respect for his own person (truthfulness)” (KpV 5:92-93).

superior to, the causal law of our inclinations. Finally, the *feeling of elevation* pertains to our discovery that the source of the moral law is, in fact, internal to our will.

6.2 The Dialectic of Respect

A RECOGNITION

Kant presents two thought-experiments early on in the second *Critique*. In the first he asks us to imagine an object of desire we could not think of resisting if it were placed right before us. He then asks how we would act if satisfying this desire would lead to our immediate execution. “Suppose,” Kant writes, “someone asserts of his lustful inclinations 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” (KpV 5:30). Our answer, of course, is that we would restrain ourselves. The desire for any specific object could never outweigh the possibility of losing our life. And that is because death would cancel out the basis of that desire, our sensible nature, and its underlying principle, happiness. Kant’s point is that we cannot rationally act on a desire that would lead to our own death, for that would amount to pursuing an object whose satisfaction would destroy the principle on which we act: the principle of satisfying our desires.

Kant then asks us to imagine how we would act if we were asked to give false testimony, also on pain of execution. He outlines the scenario of a corrupt prince who wants to bring false charges against an innocent man, and the prince has called on us to support these charges. Death awaits if we do not submit. Can we imagine refusing him? Kant does not expect us to answer as confidently as before. When we put ourselves in this situation, he does not expect us to deny the prince’s request. It would not be a sign of moral failure if we admitted that, after considering the matter seriously, we would likely give in under the threat of execution. Kant would understand that. For his thought-experiment is not meant to elicit moral heroism. His aim is more modest. Kant only wants to demonstrate that it is possible we could imagine doing the right thing, that we could deny the prince, knowing it would lead to our own death. That is all he needs from us: the recognition that we can do it because we judge it is right, that we “ought” to do it. As long as the reader can make this concession, Kant’s second thought-experiment works.

What does it prove? Not that we will act on our moral appraisals, or at least not a guarantee that we will so act. It proves, rather, that we can endorse the moral “ought” as the principle of our own will. The moral “ought”

may never manifest in our physical deeds, and Kant is fine with that. Confronted with the option of denying the corrupt prince we may wish we could be disposed with the requisite strength of will to put our denial into effect. But the determination of our will does not require this. When we give preference to moral considerations we are already exercising our will in its legislative capacity, whether or not we act that way. Re-ordering our deliberative preferences in this way is already a “deed” of reason: it is an act of subordinating the principle of self-love to the moral law. Kant’s thought-experiment is meant to capture this activity within the reader, which is why he does not require us to answer his question whole-heartedly. He simply wants us to engage in a process of deliberation without consulting self-love. If we can imagine doing the right thing even if it means ending our life, we display a capacity to drop all self-interests in favour of an unconditional practical law. In which case we display our interest in something other than life, “something in comparison and contrast with which life and all its agreeableness has no worth at all” (KpV 5:88).

B THE FEELING OF HUMILIATION

We have seen that, for Kant, moral authority is irreducible to our act of taking the law as a sufficient ground of action. This is more a “deed” than a “fact” we somehow register. In this regard, the “deed of reason” involves that I give preference to the law. And to do this I must, at the same time, restrain my own private interests. That is why the first effect of the moral law on my sensibility will be negative. In giving priority to the law, I must subordinate my personal commitments. This is not a case of rejecting my desires or inclinations. It is rather a case of shifting my preferences.⁷

But this is why we have a problem. For as a human being, I have a propensity to care for myself. The first object of self-care will be my needs as a naturally situated and social being (needs for food, shelter, community); and soon after that I will develop my own personal desires and tastes. The act of making myself responsive to the claims of morality will then have a negative influence on this tendency: I will “feel” the moral law as restraining the regard I have for myself. But it will also affect another propensity I have as a sensibly affected being, a propensity to regard the principle of my happiness as if it were a supreme practical law. This is what Kant calls “self-conceit” (*Eigendünkel*) and he is clear in distinguishing it from “self-love” (*Eigenliebe*). Self-conceit is the attempt to give deliberative preference

⁷As a “propensity,” self-conceit is what Kant calls a “ground of inclinations,” not an inclination itself. Kant thinks that the inclinations we have in virtue of our biological and social natures are predisposed to the good; they become vicious only when we rank sensibility above the moral law. Moreover, only self-conceit can be properly eliminated as a deliberative attitude because it is, for Kant, a delusion.

to one's own wants and wishes, to place self-love above morality. It is the tendency to treat one's sensible nature as if it were a source of unconditional authority, something which Kant thinks only the moral law can have.⁸ In the first instance, then, my act of giving priority to the law will give rise to a negative feeling: it will humble me.

The example Kant gives of Voltaire is meant to illustrate this: "when the common run of admirers believes it has somehow learned the badness of character of such a man (such as Voltaire) it gives up all respect for him, whereas the true scholar still feels it at least with regard to his talents, because he is himself engaged in a business and a calling that makes emulating [*Nachahmung*] such a man to some extent a law for him" (KpV 5:78; modified). In his intellectual achievements, Voltaire instantiates a law for anyone committed to the vocation of a scholar. Common admirers may lose esteem for him when they learn the details of his personal life; but the scholar continues to respect him. Why? Because the scholar recognizes in Voltaire's achievements a norm that applies to him. As someone who fulfills this norm, Voltaire confronts the scholar with an ideal image of his own self, with the image of *how he should be as a scholar*. In comparison, then, the scholar sees how much he falls short of his vocation. He feels humbled.

To sum up: One aspect of humiliation is the restraining of the propensity to care for oneself, a propensity Kant observes is "natural and active in us" before the constitution of a moral will. Self-care, then, need only be held "in check." But self-conceit must be struck down and eliminated, because it is a tendency to blind oneself to moral considerations. My representation of the law's authority elicits a feeling analogous to pain, because when I acknowledge the law as a source of unconditional reasons for action I recognize the conditional value of my inclinations. "Hence the moral law unavoidably humiliates every human being when he compares with it the sensible propensity of his nature" (KpV 5:74). Because self-care is a natural predisposition of the will to tend to its own interests, it needn't be eliminated, and any attempt at its elimination would border on enthusiasm, which Kant believes the Stoics were guilty of. But unlike the natural tendency to care for oneself, self-conceit only furnishes the appearance of authority, for sensible desires are in fact unfit to serve as laws of action.

This effect is not a contingent fact of our psychology, but a necessary consequence that Kant thinks we can determine a priori. Our propensity to self-conceit is inherently opposed to the claims of pure practical reason, and so the first effect of the moral law on our sensibility will involve a painful feeling of self-reproach when we compare the moral law to self-conceit. The

⁸As Kant writes in a later text: "A conviction of the greatness of one's moral worth, but only from failure to compare it with the law, can be called moral arrogance (*arrogantia moralis*)" (MS 6:435).

feeling of pain involves a representation of disagreement between and object and one's constitution (as we discussed in section 5.3). By definition, then, the representation of the moral law, as an incentive, will necessarily disagree with our propensity to make our wants and wishes law-giving. As we can see in Figure 6.1 below (p. 110), the feeling of humiliation has a reflexive structure. It is a feeling of "displeasure" in representing the law, because that representation discloses the limited authority of self-love and the illusory authority of self-conceit. This is what we might call the negative dimension of moral feeling.

There is a positive dimension, however. As I mentioned above, feeling humiliated before the moral law is only possible if one has recognized the law as a source of unconditional reasons for action. When I compare that authority to my sensible will, I feel pained by the realization that I must limit my impulses and inclinations.⁹ But in having turned my attention to the law in this manner, I have already adopted an appropriate attitude to the law. When I turn my attention to the moral law as a superior source of authority—that is, when I stop comparing it to my sensible will—I feel reverence. I recognize the law as a "positive and determining ground" distinct from the causal ground of my desires (KpV 5:74).

C THE FEELING OF REVERENCE

Kant illustrates this shift with the example of a humble man of virtuous character (KpV 5:77). Our initial reaction to the humble man will be similar to pain: in representing his honesty, his benevolence, his integrity of character, etc., we cannot avoid feeling displeased by seeing how much we fall short of the standard he sets before us. We feel humiliated because we are sensitive to our own lack of honesty, benevolence, and integrity when we compare ourselves to him. The "feeling" is reflexive in this way. Our representation of the humble man is, at the same time, a self-representation. We are placing our moral conduct (or lack of) in light of the example he sets before us; and the disagreement between the two effects a pain-like feeling. Kant quotes Fontanelle's saying that "*I bow before an eminent man, but my spirit does not bow,*" to which Kant adds, "before a humble common man in whom I perceive uprightness of character in a higher degree than I am aware of in myself *my spirit bows*, whether I want it or whether I do not" (KpV 5:77). Now, humiliation may also be mixed with positive feelings, like those of admiration. When we see how well the humble man handles himself—how he is self-effacing but not self-deprecating, sensitive to others but not indulgent, honest but not cruel in his honesty—we cannot help

⁹"He who has lost at play," Kant tells us, "can indeed be chagrined with himself and his imprudence; but if he is conscious of having cheated at play (although he has gained by it), he must despise himself as soon as he compares himself with the moral law" (KpV 5:37).

but hold him in high regard. In experience, we might admire the humble man one moment and feel cast down and humiliated the next. Kant's point, though, is that these two feelings are modes of a single judgment. In having judged that the humble man represents a law applicable to us, we either feel displeasure in comparing ourselves to him or admiration in witnessing his conduct and the degree of moral perfection his conduct makes visible to us.

In the second instance, then, our representation of the moral law effects a positive feeling: it awakens our reverence. Whatever inhibits the pretensions of self-love will also effect a feeling of pleasure. The inhibiting of self-love will effect a feeling of esteem for the law as a positive causality. We revere that which humiliates our false self-esteem because as painful as humiliation is, it allows us to recognize a source of deliberative authority higher than our sensible nature. This is the point of the balance-beam metaphor (KpV 5:76). The demands of self-love and self-conceit pose a counter-weight to the demands of the moral law. By removing their claim to priority in the balance of reasons, then, the moral law effects a *shifting of the scales* in its favour. The demotion (or "lowering") of the demands of self-love and self-conceit effects a painful feeling, for we feel the activity of our sensible will under constraint; but that lowering also effects a positive feeling, for in one and the same movement it constitutes the promotion (or "raising-up") of the law's authority.

D THE FEELING OF ELEVATION

This is how we identify with the humble man, how we recognize that he presents us with a law, in the same way that the scholar identifies with Voltaire and recognizes, in his accomplishments, the fulfillment of a norm applicable to him. In recognizing we should be like the humble man, we put ourselves under obligation: We hold ourselves to the claim that we should be virtuous—*like him*. The feeling of humiliation arises first because our initial point of comparison with the law will be our sensible nature (and its propensity of self-conceit). But this feeling occasions a positive feeling, because in taking up the appropriate attitude to the law—in giving the law priority in the balance of reasons—we have already exercised our autonomy of will. This occasions a feeling analogous to pleasure because we see that in giving priority to the law we have acted in a manner exemplary of a virtuous person. We see the humble man reflected back, as it were, in ourselves. For the humble man not only represents a positive determining ground of action separate from the order of desires; more importantly, he represents a determining ground *that we recognize within us*.

The elevating dimension of respect consists in an individual's discovery that the demotion of self-love is part of the law-giving "of his *own* reason"

(KpV 5:81; Kant's emphasis). In relation to the humble man, I feel humiliated when I compare my conduct to the law his example sets before me, but I feel elevated when I see that my attitude toward him has already, in fact, performed the very attitude the law requires of me. I have, in casting reprobation on myself, displayed my immediate responsiveness to the law. When I recognize this, *self-reprobation* turns into *self-approval*. Because an individual's restraining of self-love "is exercised only by the law-giving of his own reason, it also contains something elevating [*Erhebung*], 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 reason" (KpV 5:80-81):

This completes Kant's a priori phenomenology. By (A) recognizing the unconditional authority of the moral law, we are realizing our autonomy of will through a "deed" of reason. The first effect of this realization is (B) humiliation (we feel the painful inhibition of our desire-based interests), the second is (C) reverence (through the inhibition of such interests we feel the promotion of the law's causality), the third is (D) elevation (through the law's causality we discover our will in its legislative capacity). These points are summarized in Figure 6.1.

As we can see, feeling_T is the key to understanding this model. The initial representation of the law serves as an object of self-comparison to one's sensible will: the lack of agreement between the two elicits displeasure_T. But the restraining of self-love at the same time allows us to revere the law as a causal power distinct from the order of our desires. Comparing this law to our rational will elicits pleasure_T because the two are in harmony. Elevation is brought about by the discovery of one's standing as an autonomous will.¹⁰ These modes of comparison are feelings_T for the law from our point of view. They are how the moral law, as a purely rational law, gains access to our sensibility.

I hope I am now justified in calling Kant's model "dialectical," as this term comes closest to describing the four stages I have outlined. The stages are dialectical because the transition from (A) recognition to (D) elevation marks out a process of conflict and eventual resolution between the claims of the moral law and the counter-claims of self-love. It is, moreover, a process of internalization because the transition involves a shift in

¹⁰Kant later says that the feeling of elevation grounds "true humility." As he writes: "True humility follows unavoidably from our sincere and exact comparison of ourselves with the moral law (its holiness and strictness). But from our capacity for internal lawgiving and from the (natural) human being's feeling himself compelled to revere the (moral) human being within his own person, at the same time there comes exaltation (*Erhebung*) of the highest self-esteem, the feeling of his inner worth (valor), in terms of which he is above any price, and possesses an inalienable dignity (*dignitas interna*), which instills in him respect for himself (*reverentia*)" (MS 6:436).

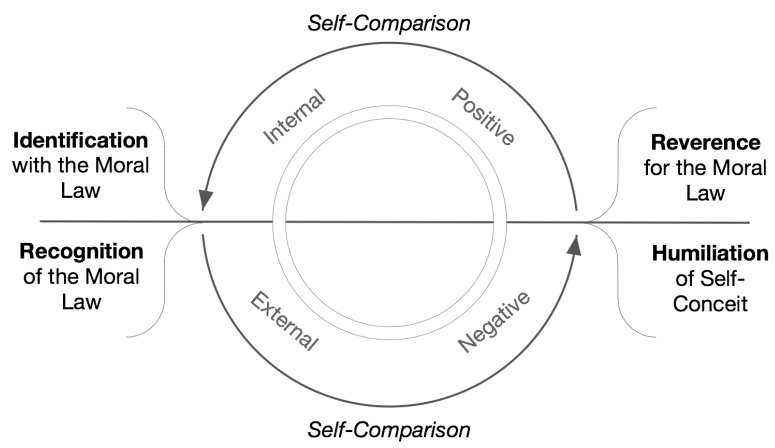


Fig 6.1: Kant's Dialectic of Respect

my consciousness of the moral law as an external imperative which humiliates my sensible will to an elevating consciousness of the law as internal to my will. In comparison to the humble man, I reproach myself for falling short of the norm he sets before me; but that reproach is an effect of my prior recognition of the norm's superiority in the space of reasons. I feel humiliated in my self-comparison because I have already—in an act of freedom—given deliberative preference to the moral law. In thus recognizing a superior ground of choice and action—superior to the order of my own inclinations—I feel reverence for the moral law; and in doing so I discover my capacity for moral action.

6.3 Completing the Experiment

We now have the resources to complete Kant's second thought-experiment. Imagine you could reject the prince's request, that you could sacrifice your love of life because you recognize you shouldn't give false testimony against an innocent man. Following Kant's account, we can now describe how you should experience this process, given your motivational psychology. In the first moment, by judging the wrongfulness of the prince's request, you have already acknowledged the authority of the moral law. In this way, you have already adopted the proper attitude to the law. And to this extent you have, in judging the wrongfulness of the request, adopted the standpoint of freedom. Why? Because you have shown you are capable of deliberating outside the order of your desires, and in doing so you have presupposed membership in a normative order. On a purely rational level, you have displayed your responsiveness to the law. All of this is contained in your answer, that it would be wrong to obey the prince.

This answer will have different effects on your sensibility. In the first case, its effect will be negative. You will feel pained by representing the disagreement between the moral law, whose authority you take to be unlimited, and your love of self, whose authority you are now forced to constrain. One aspect of self-love (the natural tendency to care for yourself) will be "held in check"; another aspect (the tendency to elevate the principle of self-love to the status of a law) will be "struck down." On the other hand, in judging the moral law as a superior principle of choice, you will feel reverence. You will recognize the law as a source of authority whose claim on you, as a rational agent, is greater than any impulse or inclination. This will in turn elicit a positive feeling. You can deny the prince because his threats of execution do not hold authority over you. They threaten your sensibility, your love of life. They do not threaten your capacity to act freely—your autonomy.

Kant believes we feel elevated in considering the example of another

dying for the sake of duty because we encounter the example dialectically. This is the space of moral education. We recognize in the other's exemplary deed a capacity within ourselves, *a capacity to act on principle against the resistance of self-love*. We feel elevated because in restraining the influence of self-love we are already making ourselves responsive to moral considerations. We are already facilitating our identity as moral beings, which appears to us in its outward perfection through the other's noble sacrifice. Although we have yet to bring this identity to perfection in our own actions, we have fulfilled our commitment to the law by exercising our freedom of will. Our feelings for morality make this commitment transparent to us.

6.4 Two Objections

At this point, two objections could be raised.

So far I have assumed that when you think about Kant's example honestly, you would get as far as conceding that you "could" deny the prince because you judge you "ought" to. But this might be asking too much for some people, and on reflection it seems outrageous if Kant's argument requires the reader to choose duty over life. One could object that the prince's threat would break the will of most people; that however courageous we feel in our imagination, a real life situation would defeat us. How might Kant answer this? To begin, he could ask you to imagine that your false testimony works: it leads to the death of an innocent man. He might then ask how this outcome sits with you. If you find you are trying to reassure or console yourself—"I had no choice," "I was forced into it"—he could point out you're still committed to judging the prince's request as wrong. And if you are still bothered by what you did, hypothetically, Kant could say you're struggling to avoid the fact that you were free in obeying the prince—and that you could have chosen otherwise. Cases of bad conscience, in Kant's view, only occur for those who were aware of acting on their own initiative. The very attempt to assuage such thoughts—"I had no choice," "I was forced into it"—only shows that you are, in Kant's sense, humiliated. So if the thought of succumbing to the prince's request gives rise to feelings of self-reprobatation or guilt, the thought-experiment still works. By casting a disapproving eye on your imaginary deed, you are giving weight to moral considerations over self-interested ones. That is all Kant's argument requires of you.

A further, and more pressing, objection could be directed at Kant's choice of examples. For he has seemingly failed to consider the possibility of sacrificing one's life on non-moral grounds. There are any number of counterexamples we could raise here. If I am able to deny my love of life for the sake of my country, for instance, it appears the second thought-experiment fails

to capture the special authority of the moral law. How might Kant respond if we cornered him here? Interestingly, Kant does work through a list of alternative examples in his brief discussion of moral education at the end of the second *Critique*, and there he argues such examples would fail to instill a proper incentive to morality. The contrast he sets up is specifically between actions represented as “noble and magnanimous” and actions represented “merely as duty” in relation to the moral law (KpV 5:158). He lists two of the former kind: first, a man who dies while swimming out to save the victims of a shipwreck, and second, a man who—out of patriotism—dies defending his country.

Kant says a young or uncultivated mind may consider such actions worthy of praise, but they will still fail to produce a genuine moral feeling. Why? In the first case, because the attempt to save the shipwrecked victims is meritorious but not necessary; in the second case, because the act of defending one’s country stands in conflict with other duties, such as self-preservation. Only actions that demand complete and unconditional compliance to the law provide grounds for knowing our freedom. And that is why Kant believes examples like refusing the prince allow us, in turn, to discover our capacity to deliberate without consulting self-love. We judge we “can” refuse the prince because we “ought” to, and that judgement is enough to realize our capacity to act out of respect for the moral law. For an example to serve as a model of emulation, then, it must perform the same disengagement from self-love that, for Kant, is the characteristic mark of freedom. That is why Kant believes the purity of an example only adds to its practical force; that an instance of one’s commitment to duty over life “apart from every view to advantage of any kind in this world or another and even under the greatest temptations of need or allurements . . . elevates the soul and awakens a wish to be able to act in like manner oneself” (G 4:411n). For the purposes of educating the child or pupil morally, then, only examples of moral sacrifice provide the conditions for a dialectical encounter.

6.5 Moral Enthusiasm

In Chapter 2 I introduced Kant’s motivation skeptic from *Groundwork* II, the skeptic who expresses a “deep regret” that we are noble enough creatures to recognize the authority of the moral law but too weak, in our constitution, to follow it. In his *Lectures on Ethics*, Kant provides a similar account of moral despair, which he describes as someone’s decision to *give up* approximating the standards of action demanded by morality. Vigilantius reports Kant saying in a 1793 lecture that moral despair is “the decision arising from doubt as to man’s capacity for ever attaining to the moral law, whereby we give up all effort to approach it, and declare ourselves inca-

pable of improving or elevating our worth” (LE 27:611). I believe Kant’s moral phenomenology is meant to address this, which is why he describes the feeling of elevation as a feeling of proper *esteem* for ourselves as rational beings.

In Ch.III, however, Kant does not mention the problem of despair, which is surprising. The problem he thinks poses a real threat is not one of giving up on morality, but rather of turning morality into an *affect*, what he calls “enthusiasm” (*Schwärmerei*). At first, this is puzzling. When the question at stake is how a purely intellectual principle can function as an incentive, why would Kant be worried about someone who is over-zealous about morality (KpV 5:82-86)? Why not focus on someone who, out of despair, risks giving up? My speculation is that Kant is less worried about the morally despondent because they are simply ignorant of the moral incentive: they need inspiration, which Kant assigns to moral education. The enthusiast, however, is caught in a “delusion” (*Täuschung*), which is, for Kant, much harder to overcome. My aim for the remainder of this chapter is to explain what this *Täuschung* is and how, according to Kant, it comes about.

The issue of enthusiasm comes up in Kant’s discussion of the love commandments (KpV 5:83), so this is a good place to start. The command to love God, or to love your neighbour as yourself, is not a command to *feel* benevolence to God or to others. It is, rather, a command to make it your maxim to *strive* to possess this disposition, even though it can only be endlessly approximated. Kant notes that it would be contradictory to command a feeling we already have or to command us to some end that excludes any feeling as its motivating ground. The love commandment does not require us to have feelings of love, or to take steps toward possessing those feelings completely. Nor does it require that our feelings are the basis of our commitment to ends assigned by reason, such as our relationship with God or our neighbours. The love commandment presents us with the moral disposition in its perfection, i.e, as an ideal, whereby one who possessed such a disposition would gladly do everything specified by the moral law. This is the idea of a “holy” will we discussed in section 2.5. The point is that sensible and imperfect beings like us can only ever approximate holiness in our actions. The human condition will forever leave us inclined to act on impulses and inclinations contrary to the moral law, which is why the love commandment, and any commandment of pure practical reason, will remain for us an obligation. We will always encounter moral laws as duties, as what we “should” do, even though we must constantly bring our actions into comparison with a will that “wants” to act on these laws. The holy will is thus an ideal: it presents us with the image our own perfectly rational self.

This is the problem with the moral enthusiast. The enthusiast attempts to render his encounter with the law a matter of inclination. He deludes

himself into thinking he possesses holiness, i.e., that he gladly wants to do everything specified by morality. And that is why, for Kant, moral enthusiasm represents the most exaggerated form of self-conceit. The enthusiast's arrogance does not merely lead him to confer authority to his inclinations so that he can freely pursue his happiness against the dictates of morality; rather, his arrogance leads him to turn those dictates into his own sources of pleasure. In this way, he creates what Kant calls a "egotistical illusion" for himself: he imagines he does not stand in a relation of constraint to the law.¹¹ Individuals guilty of this produce 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" (KpV 5:85). Kant may give the impression of criticizing the enthusiast solely on the grounds of his affections for morality, and this might appear to indicate nothing more than Kant's resistance to see value in the emotional dimension of moral life. But on closer examination we can see that the enthusiast lacks genuine moral feelings, and that's the real problem in Kant's view. Without experiencing the distance between morality and his own nature as a sensible and imperfect being, the enthusiast fails to know his proper standing vis-a-vis the law.

How, then, does enthusiasm come about? Kant does not say much to explain this, but he leaves us at least two clues. The first occurs later in the second *Critique* where Kant talks about a *vitium subreptionis* or "error of subreption" (KpV 5:116). An error of subreption is an "optical illusion" between consciousness of the law as a determining *ground* of the will and the feeling of satisfaction that arises from it as an *effect*.

The moral disposition is necessarily connected with consciousness of the determination of the will *directly by the law*. Now, this consciousness of the faculty of desire is always the ground of a satisfaction in the action produced by it; but this pleasure, this satisfaction with oneself, is not the determining ground of the action: instead, the determination of the will directly by reason alone is the ground of the feeling of pleasure, and this remains a pure practical, not aesthetic, determination of the faculty of desire. Now, since this determination has exactly the same inward effect, that of an impulse to activity, as a feeling of the agreeableness expected from the desired action would have produced, we easily look upon what we ourselves *do* as something that we merely *passively feel* and take the moral incentive for a sensible impulse, just as always happens in so-called illusion of the senses. (KpV 5:116-117; my emphasis).¹²

¹¹As a result, we could say the enthusiast does not register the existence of "perfect" or "narrow" obligations. Every moral action is, in his eyes, an "imperfect" or "wide" obligation, i.e., a noble or meritorious action he is not under constraint to perform.

¹²"Die moralische Gesinnung ist mit einem Bewußtsein der Bestimmung des Willens

The error here is in part due to the human faculty of desire, for we are set up to feel satisfaction in representing the existence of objects agreeable to us: “consciousness of a determination of the faculty of desire is always the ground of a satisfaction in the action produced by it” (KpV 5:116). For Kant, many objects elicit pleasure in us, but only those represented by reason ground *interests*, i.e., motives for making such objects real through our own actions. The feeling of satisfaction is identical to the interest we take in actions judged available to us, because an interest is that by which reason becomes a faculty of desire, an “impulse to activity,” without which our will would remain inactive, as when we engage in empty wishing. There are two options here. On the one hand, if the object is agreeable to my desires, or if its representation produces a desire in me, then my interest in it will be “sense-based” or pathological. On the other hand, if the object is agreeable with my spontaneity of choice, as the universality of an action agrees with my capacity to give universal law, then my interest in it will be “sense-free” or purely practical. Because the universality of an action has nothing to do with its effects, we can say that a practical interest only concerns the action itself, or the principle of reason it represents. For Kant, every instance of moral action involves a sense-free interest, and while it does not arise from the representation of an action’s effects, it is still the basis of a *feeling*.

The possibility of illusion arises here, I believe, because we may take the feeling of interest to be the determining ground of the will. Our consciousness of the law produces an interest because we are aware of the agreement between the law and our capacity for rational choice. Because this consciousness of agreement grounds a *feeling of satisfaction* we are prone to confuse the subjective effect with the objective ground, and thus commit an error of subreption. As Kant puts it: “this determination *has exactly the same inward effect*, that of an impulse to activity, as a feeling of agreeableness expected from the desire action would have produced, we easily look upon what we ourselves do as something that we merely passively feel” (KpV 5:117; my emphasis). This optical illusion may explain how an agent could replace (in his mind) a command of moral action with his feeling of

unmittelbar durchs Gesetz nothwendig verbunden. Nun ist das Bewußtsein einer Bestimmung des Begehrungsvermögens immer der Grund eines Wohlgefallens an der Handlung, die dadurch hervorgebracht wird; aber diese Lust, dieses Wohlgefallen an sich selbst, ist nicht der Bestimmungsgrund der Handlung, sondern die Bestimmung des Willens unmittelbar, bloß durch die Vernunft, ist der Grund des Gefühls der Lust, und jene bleibt eine reine praktische, nicht ästhetische Bestimmung des Begehrungsvermögens. Da diese Bestimmung nun innerlich gerade dieselbe Wirkung eines Antriebs zur Thätigkeit thut, als ein Gefühl der Annehmlichkeit, die aus der begehrten Handlung erwartet wird, würde gethan haben, so sehen wir das, was wir selbst thun, leichtlich für etwas an, was wir bloß leidentlich fühlen, und nehmen die moralische Triebfeder für sinnlichen Antrieb, wie das allemal in der sogenannten Täuschung der Sinne (hier des innern) zu geschehen pflegt.” Interestingly, Kant says that even “the most practiced” cannot always avoid this illusion (KpV 5:116).

desire to perform it. For example, the enthusiast may see that *x*-ing is valid as a possible law for all rational beings, and so he will judge it good by the standards of the categorical imperative. This judgment will then give rise to a feeling of satisfaction in *x*-ing, for he will represent the agreeableness between the judgment and his capacity for rational choice. But the enthusiast will in turn commit a *vitium subreptionis*: he will confuse this feeling, the satisfaction produced by the thought of *x*-ing, with his recognition of what makes *x*-ing good, namely, its universality. And so he will take his feeling for the determining ground of *x*-ing.

Now, as far as this process is concerned, one could argue that the error of subreption does not explain the possibility of moral enthusiasm, since subreption is an error of judgment (a confusion of the ground-consequent relationship), but that does not undermine the morality of the action. In the second *Critique*, Kant's claim is much stronger: that subreption "demeans and deforms" the real incentive to morality, i.e., "respect" (*Achtung*). But he does not say how this comes about. To account for this, I think we need to look ahead a few years to the *Metaphysics of Morals*, where Kant provides a brief but important definition of moral enthusiasm. He says that "only the apparent strength of someone feverish lets a lively sympathy even for *what is good* rise into an affect, or rather, degenerate into it." And later: "an affect, even one aroused by the thought of *what is good*, is a momentary, sparkling phenomenon that leaves one exhausted" (MS 6:409). For the sake of technical precision, I think Kant should have said enthusiasm is an affect aroused by the feeling of satisfaction brought about by the thought of the good. This makes sense on the general account of desire Kant presents us. The thought of the good must ground a practical interest, because the thought agrees with our capacity for rational choice, and this agreement is the basis of our satisfaction (our impulse to activity) in representing the good. What he "demeans and deforms" is strictly speaking not the thought of the moral law, but the practical interest this thought produces in him. So the enthusiast demeans his sense-free interest in morality by turning it into an *affect*, a pathological feeling of pleasure, thus rendering his satisfaction in what is good a "momentary, sparkling phenomenon" which flares up in his mind only to leave him exhausted once his energy is spent. In this regard, the enthusiast manages to accomplish in practice what the sentimentalist argues on theoretical grounds.

This may explain why enthusiasm is a more serious practical failure than despair, although we can see that Kant is responding to both in Ch.III. The feeling of elevation we have for our rational nature is, in Kant's view, a source of moral *confidence*. We realize it is within our power to approximate the moral law. This feeling is "*so little displeasure*" that "once one has laid self-conceit aside and allowed practical influence to that respect,

one can in turn never get enough of contemplating the majesty of this law, and the soul believes itself elevated in proportion as it sees the holy elevated above itself and its frail nature” (KpV 5:78). Kant’s subsequent worry about enthusiasm is, I think, related to this. For if morality has the potential to present us with a highest incentive, outranking every possible empirical motive, then the risk that it will degenerate into affect is all the more great. For this reason, we can understand why Kant recommends to moral educators that they emphasize examples of virtuous conduct that come at a great sacrifice—even at the sacrifice of one’s life. So between the two moments of respect I have outlined in section 6.2—the moment of humiliation and the moment of elevation—Kant thinks we should put more emphasis on the former, at least in our teaching of morals. “It is quite advisable to praise actions in which a great, unselfish, sympathetic disposition or humanity is manifested. But in this case one must call attention not so much to the *elevation of the soul*, which is very fleeting and transitory, as to the *subjection of the heart to duty*, from which a more lasting impression can be expected” (KpV 5:155n).

Kant may have overstated his point. In this thesis I have put more emphasis on the “elevation of the soul,” since this overcomes the worry many readers of Kant have, namely, that our sensibility is only rebuffed by practical reason, and that Kant is unable to account for a positive connection between the moral law and the heart’s desire. I have argued that this worry is unfounded; but it does contain an important point. For Kant just as well could have finished Ch.III with an account of moral despair, showing how his phenomenology provides resources for educators who are challenged by the morally despondent, or those who have dabbled in philosophy and express a “deep regret” over the frailty of human nature. Clearly, the negative-positive structure of moral feeling gives us material to address both problems, moral arrogance and moral despair. Was Kant justified in thinking the former is a greater practical failure? I am not sure. The answer will depend on how we understand Kant’s theory of moral education. But that is a topic I will have to save for another occasion.

6.6 Closing Remarks

As I mentioned in the Introduction, Lewis White Beck is one of the few commentators to recognize the argumentative weight Kant’s phenomenology must carry. The mystery of how a practical determination of our faculty of desire is possible must be removed “from the phenomenological surface,” as he puts it, “for the thing is so puzzling that doubts of its reality can have the actual effect of reducing the effectiveness of this incentive” (1960, p. 211). If we are unable to explain how practical reason can access our faculty

of desire in the same way that empirical incentives do, “we leave the door open to those who would give simpler explanations that have the advantage of being readily comprehended but the disadvantage of being theoretically wrong and damaging to morality itself” (Beck 1960, p. 211). This also helps refocus the criteria by which we should judge a moral theory’s ability to respond to skepticism. For Kant, in attempting to convince the skeptic why he should be moral, we commit the same error underlying his question. We assume that morality is a domain we can enter or exit, and that we are able to formulate reasons for why those who stand on the threshold, the moral outsiders, should come inside.

This is the assumption that goes unquestioned when the stakes of moral justification are at their highest. Kant’s point is that the real threat of skepticism does not have its origins external to morality. At least the problem worth addressing, the worry that our will may be unfit for morality, arises for those “inside” morality. I do not think Kant fully understood this in *Groundwork* III, which is one of the reasons for its failure. But what Kant did see clearly in *Groundwork* III, and possibly before then, is that our deepest moral concerns do not find their best expression in the question “Why should I be moral?” One might complain that until this question is addressed, skepticism continues to remain a threat. For Kant, as I argued in Chapter 2, this is just one side of the moral question, what I called the Question of Authority. And Kant thinks there are much more pressing worries to address, such as our standing as rational agents (the Question of Identity), our experience of constraint (the Question of Obligation), and our interest in morality (the Question of Motivation).

My aim in this thesis has been to interpret, and sometimes reconstruct, Kant’s answers to these questions. Along the way, I have tried to present Kant’s project of moral justification in its best possible light, defending his arguments against a number of long-standing criticisms. In Chapter 3, I argued that Kant’s worry of a “hidden circle” arises because he anticipated the reader would ascribe freedom to himself without argument; and that unless we make a critical distinction between “two worlds,” we will fall short of understanding ourselves as agents. Later in Chapter 3, I argued that the world of understanding and the world of sense are not distinct ontological realms, but rather models or archetypes for understanding our agency. Against a prevalent view in the literature, I maintained that the direction of Kant’s argument remains the same in *Groundwork* III and the second *Critique*: both try to justify the presupposition of freedom through our practical cognition of the moral law. My criticism of *Groundwork* III was that it missed a final *Übergang* that would explain how the moral law connects to our sensibility, a final shift of frameworks from a “critique” to an “aesthetic of pure practical reason.” My aim in Part II was to understand how Kant overcame

this failure in 1788. In Ch.III of the second *Critique*, I argued that we find a “dialectical” account of the effects practical reason has on our sensibility—in particular, its “restraining” of our predisposition of *self-love*, and its “striking down” of our propensity of *self-conceit*. For lack of space, I was unable to offer a full discussion of the criticisms raised by Kant’s immediate successors; but I believe that my reconstruction avoids what I called Hegel’s Challenge, that Kantian morality is unable to establish an a priori connection between reason and feeling.

If I am right, then Kant’s moral phenomenology is meant to relieve a skeptical pressure: the pressure put on us by the apparent failure of fit between our sensible will and the demands of the moral law. The pressure may remain on those who refuse to accept the force of an account, like Kant’s, that presupposes our active recognition of moral authority. Too much is being assumed here, on too little grounds. The weakness of a phenomenological account of morality’s application may strike many as a second-best account, one we turn to when theoretical knowledge fails to deliver us the kind of explanation and justification we’re seeking. It may be wondered, though, what such a strong theoretical account could accomplish, were it available to us. Even if we could explain our responsiveness to morality in theoretical terms, how would those terms speak to us? The skeptical problem we face here, the gap between what morality requires of us and what we’re capable of doing, arises from *our* perspective. Perhaps the Kantian insight that must be taken seriously today is that the burden of moral theory is not to justify the claims of morality for those supposedly outside of morality’s domain, the amoralist or the radical skeptic. The burden is to make those claims, their application and practicality, transparent to those within morality’s domain. This helps rethink and refocus the purpose of a phenomenological argument. For if skeptical doubts have their origin in the human experience of morality, then the only solution is to develop our answers from the human standpoint itself.

Abbreviations

Passages from the *Critique of Pure Reason* are cited from the 1781 (“A”) and 1787 (“B”) editions. I will use the following abbreviations—volume and page number—from the “Academy edition,” *Kant’s gesammelte Schriften*, edited by the Royal Prussian (subsequently German, then Berlin Brandenburg) Academy of Sciences (Berlin: Georg Reimer, subsequently Walter de Gruyter and Co., 1900). Unless otherwise noted, English translations will come from the Cambridge Edition of Kant’s works. I will mark any change to the translation with “modified.”

- Anth** *Anthropologie in pragmatischer Hinsicht* (1798), trans. Robert B. Louden, in *Lectures on Pedagogy in Anthropology, History, and Education*, ed. Günter Zöller and Robert B. Louden (Cambridge: Cambridge University Press, 2007).
- Br** *Briefe*, trans. Arnulf Zweig, *Correspondence* (Cambridge: Cambridge University Press, 1999).
- G** *Grundlegung zur Metaphysik der Sitten* (1785), trans. Mary K. Gregor, *Groundwork of the Metaphysics of Morals in Practical Philosophy*, ed. Allen Wood (Cambridge: Cambridge University Press, 2006).
- GSE** *Beobachtungen über das Gefühl des Schönen und Erhabenen* (1764), trans. Paul Guyer, *Observations on the Feeling of the Beautiful and the Sublime*, in *Lectures on Pedagogy in Anthropology, History, and Education*, ed. Günter Zöller and Robert B. Louden (Cambridge: Cambridge University Press, 2007).
- DG** “Untersuchung über die Deutlichkeit der Grundsätze der natürlichen Theologie und der Moral” (1765), trans. David Walford and Ralf Meerbote, “Inquiry concerning the Distinctness of the Principles of Natural Theology and Morality” [otherwise known as the “Prize Essay”] in *Theoretical Philosophy 1755-1770* (Cambridge: Cambridge University Press, 1992).

-
- KpV** *Kritik der praktischen Vernunft* (1788), trans. Mary K. Gregor, *Critique of Practical Reason in Practical Philosophy*, ed. Allen Wood (Cambridge: Cambridge University Press, 2006).
- KrV** *Kritik der reinen Vernunft* ([A] Edition: 1781/ [B] Edition: 1787), trans. Allen Wood and Paul Guyer, *Critique of Pure Reason* (Cambridge: Cambridge University Press, 1998).
- KU** *Kritik der Urteilskraft* (1790), trans. Paul Guyer and Eric Matthews *Critique of the Power of Judgment* (Cambridge: Cambridge University Press, 2000).
- LE** *Lectures on Ethics*, trans. Peter Heath, ed. J.B. Schneewind and Peter Heath (Cambridge: Cambridge University Press, 1997).
- LM** *Lectures on Metaphysics*, trans. Karl Ameriks and Steve Naragon (Cambridge: Cambridge University Press, 1997).
- MS** *Die Metaphysik der Sitten* (1797), trans. Mary K. Gregor, *Metaphysics of Morals in Practical Philosophy*, ed. Allen Wood (Cambridge: Cambridge University Press, 2006).
- MpVT** “Über das Mißlingen aller philosophischen Versuche in der Theodicee” (1791), trans. George Di Giovanni, “On the Miscarriage of all Philosophical Trials in Theodicy,” in *Religion and Rational Theology*, ed. Allen Wood and George Di Giovanni (Cambridge: Cambridge University Press, 2005).
- N** “Nachricht von der Einrichtung seiner Vorlesungen in dem Winterhalbenjahre von 1765-1766,” trans. David Walford and Ralf Meerbote, *Metaphysics of Morals in Theoretical Philosophy 1755-1770* (Cambridge: Cambridge University Press, 1992).
- PR** *Vorlesungen über die philosophische Religionslehre* (1817), trans. Allen Wood, *Lectures on the Philosophical Doctrine of Religion in Religion and Rational Theology*, ed. Allen Wood and George Di Giovanni (Cambridge: Cambridge University Press, 2005).
- R** *Notes and Fragments*, trans. Curtis Bowman, Paul Guyer, Frederick Rauscher, ed. Paul Guyer (Cambridge: Cambridge University Press, 2005).
- RGV** *Die Religion innerhalb der Grenzen der bloßen Vernunft* (1793), trans. Allen Wood, *Religion within the Boundaries of Mere Reason in Religion and Rational Theology*, ed. Allen Wood and George Di Giovanni (Cambridge: Cambridge University Press, 2005).

Bibliography

- Allison, H. 1981. "Transcendental Schematism and the Problem of the Synthetic *A Priori*." *Dialectica* (35:1): 57-83.
- 1989. "Justification and Freedom in the *Critique of Practical Reason*," in *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, ed. Eckart Förster, Stanford: Stanford University Press.
- 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- 1996. "Autonomy and Spontaneity in Kant's Conception of the Self." *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge: Cambridge University Press.
- 2002. "On the Very Idea of a Propensity to Evil." *The Journal of Value Inquiry* (36): 337-348.
- 2004. *Kant's Transcendental Idealism*. New Haven: Yale University Press.
- Ameriks, K. 2003. *Interpreting Kant's Critiques*. Oxford: Clarendon Press.
- 2006. "Kant, Hume, and the Problem of Moral Motivation," in *Kant and the Historical Turn: Philosophy as Critical Interpretation*. Oxford: Clarendon Press, 2006: 89-107.
- Aune, B. 1979. *Kant's Theory of Morals*. Princeton: Princeton University Press.
- Baumgartens, A.G. 1739. *Metaphysik*. Jena: Scheglmann, 2004.
- Beck, L.W. 1960. *A Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press.
- 1965. "The Fact of Reason: An Essay on Justification in Ethics," in *Studies in the Philosophy of Kant*. Indianapolis: Indiana University Press.
- Benton, R. 1978. "The Transcendental Argument in Kant's *Groundwork of the Metaphysics of Morals*." *The Journal of Value Inquiry* (12:3): 225-238.
- Broadie, A, and Pybus, EM. 1975. "Kant's Concept of 'Respect.'" *Kant-Studien* (66): 58-64.

-
- Budick, S. 2000. "Kant's Miltonic Test of Talent: 'When I Consider' in the *Groundwork of the Metaphysics of Morals*." *Modern Language Quarterly* (61:3): 481-518.
- Borges, M. 2004. "What Can Kant Teach Us About the Emotions?" *The Journal of Philosophy* (101:3): 140-158.
- Burge, T. 1998. "Reason and the First Person," in *Knowing Our Own Minds*, ed. C. Wright, B. Smith and C. Macdonald. Oxford: Oxford University Press.
- Cavell, S. 1976. "Knowing and Acknowledging," in *Must We Mean What We Say? A Book of Essays*. Cambridge: Cambridge University Press.
- 1979. *The Claim of Reason: Wittgenstein, Skepticism, Morality, and Tragedy*. Clarendon: Oxford University Press.
- Copp, D. 1992. "The 'Possibility' of a Categorical Imperative: Kant's Groundwork, Part III." *Philosophical Perspectives* (6): 261-284.
- Crowell, S. 2003. "Kantianism and Phenomenology," in *Phenomenological Approaches to Moral Philosophy*, ed. John Drummond and Lester Embree. Kluwer.
- Darwall, S. 1977. "Two Kinds of Respect." *Ethics* (88:1): 36-49.
- 1995. *The British Moralists and the Internal 'Ought': 1640-1740*. Cambridge: Cambridge University Press.
- 2006. *The Second-Person Standpoint: Morality, Respect, and Accountability*. Cambridge MA: Harvard University Press.
- Derrida, J. 1982. "The Ends of Man," in *Margins of Philosophy*, trans. Alan Bass. Chicago: University of Chicago Press.
- Engstrom, S. 1993. "Allison on Rational Agency." *Inquiry* (36:4): 405-418.
- 1994. "The Transcendental Deduction and Skepticism." *Journal of the History of Philosophy* (32:3): 359-380.
- 2009. *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge MA: Harvard University Press.
- Enoch, D. 2006. "Agency, Shmagency: Why Normativity Won't Come from What Is Constitutive of Action." *Philosophical Review* 115(2): 169-198.
- Falk, W.D. 1948. "Ought and Motivation," in *Ought, Reasons, and Morality: Collected Papers of W. D. Falk*. Cornell: Cornell University Press.
- Forster, M. 2008. *Kant and Skepticism*. Princeton: Princeton University Press.
- Förster, Eckart, ed. 1989. *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*. Stanford: Stanford University Press.
- Frankena, W.K. 1958. "Obligation and Motivation in Recent Moral Philosophy," in *Perspectives on Morality: Essays by William K. Frankena*, ed. K.E. Goodpaster. Notre Dame: Notre Dame University Press, 1976.
- Franks, P. 1997. "Freedom, *Tatsache* and *Tathandlung* in the Development of Fichte's Jena *Wissenschaftslehre*." *Archiv Für Geschichte der Philosophie* (79): 310-323.

-
- 1998. “Transcendental Arguments, Reason, and Skepticism: Contemporary Debates and the Origins of Post-Kantianism,” in *Transcendental Arguments: Problems and Prospects*, ed. Robert Stern. Oxford: Oxford University Press.
- 2005. *All or Nothing: Systematicity, Transcendental Arguments, and Skepticism in German Idealism*. Cambridge, MA: Harvard University Press.
- Frankfurt, H. 1998. *The Importance of What We Care About: Philosophical Essays*. Cambridge: Cambridge University Press.
- Frierson, P. 2003. *Freedom and Anthropology in Kant’s Moral Philosophy*. Cambridge: Cambridge University Press.
- Galvin, R. 1991. “Does Kant’s Psychology of Morality Need Basic Revision?” *Mind* (100): 221-236.
- Gibbons, S.L. 1994. *Kant’s Theory of Imagination: Bridging Gaps in Judgement and Experience*. Oxford: Oxford University Press.
- Ginsborg, H. 1998. “Korsgaard on Choosing Nonmoral Ends.” *Ethics* (109): 5-21.
- Guevara, D. 2000. *Kant’s Theory of Moral Motivation*. Boulder: Westview.
- Guyer, P. 2007. *Kant’s ‘Groundwork for the Metaphysics of Morals’: A Reader’s Guide*. Continuum.
- 2008. *Knowledge, Reason, and Taste: Kant’s Response to Hume*. Princeton: Princeton University Press.
- Grenberg, J. 2009. “The Phenomenological Failure of *Groundwork* III.” *Inquiry* (52:4): 335-356.
- Hegel, G. W. F. 1807. *Phenomenology of Spirit*, trans. A.V. Willer. Oxford: Oxford University Press, 1977.
- Heidegger, M. 1927. *The Basic Problems of Phenomenology*, trans. Albert Hofstadter. Indianapolis: Indiana University Press, 1988.
- Henrich, D. 1960. “The Concept of Moral Insight and Kant’s Doctrine of the Fact of Reason,” in *The Unity of Reason: Essays on Kant’s Philosophy*, ed. Richard L. Velkley. Cambridge MA: Harvard University Press, 1996.
- 1963. “Ethics as Autonomy,” trans. Louis Hunt, in *The Unity of Reason: Essays on Kant’s Philosophy*, ed. Richard L. Velkley. Cambridge MA: Harvard University Press, 1996.
- 1968. “The Proof-Structure of Kant’s Transcendental Deduction.” *Review of Metaphysics* (22): 640-59.
- 1975. “Die Deduktion des Sittengesetzes, Über die Gründe der Dunkelheit des letzten Abschnitts von Kants ‘Grundlegung zur Metaphysik der Sitten.’” in *Denken im Schatten des Nihilismus*. Darmstadt.
- 1989. “Kant’s Notion of a Deduction and the Methodological Background of the First *Critique*,” in *Kant’s Transcendental Deductions*, ed. Eckart Förster. Stanford: Stanford University Press.

-
- Herman, B. 1989. "Justification and Objectivity: Comments on Rawls and Allison," in *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, ed. Eckart Förster, Stanford: Stanford University Press.
- 2005. "Transforming Incentives: Feelings and the Making of the Kantian Moral Agent," in *Philosophical Aspects of the Emotions*, ed. Asa Carlson. Stockholm: Thales.
- 2007. *Moral Literacy*. Cambridge MA: Harvard University Press.
- Hill, T. 1985. "Kant's Argument for the Rationality of Moral Conduct." *Pacific Philosophical Quarterly*: 3-23.
- Horgan, T, and Timmons, M. 2005. "Moral Phenomenology and Moral Theory." *Philosophical Issues* 15(1): 56-77.
- 2008. "What does Moral Phenomenology Tell Us About Moral Objectivity?" *Social Philosophy and Policy Foundation* (25:1) :267-300.
- Hume, D. 1739. *A Treatise of Human Nature*, Second Edition, ed. P.H. Nidditch. Oxford: Clarendon Press, 1978.
- Kain, P. Forthcoming. "Practical Cognition, Intuition, and the Fact of Reason," in *God, Freedom, and Immortality: Kant's Moral Metaphysics*. ed. Benjamin Lipscomb and James Krueger. Berlin: DeGruyter.
- Kantorowicz, E. 1961. "The Sovereignty of the Artist. A Note on Legal Maxims and Renaissance Theories of Art." *De Artibus Opuscula XL. Essays in Honor of Erwin Panofsky*. New York.
- Korsgaard, C. 1996a. "Morality as Freedom," in *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- 1996b. "Skepticism about Practical Reason," in *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- 1996c. *The Sources of Normativity*. Cambridge: Cambridge University Press.
- 1997. "The Normativity of Instrumental Reason," in *Ethics and Practical Reason*, ed. Garrett Cullity and Berys Gaut. Oxford: Oxford University Press.
- 1999. "Self-Constitution in the Ethics of Plato and Kant." *The Journal of Ethics* (3): 1-29.
- Lau, C. 2008. "Freedom, Spontaneity and the Noumenal Perspective." *Kant-Studien* (99): 312-338
- Lavin, D. 2004. "Practical Reason and the Possibility of Error." *Ethics* (114): 424-457.
- Longuenesse, B. 1998. *Kant and the Capacity to Judge*. Princeton: Princeton University Press.
- 2005. "Moral Judgment as a Judgment of Reason," in *Kant on the Human Standpoint*. Cambridge: Cambridge University Press.

-
- Lukow, P. 1993. "The Fact of Reason: Kant's Passage to Ordinary Moral Knowledge." *Kant-Studien* (84:2): 204-221.
- Mandelbaum, M. 1955. *The Phenomenology of Moral Experience*. Illinois: Free Press.
- McCarty, R. 1993. "Kantian Moral Motivation and the Feeling of Respect." *Journal of the History of Philosophy* (31:3): 421-435.
- 1994. "Motivation and Moral Choice in Kant's Theory of Rational Agency." *Kant-Studien* (85-1): 15-31.
- 2009. *Kant's Theory of Action*. Oxford: Oxford University Press.
- McCarthy, M. 1985 "The Objection of Circularity in *Groundwork* III." *Kant-Studien* (76:1): 28-42.
- McDowell, J. 1995. "Might there be External Reasons?" in *World, Mind, and Ethics: Essays on the Ethical Philosophy of Bernard Williams*, ed. J.E.J. Altham and Ross Harrison. Cambridge: Cambridge University Press.
- Moran, R. 2001. *Authority and Estrangement: An Essay on Self-Knowledge*. Princeton: Princeton University Press.
- Morgan, S. 2005. "The Missing Formal Proof of Humanity's Radical Evil in Kant's *Religion*." *Philosophical Review* (114:1): 63-114.
- Unpublished Manuscript. "Kant on Self-Conceit"
- Morrisson, I. 2008. *Kant and the Role of Pleasure in Moral Action*. Ohio: Ohio University Press.
- Nagel, T. 1970. *The Possibility of Altruism*. Princeton: Princeton University Press.
- O'Neill, O. 1989. *Constructions of Reason: Explorations in Kant's Practical Philosophy*. Cambridge: Cambridge University Press, 1989.
- Onof, C. 2009. "Reconstructing the Grounding of Kant's Ethics: A Critical Assessment." *Kant-Studien* (100): 496-517.
- Paton, H. J. 1947. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. London: Hutchinson's University Library.
- Pennington, K. 1993. "Innocent III and the Divine Authority of the Pope," in *Popes, Canonists, and Texts: 1150-1550*. Aldershot.
- Piper, A. 2001. "Kant's Intelligible Standpoint on Action," in *Systematische Ethik mit Kant*, ed. Hans-Ulrich Baumgarten und Carsten Held. München/Freiburg: Alber.
- Pippin, R. 1987. "Kant and the Spontaneity of Mind." *Canadian Journal of Philosophy* (17:2): 449-476.
- Post, G. 1972. "Vincentius Hispanus, 'Pro Ratione Voluntas,' and Medieval and Early Modern Theories of Sovereignty." *Traditio* (28): 159-184.
- Proops, I. 2003. "Kant's Legal Metaphor and the Nature of a Deduction." *Journal of the History of Philosophy* (41:2): 209-229.
- Pryor, J. 2000. "The Skeptic and the Dogmatist." *Nous* (34:4): 517-549.

-
- Quarfood, M. 2005. "The Circle and Two Standpoints (GMS III, 3)," in *Groundwork for the Metaphysics of Morals*, ed. Christoph Horn and Dieter Schönecker. Berlin: DeGruyter.
- Raphael, D.D. 1991. (ed.) *British Moralists 1650-1800* (2 Volumes). Indianapolis: Hackett.
- Rawls, J. 1989. "Themes in Kant's Moral Philosophy," in *Kant's Transcendental Deductions: The Three Critiques and the Opus Postumum*, ed. Eckart Förster, Stanford: Stanford University Press.
- 2000. *Lectures on the History of Moral Philosophy*, ed. Barbara Herman. Cambridge, MA: Harvard University Press.
- Reath, A. 1989. "Kant's Theory of Moral Sensibility. Respect for the Moral Law and the Influence of Inclination." *Kant-Studien* (80:3): 284-302.
- 2006. *Agency and Autonomy in Kant's Moral Theory*. Oxford: Oxford University Press.
- Reich, K. 1939. "Kant and Greek Ethics." (Part I and II) *Mind* (48): 446-51.
- Ripstein, A. 2009 *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge MA: Harvard Univ Press.
- Rousseau, J.J. 1755. "Discours sur l'origine et les fondements de l'inégalité parmi les hommes," trans. Donald A. Cress, "Discourse on the Origin and Basis of Inequality Among Men" [Second Discourse], in *Basic Political Writings*, ed. Peter Gay. Indianapolis: Hackett, 1987.
- 1762. *Émile, ou De l'éducation*, trans. Allan Bloom, *Emile, or On Education*. New York: Basic Books, 1979.
- Russell, P. 2006. "Practical Reason and Motivational Skepticism," in *Moralische Motivation: Kant und die Alternativen*, eds. Heidener F. Klemme, Manfred Kühn, Dieter Schönecker. Hamburg: Meiner.
- Schiller, F. 1793. "Über Anmut und Würde," in 'On Grace and Dignity,' in *Its Cultural Context*, ed. Curran and Fricker. Camden House, 2005.
- Schneewind, J.B. 1998. *The Invention of Autonomy*. Cambridge: Cambridge University Press.
- 2003. (ed.) *Moral Philosophy from Montaigne to Kant*. Cambridge: Cambridge University Press.
- Schönecker, D. 1999. *Grundlegung III: Die Deduktion des Kategorischen Imperativs*. Freiburg: Karl Alber.
- 2005. "How is a Categorical Imperative Possible?," in *Groundwork for the Metaphysics of Morals*, ed. Christoph Horn and Dieter Schönecker. Berlin: DeGruyter.
- 2007. (with Allen Wood). *Kants 'Grundlegung zur Metaphysik der Sitten': Ein einführender Kommentar*, Third Edition. Paderborn: Schöningh-Verlag.

-
- Schroeter, F. 2005. "Normative Concepts and Motivation." *Philosophers' Imprint* (5:3):1-23
URL: www.philosophersimprint.org/005003/
- Sellars, W. 1970. "...this I or He or It (The Thing) which thinks..." *Proceedings and Addresses of the American Philosophical Association* (44): 5:31.
- Sensen, O. Forthcoming. "Kant's Conception of Inner Value." *European Journal of Philosophy*.
- Setiya, K. 2007. *Reasons Without Rationalism*. Princeton: Princeton University Press.
- Smith, M. 1995. "Internal Reasons." *Philosophy and Phenomenological Research* (55:1) 109-131.
- Smith, D.W. 2008. "Phenomenology" in the Stanford Encyclopedia of Philosophy:
URL: <http://plato.stanford.edu/entries/phenomenology/>
- Stratton-Lake, P. 2001. *Kant, Duty and Moral Worth*. London: Routledge.
- Strawson, P. 1966. *The Bounds of Sense: An Essay on Immanuel Kant's 'Critique of Pure Reason'*. Oxford: Oxford University Press.
- Sussman, D. 2005. "Perversity of the Heart." *Philosophical Review* (114:2): 153-177.
- 2008. "From Deduction to Deed: Kant's Grounding of the Moral Law." *Kantian Review* (13:1): 52-81.
- Svavarsdottir, S. 1998. "Moral Cognitivism and Motivation." *Philosophical Review* (108:2): 161-219.
- Tenenbaum, S. 2003. "Speculative Mistakes and Ordinary Temptations: Kant on Instrumentalist Conceptions of Practical Reason." *History of Philosophy Quarterly* (20:2): 203-223.
- Forthcoming. "The Idea of Freedom and Moral Cognition in *Groundwork* III." *Philosophy and Phenomenological Research*.
- Thiel, U. 1997. "Varieties of Inner Sense: Two Pre-Kantian Theories." *Archiv Für Geschichte der Philosophie* (79): 58-79.
- Timmerman, J. 2007. *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Cambridge: Cambridge University Press.
- Velleman, J.D. 1996. "The Possibility of Practical Reason." *Ethics* (106: 4): 694-726.
- 2006. "Willing the Law," in *Self to Self: Selected Essays*. Cambridge: Cambridge University Press.
- Wallace, J. 1990. "How to Argue about Practical Reason." *Mind* (99): 355-385.
- Ware, O. 2009. "The Duty of Self-Knowledge." *Philosophy and Phenomenological Research* (79.3): 671-698.

-
- 2010. “Fichte’s Voluntarism.” *European Journal of Philosophy*. (18:1): 262-282.
- Williams, B. 1981. “Internal and External Reasons,” in *Moral Luck: Philosophical Papers 1973-1980*. Cambridge: Cambridge University Press.
- 1996. “History, Morality, and the Test of Reflection,” in “Replies” to *The Sources of Normativity*. Cambridge: Cambridge University Press.
- Wood, A. 1999. *Kant’s Ethical Thought*. Cambridge: Cambridge University Press.
- 2005. *Kantian Ethics*. Cambridge: Cambridge University Press.
- Zinkin, M, H. 2006. “Respect for the Law and the Use of Dynamical Terms in Kant’s Theory of Moral Motivation.” *Archiv Für Geschichte der Philosophie* (88:1): 31-53.

Index

- Act of Reason, 78, 114
Affectivism, 15, 84
Ainslie, Don, ii
Allison, Henry, 58
Allison, Henry, 5, 84
Ameriks, Karl, 84
Analytic Kantianism, 4
- Bad Conscience, 123
Beautiful Soul, 105
Beck, Lewis White, 14, 130
Burge, Tyler, 45
- Capacity to Feel, 82, 100
Categorical Imperative, 4, 17, 32
Causa Noumenon, 94
Causality of Reason, 78
Clark, Phil, ii
Cohen, G.A., 74
Comay, Rebecca, ii
Constitutivism, 46
Corrupt Prince, 114
- Deduction, 17, 27, 31, 54, 70, 88
Despair, 128
Dogmatism, 36
- Ectypal Nature, 85, 90
Elevation, 119
Empty Formalism, 4
Engstrom, Steve, ii
Enoch, David, 46
Enthusiasm, 124
Error of Subreption, 126, 128
- Fact of Reason, 4, 36, 42
Faculty of Desire, 99
Franks, Paul, ii, 37
- Gibbs, Bob, ii
Guyer, Paul, 55
- Happiness, 50
Hardened Villain, 72
Harms, Anna Leah, ii
Hegel's Challenge, 101, 102, 107, 131
Hegel, G.W.F., 101
Heidegger, Martin, 99
Henrich, Dieter, 67, 102
Herman, Barbara, ii, 4
Herz, Marcus, 11
Hidden Circle, 52, 130
Hill, Thomas, 5, 19
Holy Will, 28
Human Standpoint, 131
Hume, David, 5
Humiliation, 115
Hutcheson, Francis, 10, 110
- Idea of Freedom, 44
Intellectualism, 15, 84
- Justification, 4
Juvenal, 37
- Kain, Patrick, 37
Korsgaard, Christine, 5, 8, 22, 75
- Lavin, Doug, 46
Longuenesse, Béatrice, 89
- Mafioso, 75
McCarty, Richard, 84, 85
Metaphysics of Morals, 26
Moral Feeling, 10, 13, 23
Moral Incentive, 17
Moral Motivation, 3

Moral Psychology, 3
Moral Sensibility, 85

Nagel, Thomas, 7
Naturalism, 5, 68
Noumenal Agency, 5

O'Neill, Onora, 67
Obligatoriness, 68, 70
Onof, Christian, 27
Outermost Limit, 65

Phantom of the Brain, 6
Phenomenology, 6, 7, 109, 119
Philosophers' Stone, 1, 9, 13
Piper, Adrian, 68
Popular Philosophy, 22
Practical Aesthetic, 85, 89, 131
Practical Ideals, 29
Practical Sensibility, 87
Proops, Ian, 37

Question 'Quid Juris', 32
Question of Authority, 34
Question of Identity, 34, 43, 62
Question of Motivation, 34, 65, 78, 79, 104
Question of Obligation, 34

Rattan, Gupreet, ii
Rawls, John, 4
Reciprocity Thesis, 41
Respect (*Achtung*), 5
Reverence, 117
Ripstein, Arthur, ii

Schönecker, Dieter, 49
Schiller, Fredrich, 101, 105
Schlösser, Ulrich, ii
Schmagency, 46
Self-Activity, 56, 57
Self-Conceit, 115
Self-Deception, 22
Self-Love, 72
Sellars, Wilfred, 52
Sensibility, 14
Sensible Nature, 55
Simultaneity Thesis, 70
Skepticism, Content, 8, 21
Skepticism, Motivation, 8, 21, 76, 103
Skepticism, Transcendental, 26
Stern, Bob, ii
Stoic Sage, 29
Strawson, Peter, 5

Tenenbaum, Sergio, ii, 60
Third Antinomy, 94
Timmerman, Jens, 57
Transcendental Freedom, 6, 41, 50, 77
Transcendental Idealism, 55
Transcendental Schema, 92
Two Frameworks, 69
Two Standpoints, 55
Two Worlds, 17, 53, 71
Typic, 91

Velleman, David, 5, 45

Whiting, Jennifer, ii
Wood, Allen, 5, 19, 48

Zinkin, Melissa, 46