

THE UNIVERSITY OF CHICAGO

FREEDOM, FEELING AND CHARACTER: THE UNITY OF REASON AND SENSIBILITY
IN KANT'S PRACTICAL PHILOSOPHY

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE SOCIAL SCIENCES
JOHN U. NEF COMMITTEE ON SOCIAL THOUGHT
AND
THE FACULTY OF THE DIVISION OF THE HUMANITIES
DEPARTMENT OF PHILOSOPHY
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
ANASTASIA NETTA ARTEMYEV BERG

CHICAGO, ILLINOIS

AUGUST 2017

Certainly an authority as such is never a justification, and something is not true just because Kant has said it. Nevertheless, Kant has *the* immense significance in education for scientific, philosophical work; and one can trust him fully. In Kant as in no other thinker one has the immediate certainty that he does not cheat. And the most monstrous danger in philosophy consists in cheating, because all efforts do not have the massive character of a natural scientific experiment or that of an historical source. But where the greatest danger of cheating is, there is also the ultimate possibility for the genuineness of thinking and questioning. The meaning of doing philosophy consists in awakening the need for this genuineness and in keeping it awake.

—Martin Heidegger,
Phenomenological Interpretation of Kant's Critique of Pure Reason

Table of Contents

Acknowledgements.....	iv
Abstract.....	vi
Introduction.....	1
Chapter 1: Overcoming the Dualism of Reason and Sensibility	9
Chapter 2: Moral Respect	40
Chapter 3: Moral Character	68
Chapter 4: Moral Education and Moral Evil.....	101
Bibliography	130

Acknowledgments

I wish to first thank the members of my dissertation committee.

Matthew Boyle kindly agreed to join my committee at a late stage, but his philosophical influence is visible from the first page of the first chapter. He read my work with more care and attention than I thought I had the right to expect.

Stephen Engstrom's work has set for me a critical standard for the philosophical interpretation of Kant. The conversations he generously found room for were some of the most intense philosophical discussions I have had on the questions I raise in the dissertation.

James Conant agreed to join my committee on what I felt was blind faith and continued to offer me much needed support along the way. Philosophically, he has taught me to be generous with my opponents, to characterize them in the strongest terms possible by acknowledging their virtuous philosophical ambitions, thereby showing that it is by their own standards that they fail, not mine.

Robert Pippin has provided fair, incisive feedback and criticism on every word I have ever sent him and he always allowed me to try new directions of inquiry and argument. His kindness and trust gave me a sense of intellectual freedom without which I could not have written this essay.

In addition, I'd like to briefly acknowledge a few colleagues and friends.

Amichai Amit helped me edit the dissertation proposal as well as the original versions of the first and third chapters. He challenged me to question the role of the history of philosophy and sometimes the pretensions of philosophy itself. In so doing, he pushed me to scrutinize my own convictions, and for that I am a better philosopher.

Rory O'Connell offered valuable intellectual and technical help over the years and reviewed the final version of this essay in its entirety. Of all the fellow students who have graciously offered me feedback, challenges and suggestions over the years, Rory is the one with whom I have felt the deepest philosophical affinity, and for that I'm very grateful.

Rachel Wiseman and Daniel Phillips helped proofread and format the final document. I want to thank them both for their patience and friendship.

Finally, I would like to thank Irad Kimhi. The greater the debt of gratitude you owe an individual, the harder it is to acknowledge it. It would therefore be impossible for me to adequately express how grateful I am to him. What I can say with certainty is that of all the gifts Irad has given me, the most remarkable is the gift of riddles: open questions, hints, dim arrows pointing in directions in which I would otherwise not think to look. He did not tell me what it was I was supposed to find, or if I had succeeded in doing so. I want to believe this is because when you solve a riddle, what you discover you make your own. This essay came to be through one such an expedition.

Abstract

The dominant reception of Kant accords him the view that our capacity for feeling and our capacity for self-determination are essentially independent of one another: feelings, therefore, are essentially un-free. The negative aim of the dissertation is to argue against this standard interpretation; the positive aim is to offer an alternative.

I demonstrate that the standard interpretation is not only alien to our ordinary ways of self-understanding but that it moreover threatens the internal coherence of the Kantian account itself. I develop an alternative by examining Kant's account of how reason *motivates* the agent: first, in the account of the feeling of moral respect, and, second, in the account of moral character. I argue that moral respect does not name one particular feeling among many but that implicit in Kant's account is the previously unrecognized idea that human feeling is a unique mode of self-consciousness—disclosing the subject to herself as rational and efficacious, i.e., as a moral agent. The distinctively human capacity for feeling emerges as the form of self-consciousness *constitutive of practical agency*, i.e., of freedom. This understanding of feeling allows us to reevaluate Kant's account of moral character. Drawing on an Aristotelian understanding of the logical structure of capacities and activities I argue that character is the *activity* of maintaining one's identity as a practical agent—an activity that consists in maintaining the agent's structure of motivation. I then attend to the apparent tension between Kant's rigorism—the claim that an agent is of either wholly good or wholly evil character—and his nuanced account of the grades of moral imperfection. In addition to good and bad character, I claim, we must find room for moral *immaturity*, or the partial acquisition of moral character. To do this we must recognize the *acquisition* of moral character as a form of rational

accomplishment: the development and determination of our rational capacities for feeling. Thus, on this alternative interpretation of Kant's account, feeling and character do not oppose, but are rather the constitutive conditions of freedom.

Introduction

Kant's moral philosophy is often thought to turn on a sharp opposition between freedom and feeling. The dominant reception of Kant accords him a view according to which we are essentially rational, free beings capable of self-determination, who are nonetheless subject to our feelings: forces that are produced by nature and habituation, and that lie outside of our rational control. On this interpretation, Kant is but one more proponent, albeit a very influential one, of the traditional view that emotions are irrational and are therefore at best a matter of indifference—at worst an obstacle—to a life of freedom and rationality.

I shall argue that not only is this a misreading, but that properly understood Kant is one of the most forceful critics of this dualistic conception of our practical cognitive capacities and that moreover he proves to be an insightful guide on the path to a coherent non-dualistic view of their relationship.

I argue in Chapter One that this putative dualism is not only alien to our ordinary self-understanding but that it also threatens the internal coherence of the Kantian account. It renders incoherent the idea that we can incorporate the pursuit of the objects of our feelings, i.e., of our particular concerns, even in morally permissible actions. Moreover, it has the even graver consequence of rendering unintelligible the idea of moral motivation itself: in a picture where our affective lives are divorced from our lives as free and rational beings, it becomes impossible to see how a concern with the right thing to do could ever *per se* be what *moves* us to act.

Twentieth-century scholars that are sympathetic to Kant's account of morality and who take it to be uniquely capable of securing *objective universal* moral claims have tried to resist this picture of Kant by rehabilitating his treatment of human feeling. Specifically, commentators

have recognized that the feeling of moral respect holds the key to Kant's understanding of the relation between reason and feeling in Kant. I argue that the manner in which they have attempted to spell out the role of respect in Kant's practical philosophy nevertheless completely fails to appreciate what is philosophically most distinctive and profound in Kant's account of the distinctive role of feeling in the life of a practical rational agent.

In Chapter Two, I argue that implicit in Kant's account of the feeling of moral respect is the remarkable and previously unrecognized idea that human emotion is a unique mode of self-consciousness—of self-revelation and self-understanding—one which discloses the subject to herself as rational, embodied and capable of freely determining how she acts in the world. Human emotion, the distinctive human capacity for feeling, emerges on this account as the form of self-consciousness *constitutive of practical agency*, i.e., of freedom.

Kant's account of the feeling of moral respect is supposed to answer the question of how Kantian moral judgments are *motivating*. It poses however an apparently insoluble exegetical challenge: on the one hand, moral action is supposed to be autonomous and as such independent of feelings. On the other hand, the "feeling of moral respect" is necessary for moral action. Interpreters are divided between *intellectualists*, who jettison Kant's account of the involvement of feeling in moral motivation (rendering mysterious how moral judgments motivate at all), and *affectivists*, who claim that respect is "non-pathological" and can therefore safely motivate moral action. I demonstrate that affectivists fail to secure a characterization of "non-pathological" feeling that's adequate to account for Kant's incisive critique against the involvement of feeling in *moral* motivation. I claim that the distinction between pathological and non-pathological feeling should instead be understood by reference to Kant's distinction between the lower and

higher faculty of desire: between the kind of desire that subrational animals have and *our* rational faculty of desire, the will. If our faculty of desire is unique so must be our faculty of feeling. This faculty of feeling Kant shows us must be understood as a form of self-consciousness, which (1) constitutes and reveals the agent to be practically rational and (2) is the basis of all particular feelings.

This reading of moral respect opens up a new perspective on the topic of *moral character*. On standard accounts of Kant's practical philosophy character is interpreted as an aggregate of dispositions to act that are the product of merely empirical habituation. This renders completely mysterious how character is supposed to play the role that it does on Kant's account. According to Kant, any action that has moral worth is grounded in moral character, an overarching principle of action that determines the agent's structure of motivation. Drawing on an Aristotelian understanding of the logical structure of capacities and activities, I argue in Chapter Three that character is in fact the *activity* of maintaining one's identity as a practical agent.

In Chapter Three I elucidate Kant's account of moral character, the moral *Gesinnung*. The *Gesinnung* is a freely determined moral attitude. This moral attitude is identified with a maxim—a principle of activity—of utmost generality, which serves as the fundamental rule for the exercise of the free power of choice. It consists in the admission of *Triebfedern*, kinds of grounds of desire, into the maxim of utmost generality. I grapple with the difficulty of accounting for the way in which this moral attitude is *freely* determined. Kant claims that unlike ordinary free actions, which are “sensible deeds”, the determination of the moral attitude constitutive of character, takes place in an “intelligible deed.” I argued that we should consider

interpreting the distinction between intelligible and sensible deeds against the Aristotelian distinction between kinetic and energetic activities, activities that aim at the attainment of an end and activities that contain their end in themselves. I further argued that to admit a *Triebfeder*, a ground of desire, into one's character-constituting maxim is to determine a structure of motivation. The faculty of desire is one's capacity to act from the representation of ends. Accordingly, to determine grounds for one's faculty of desire is to determine certain kinds of ends to be grounds of activity.

I further propose that Kant's distinction between sensible and intelligible deeds, or activities, should guide our interpretation of the two kinds of *Triebfedern*, grounds of desire, that are admitted into the character-constituting maxim: sensible and moral. To act on a sensible ground is to act for a *kinetic* end, an end at which the action aims and that is accomplished when the action is completed. To act on a moral ground is to act for an *energetic* end, an end accomplished in every instance of the exercise of the capacity for *energetic* activity. Specifically, it is to act in recognition of the moral law. Finally I suggest a reading of Kant's claim that in the *Gesinnung*-constituting maxim the two kinds of *Triebfedern* are subordinated to one another, as a claim concerning the constitution of the structure of motivation whereby the pursuit of one kind of end—sensible-*kinetic* or moral-*energetic*—is subordinated to the pursuit of the other.

Good and bad characters reflect the way these *Triebfedern* are subordinated to one another, namely, which kind of *Triebfeder* is made supreme, or the pursuit of what kind of end is made the condition of pursuing an end of the other kind. The *Gesinnung*-constituting maxim determines whether a person is of good or bad character by determining which of the two kinds of *Triebfedern* is subordinated to the other. Moral goodness consists in respect for the moral law

being admitted “as an incentive [*Triebfeder*], sufficient by itself, of the power of choice” (R 6:27). Moral evil consists in the sensible *Triebfedern* being admitted into the *Gesinnung*-constituting maxim “as by themselves sufficient for determining the power of choice” (R 6:36). I suggest that moral goodness should therefore be understood as actions grounded in a motivational structure whereby the pursuit of particular, sensible-*kinetic* ends is subordinated to the pursuit of the moral-*energetic* end, namely, the activity whose principle is the moral law. Evil character should be understood, on the contrary, as giving in one’s *Gesinnung*-constituting maxim sensible *Triebfedern*, i.e., the pursuit of *kinetic*, particular ends, *absolute* priority in action.

In Chapter Four I attend to a possible tension in Kant’s account of moral imperfection. The general account of evil character—a state of systematic, stable selfishness—requires giving sensible grounds absolute priority in action. This account of evil leaves, it may seem, no room for Kant’s own, important, account of the levels (*Stufen*) of evil: frailty, impurity, and wickedness. First, the definition of wickedness, the third and highest level, is *identical* to the general definition of evil: wickedness “reverses the moral order in regard to the incentives [*Triebfedern*] of *free* power of choice” (R 6:30). Second, the other two levels resist the general definition: frailty is not a stable disposition, and impurity requires the cooperation of sensible incentives [*Triebfedern*] but does not prioritize them. Both seem therefore not to be cases of evil character, at all. Philosophically, dismissing frailty and impurity would exact a high price. First, they define fundamental phenomena of moral life, and, second, they are necessary to make sense of differences among characters that are not wholly good but are not wholly, stably, bad, either—a distinction necessary to account for moral development.

I argue that we should understand frailty and impurity as states of moral *immaturity*, the condition that precedes the acquisition of a stable moral character. To substantiate this claim I argue that the idea of the *acquisition* of character must play a wholly different role in Kant's practical philosophy than that accorded to it on standard readings. The acquisition of character emerges as a rational accomplishment: the development and determination of our uniquely rational capacities for feeling.

The resulting account of feeling and moral character resolves the problem of moral motivation that has vexed much of the Kant literature, and offers an understanding of practical cognition according to which our capacity for feeling and our capacity for having a character grounded in feeling are not obstacles to ethical life, but the necessary conditions of freedom.

*

After Kant, this essay is most intellectually indebted to the thought of Martin Heidegger. This is true in many ways, of which I will mention only two. First, Heidegger was the first to unequivocally recognize the centrality of moral respect to Kant's own philosophical project as well as the depth of the insight that Kant offers with this account. "Kant's interpretation of the phenomenon of respect," Heidegger pronounced, "is probably the most brilliant phenomenological analysis of the phenomenon of morality that we have from him."¹ Heidegger claimed that the moral I, which constitutes the "authentic self and essence of man," is constituted essentially by the feeling of respect, which is not only constitutive of moral action ("respect for the law means *eo ipso* action"),² but moreover "must present a way of Being-self-

¹ Martin Heidegger, *The Basic Problems of Phenomenology*, trans. Albert Hofstadter (Bloomington: Indiana University Press, 1981), 133.

² *Ibid.*, 137.

conscious”.³ Kant’s moral respect, Heidegger averred, is “the true mode in which man’s existence becomes manifest”.⁴ But Heidegger did not explicitly treat Kant’s practical works in anything like the detail with which he considered his theoretical philosophy. This essay constitutes, among other things, an attempt to recover a reading of Kant that elucidates and justifies Heidegger’s suggestive but brief remarks on this topic.

The second way in which this essay is beholden to Heidegger is in its philosophical method. Even a brief overview of the project is likely to invite the reader to wonder whether I have not strayed so far from Kant’s text that it is uncertain—whatever philosophical merits the account may have—whether or not it properly serves as an interpretation of Kant. I trust the analysis of the evidence will substantially assuage a significant number of these worries, and I try to show throughout how the arguments I bring to bear are not external to but rather lie at the heart of Kant’s own philosophical commitments. Yet behind this concern lies a deeper one regarding the meaning of the activity of interpreting a work of philosophy. How could an interpretation that is self-avowedly radical claim to be an adequate elucidation of a philosopher’s own project?

In his interpretation of the *Critique of Pure Reason* Heidegger argues for the radical claim that underlying sensibility and understanding there exists a third fundamental faculty, the faculty of imagination, which constitutes the common root from which sensibility and understanding stem. Heidegger defended the validity of his interpretation qua interpretation in the following terms:

³ Martin Heidegger, *Kant and the Problem of Metaphysics*, trans. Richard Taft (Bloomington: Indiana University Press, 1962), 110.

⁴ Martin Heidegger, *The Basic Problems of Phenomenology*, trans. Albert Hofstadter (Bloomington: Indiana University Press, 1981), 137.

For us, of course, the inquiry [...] does not originate in an unfounded inclination toward radicalization, only in order simply to excel Kant in this respect. [...] But at the same time it will be shown that everywhere we have actually added too much to Kant and that in fact the problem which he poses is not so explicit and original as we have presented it. Our interpretation first achieves a clarification of the whole, in order then to see sharply and in retrospect precisely what in Kant is unbalanced and questionable. This will make it possible that coming to terms [with Kant] addresses the crucial difficulties. Not only will we let Kant say what he intended to say; but we shall let him say more, even if this “more” rests only on philosophical interpretation. Then taking back later the “more” that we have added to Kant can readily be done on the basis of the philosophical understanding that we have achieved. However, it is difficult if not impossible to aim directly at what Kant wanted to say, unless we would simply repeat Kant’s sentences.⁵

In the reading of Kant’s practical philosophy that I present in this essay I have tried, to the best of my abilities, to follow the method of interpretation gestured at here.

⁵ Heidegger, *PICPR*, 64.

Chapter 1

Overcoming the Dualism of Reason and Sensibility in Kant's Practical Philosophy

§1. Introduction

Kant's practical philosophy is grounded in the following claim: absolutely good activity, activity that is good in itself and not merely relative to an end whose goodness is independent of the goodness of that activity, is grounded in that activity having its self-conscious source in reason. Absolutely good activity is not determined by contingent laws of nature, but by necessary laws that command universally, and this can only be achieved if the activity stems from a consciousness of necessary universal legislation.

For Kant, reason is the “faculty of principles” (A299/B356). Desire is “a being’s faculty to be by means of its representations the cause of the reality of the objects of these representations” (*KpV* 5:9n).¹ A being’s faculty of desire is a being’s capacity to be causally

¹ ABBREVIATIONS. In the case of the *Critique of Pure Reason*, I follow the standard practice of referring to the 1781 (A) and 1787 (B) editions. For all other texts, citations appear in the order of abbreviation, volume number, and page number from the *Akademie Ausgabe* (AA), *Kants Gesammelte Schriften*, edited by *Königlich Preussische Akademie der Wissenschaften* (29 vols. Berlin: de Gruyter, 1900–). All translations come from *The Cambridge Edition of the Works of Immanuel Kant*, edited by Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1992–).

- G* *Grundlegung zur Metaphysik der Sitten* (AA 4), *Groundwork for the Metaphysics of Morals*, trans. Mary Gregor.
KpV *Kritik der praktischen Vernunft* (AA 5), *Critique of Practical Reason*, trans. Mary Gregor.
KU *Kritik der Urteilskraft* (AA 5), *Critique of the Power of Judgment*, trans. Paul Guyer and Eric Matthews.
MS *Die Metaphysik der Sitten* (AA 6), *The Metaphysics of Morals*, trans. Mary Gregor.
MPC *Moralphilosophie Collins* (AA 27), *Moral Philosophy Collins [Lectures on Ethics]*, trans. Peter Heath.

efficacious through representations, i.e., it is the capacity to act. It follows that if we are to be capable of absolutely good activity, our faculty of desire must be one that is efficacious through the representation of principles. Kant calls this faculty a “will.” The will is therefore the capacity to be the cause of the objects of representations that are themselves representations of principles. Thus, it is the capacity to be cause of the reality of principles. Alternatively, we can say that the will is the faculty of reason in its specifically practical employment. Reason, both in its theoretical and its practical employment, is autonomous, which means that it is self-determining. For Kant, therefore, good activity is that through which the subject determines herself: it is spontaneous.

However, Kant also acknowledges that acting well involves receptive capacities. As finite, rational beings, we are not self-sufficient but require external goods for our well-being. Receptive awareness is required in order have knowledge of the particular material conditions necessary for life. To deny the involvement of receptivity in the exercise of our will would therefore render mysterious how good action could be reliably directed towards the concrete conditions of our well-being—let alone be adequate for obtaining them.

By material conditions of well-being I do not merely mean to refer to the material conditions necessary for the *execution* of a command of reason. Obtaining *these* would of course require theoretical receptive capacities, since it would require theoretical cognition of objects, objects that do not depend for their reality on our activity of representing them. Moreover, Kant acknowledges that we cannot be required to do what is strictly impossible (“the action to which

R *Die Religion innerhalb der Grenzen der bloßen Vernunft* (AA 6), *Religion Within the Boundaries of Mere Reason*, trans. George Di Giovanni.

the ‘*ought*’ applies must indeed be possible under natural conditions,” KrV, A548/B576).² Since you are required to perform an action only if it is possible for you to do so, it follows that the material conditions that are necessary for performing a required action are always sufficient. What I mean to refer to, therefore, is the involvement of receptivity that is necessary for orienting the person with respect to the material conditions of their own well-being.

Any reading of Kant’s practical philosophy that seeks to do justice to both elements—the autonomy of reason, on the one hand, and the necessary involvement of receptivity in our activity, on the other—faces a challenge: how to conceive of this involvement of receptivity in our willing while preserving Kant’s insistence on a sharp distinction between the autonomous good will and the heterogeneous inclinations as distinct sources of ends and motivation.

In the literature, the problem is most perspicuously acknowledged in expositions of its canonical solution: known in the literature by the name given to it by Henry Allison, “Kant’s Incorporation Thesis” (henceforth IT). In what follows I evaluate and reject IT as philosophically inadequate on grounds that belong to Kant’s own account of theoretical cognition. From consideration of the Incorporation Thesis there emerges a constraint on any adequate solution to the problem: It must give an account of the involvement of receptive capacities in a rational being’s determination of action. A description of the activity of a rational being cannot begin with or include a role for a wholly receptive capacity, intelligible independently of the subject’s possession of rational capacities. In particular, a description of the activity of a rational actor cannot begin with or include a role for a capacity to “sensibly desire” that does not presuppose an essential connection to the exercise of reason. Rational desire is not constituted by a capacity to monitor and regulate our sensible desires, or pick and choose among them in accordance with

² Cf. *Religion*, 6:50.

rational requirements. The power of rationality, it will emerge, is not something that could be added to a mind that already forms a system that is independently intelligible of this addition.

Insofar as Henry Allison's "Incorporation Thesis" depends on a distinction between the rational will and pathological desire understood as *wholly* sensible, i.e., whose exercise does not involve the exercise of rational capacities, we must reject it. By the end of the chapter, we will come to appreciate that both philosophically and as a matter of local exegesis the viability of Kant's practical philosophy depends on there being an alternative approach. I will argue for just such an alternative in subsequent chapters.

§2. The Incorporation Thesis

The question of how to understand the involvement of receptive capacities in the exercise of a capacity for rational activity receives a unique inflection in Kant interpretation. It is usually introduced in the context of two central challenges facing Kant's account.³ First, to explain how we are responsible for acting badly. Secondly, to explain the way in which we may act well while pursuing a particular end our orientation to which depends on receptivity. I articulated the general form of the second concern above. The difficulty of the first challenge is as follows. An agent performing morally evil action is motivated, by definition, not by the moral law, the necessary law of freedom, but by her contingent inclinations. If this heteronomy means that evil action is simply caused by mechanistic forces, it would amount to a mere happening, for which it

³ Tamar Schapiro claims that it is usually introduced "as a way of correcting a common misreading of Kant's account of heteronomous action." That misreading, namely the claim that we are not responsible for our actions when we do not act from duty, is not common among Kant interpreters since its denial is so absolutely central to Kant's project. Tamar Schapiro, "Foregrounding Desire: A Defense of Kant's Incorporation Thesis," *J Ethics* 15, no. 3 (September 2011), 148.

would be nonsense to hold a rational agent accountable.⁴ If the capacity to act freely requires self-determination, this challenge amounts to a request for elucidation of the sense in which an agent be thought of as freely determining herself to act unfreely.

Allison was the first to draw attention to a passage in Kant's late work, *Religion within the Bounds of Bare Reason*, that potentially key to resolving this problem. Kant claims there that "the power of choice [*Willkür*] has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*] except *insofar as the human being has admitted the incentive [*Triebfeder*] into his maxim [nur sofern der Mensch sie in seine Maxime aufgenommen hat]* (has made this a universal rule for himself, according to which he wills to conduct himself)" (R 6:23–4).

Allison bases IT on an interpretation of this passage. According to Allison's reading, actions, whether good or evil, are "taken as genuine expressions of agency, and therefore, as imputable" insofar as "they are thought to involve an act of spontaneity on the part of the agent, through which the inclination or desire is deemed or taken as an appropriate basis of action. [...] this occurs by subsuming the inclination or desire under a practical rule or principle."⁵ Whether the action is good or evil therefore depends on whether this subsumption is warranted.

Proponents of IT celebrate it as a correction to those interpretations of Kant's practical philosophy that posit a crass dualism, according to which the faculties of sensible and rational desire are entirely conceptually distinct from one another and operate independently of one another in the practical subject. In this picture, contingent motivations play no role in the determination to good action, and furthermore it is left mysterious how could anyone be held

⁴ Ibid., 148.

⁵ Henry E. Allison, *Kant's Theory of Freedom* (New York: Cambridge University Press, 1990), 39.

responsible for acting poorly. IT promises to assuage both worries by reconciling the possibility of acting from inclination with the self-determination required for responsibility.

However, the difficulty Allison is responding to directly is the second: the possibility that free action is heteronomous. The Incorporation Thesis is supposed to show us how to reconcile the possibility of evil action with the fundamental Kantian claim that “spontaneity is an ineliminable component in rational agency.” The spontaneity of this alleged act of deeming or taking desire as an appropriate basis of action secures the agent’s freedom and therefore the action’s imputability:

This spontaneity, which is the practical analogue of the spontaneity of the understanding, may be characterized provisionally as the capacity to determine oneself to act on the basis of objective (intersubjectively valid) rational norms and, in light of these norms, to take (or reject) inclinations or desires as sufficient reasons for action. According to this model, then, the intentional actions of a rational agent are never “merely” the causal consequences of the agent’s antecedent psychological state (or any other antecedent conditions for that matter) but require, as necessary condition, an act of spontaneity.⁶

IT has come to be considered as a canonical reading of what, according to Kant, is “necessary for a finite rational being like ourselves to make the transition from having a desire to acting on it.”⁷ That is, IT has been understood as supplying an account of the way in which we may act well while pursuing an end that is revealed to me when I find myself with a particular desire.

⁶ Allison, *Kant’s Theory of Freedom*, 5.

⁷ Schapiro, 148.

Interpreters are in nearly unanimous agreement that Allison's interpretation of the passage captures a central Kantian commitment and most agree that Kant was right to hold it.⁸ I will argue that on the contrary Kant had no such commitment, and would have erred if he had.

§3. The Common Assumption of Interpretations of the Incorporation Thesis

A single common assumption grounds the various interpretations of the IT: an account of practical rationality that has room for an account of an independent capacity to desire ("sensible desire") whose operations can be defined without reference to its possessor being the bearer of rational capacities.

Allison's commitment to this assumption is implicit in his definition of "the relationship between rational agents and their inclinations or desires,"⁹ whereby spontaneity makes an appearance only in the act of subsuming an inclination under a practical rule.

Christine Korsgaard follows Allison in attributing to Kant a picture in which desire is located outside of reason and must be endorsed by reason if it is to be acted upon:

He says, "We cannot conceive of a reason which consciously responds to a bidding from the outside with respect to its judgments." If the bidding from outside is desire, then the point is that the reflective mind must endorse the desire before it can act on it, it must say to itself that the desire is a reason. As Kant puts it, we must *make it our maxim* to act on the desire. Then although we may do what desire bids us, we do it freely.¹⁰

⁸ Cf. Allen W. Wood, *Kant's Ethical Thought*, Modern European Philosophy (Cambridge: Cambridge University Press, 1999), 51–53. Christine M. Korsgaard, *The Sources of Normativity* (Cambridge: Cambridge University Press, 1996), 94, and *Self-Constitution: Agency, Identity, and Integrity* (Oxford: Oxford University Press, 2009), 105 and 115. Thomas E. Hill Jr., *Dignity and Practical Reason in Kant's Moral Theory* (Ithaca, NY: Cornell University Press, 1992), 86.

⁹ Allison, *Kant's Theory of Freedom*, 39.

¹⁰ Korsgaard, *Sources of Normativity*, 94.

Tamar Schapiro puts the point succinctly: “When I have an inclination, the reflecting part of me is aware of the non-rational principle that shapes my inner animal’s way of seeing and responding to the world.”¹¹

Allen Wood distinguishes the faculty of desire in brutes and in rational beings as essentially two different powers of self-determination to action, but nevertheless reserves an independent role for what he calls “empirical impulse,” an account of which “applies equally to rational beings and to brute animals.”¹² The two kinds of powers of desire, the brute and the rational, are distinguished, according to Wood, by how they respond to the alleged impulse. Whereas for the mere brute the power of choice is determined mechanically by the sensuous, empirical impulse (i.e., the empirical impulse is necessary and sufficient for action), rational beings are able to resist such impulses and moreover, when their faculty of desire does act on them “it is never determined by them mechanically.”¹³

Although they aim to secure the intelligibility of the interaction between our receptive and spontaneous powers I will argue below that proponents of IT have succumbed to the threat of a philosophically unacceptable dualism.

§4. Critique of IT: Philosophical Considerations

In its various guises IT attributes to Kant the claim that a (motivationally-laden) representation whose occurrence does not itself require the exercise of rational capacities, can nevertheless supply a putative reason that is available to the reflection and intervention of a rational subject. The subject can reflect upon the permissibility of pursuing the end suggested by

¹¹ Schapiro, 165

¹² Wood, *Kant’s Ethical Thought*, 51.

¹³ *Ibid.*

the impulse, inclination, or desire, and can determine whether to gratify the impulse by following its motivational lead, or else to deny it satisfaction. I claim that this is impossible and that proponents of IT therefore saddle Kant with a philosophically untenable position.¹⁴

The contemporary critique of additive theories of rationality takes as its starting point the dogma of the Myth of the Given. The Myth of the Given appears in the history of philosophy in various forms. But it first appears with this title in the context of a debate concerning theoretical knowledge:

McDowell, who takes himself to be following Sellars, describes the dogma as follows:

Givenness in the sense of the Myth would be an availability for cognition to subjects whose getting what is supposedly Given to them does not draw on capacities required for the sort of cognition in question.¹⁵

To challenge this dogma is to challenge the idea that sensibility can make objects available to cognition without the involvement of rational capacities—that sensibility can make objects available in a way that renders intelligible our capacity to make empirical judgments about the world based on our perceptual experience. If one is to attribute an epistemic role to sensory experience—one that attributes to experience the capacity to warrant or justify belief—one must therefore characterize it as a capacity whose exercises are *essentially* able to stand in rational relations. This specifically precludes the possibility that experiences understood in merely causal terms can play any epistemic or justificatory role in our thinking.

¹⁴ In what follows I rely on Matthew Boyle’s “Additive Theories of Rationality: A Critique.” I try to be brief and refer readers to that article as well as his “Essentially Rational Animals” for details. See also: Matthias Haase, “Life and Mind,” in *The Freedom of Life: Hegelian Perspectives*, ed. Thomas Khurana, (Berlin: August Verlag 2013), 69–109.

¹⁵ John McDowell, “Avoiding the Myth of the Given,” in *Having the World in View: Essays on Kant, Hegel, and Sellars* (Cambridge, MA: Harvard University Press, 2013), 256.

Matthew Boyle approaches this problem by inquiring after the kind of unity that constitutes the faculties of a rational subject.¹⁶ He criticizes a formal structure that belongs to what he calls “additive theories of rationality.” These theories hold that accounts of the mind can begin with or include an independent account of what it is to perceive and desire. Additive accounts do not presuppose within their characterization of perception and desire any essential connection between these capacities and the rational capacities of empirical and practical judgment. Rationality is understood as a capacity for rational thought that can be added on after the fact as an independent capacity to “monitor” and “regulate” our “believing-on-the-basis-of-perception” and “our acting-on-the-basis-of-desire.”¹⁷ Thus such theories regard the power of rationality as something that might be “tacked on” to a mind that already forms an *system that is intelligible apart from this addition*. The added rational capacity can interact with our capacity for perception and desire (it can reflect on its deliverances and, in the case of practical judgment, choose which impulses, inclinations, or desires to follow and which to deny satisfaction).

Proponents of IT, as we’ve seen, attribute an account of just this structure to Kant. Boyle argues that this approach faces a fundamental difficulty which he calls the Interaction Problem (4.1)¹⁸ After I introduce this problem in abstract terms I will examine how it applies specifically to Kant’s account of sensible desire (4.2).

¹⁶ Matthew Boyle, “Additive Theories of Rationality: A Critique,” *European Journal of Philosophy* 24, no. 2 (2016).

¹⁷ Boyle, “Additive Theories of Rationality,” 527.

¹⁸ Boyle likewise defends the claim that additivist accounts face what he calls “the Unity Problem.” This concerns “how an additive theorist can explain a rational subject’s entitlement to conceive of the animal whose perceptual and desiderative life he oversees as ‘I’ rather than ‘it.’” The rational subject is understood as the subject of its beliefs and agent of its actions, its sentient capacities, and hence its relation to its perceptions and desires, are rendered alien to it. It is “a difficulty about how to account for the intuitive idea that the same subject is both a certain animal and the subject who thinks.” Boyle, “Additive Theories of Rationality,” 527. I leave this challenge to the side here because at least some of

§ 4.1. The Interaction Problem

The Interaction Problem concerns “how capacities conceived as intrinsically independent of the power of reason can *interact* with this power in what is intuitively the right way.” (It is important to note that the interaction problem does not imply that a solution must, strictly speaking, secure interaction between parts. I will go on to argue that the distinctions possible and necessary in an account of practical cognition are not best thought of as distinctions between parts at all.) As it pertains to IT, the problem is as follows: According to IT, my desiring X can present me with a reason to act which I can step back from, reflect upon, and ask myself, in Korsgaard’s formulation, “Is this desire really a reason to act?”¹⁹ If this question is to make sense, we must presuppose that a desire *can* be found to constitute a reason to act. Or, in other words, my desire for X could figure in the explanation of why I acted as I did (e.g., I felt like doing X, upon reflection saw it would be permissible to X and that therefore I X-ed).

Proponents of IT hold that our faculty of desire is in essence identical to that of a non-rational animal, one which involves no capacity to consider whether there is reason to pursue X. The subject’s desiring X is not of itself an exercise of a capacity to represent X as an end that there is reason to pursue, i.e., to desire X is not even to putatively consider it choice-worthy: worthy of pursuit in a sense that admits of rational justification. One way of putting the point is this: if inclination is understood as merely a brute fact, it can be experienced *as* nothing but itself a brute fact, akin to finding oneself pulled in one direction rather than another. Brute pull can perhaps be observed with approval or disapproval by a rational subject, as when one finds herself

the proponents of IT embrace, at the very least on behalf of Kant, the picture the Unity Problem alleges is unacceptable.

¹⁹ Korsgaard, *Sources of Normativity*, 92–93.

blown away by a hurricane. One can even try to resist it or resign oneself to it, and even on rational grounds (e.g., resistance would too costly, or, for independent reasons, that pull can be gratified without undesired consequence, or if, for independent reasons, it could promote an end of mine) but the act of resignation would not amount to deeming that brute pull itself *a reason to act*.

Alternatively, if we are to think of the practical subject's desires as the motivationally-laden representations of X, where X is represented as an aim there is reason to pursue (pending further considerations), i.e., where X is represented as possibly choice-worthy, then her desires would bear a structure that a brute animal's would necessarily lack. If that is true, receptive practical experience cannot be described in merely causal terms and then, in a second step, be made to assume the role of a "reason" that can warrant and justify action.

Proponents of IT face a problem of meeting this challenge because they hold that a rational subject's faculties of perception and desire are not intrinsically different from the faculties of perception and desire in a non-rational animal, faculties that lack the power to assess whether there is reason to judge that p or pursue X. Additive theorists must therefore hold that a rational subject's perceiving p is not *itself* a reason for her to judge that p, and that a rational subject's desiring X does not *itself* engage her capacity to represent X as an aim there is reason to pursue. This account calls for an explanation of how a perception or a desire could ever come to be considered as a reason for judgment or action: an explanation the Interaction Problem threatens to render unintelligible.

These arguments, if valid, establish that Kant would have erred in being committed to IT. But they leave open the question of whether or not he in fact held it or was otherwise committed

to it implicitly. A full answer to this question will emerge in the subsequent chapters of this essay, in an examination of his accounts of moral respect and moral character. In what follows I would like to offer preliminary arguments for the plausibility of rejecting IT *on behalf of Kant*.

§5. Engstrom's Account of IT

Stephen Engstrom offers a version of IT which might seem to avoid the difficulty I've presented under the name of the Problem of Interaction, namely, that an additive account cannot make sense of how a purely sensible desire can, of itself, present practical reason with a reason to act. Engstrom's interpretation seeks to avoid this by stipulating that sensible desire does not of itself present putative reasons, which the subject steps back from and evaluates: it is rather an act of practical reason in which the object of sensible desire, i.e., the agreeable, is deemed good.

This activity of the subsumption of the agreeable under the good is manifest on particular occasions but is grounded in two moments of subsumption:

1. The subsumption of the agreeable under the concept of the good in a specific practical judgment. These judgments follow the following principle: "*To the extent that it can be an object of practical knowledge, what one finds agreeable is simply good.*"²⁰
2. The subsumption of the agreeable under the concept of the good in the fundamental practical judgment of the wish for happiness. The particular practical judgments in which particular agreeable objects are deemed simply good (namely, insofar as they can be the objects of practical knowledge), depend on "a primitive act of practical judgment, in which what one finds agreeable *in general* is made into an object of one's will—whether

²⁰ Engstrom, Stephen. *The Form of Practical Knowledge: A Study of the Categorical Imperative* (Cambridge, MA: Harvard University Press, 2009), 69.

absolutely or only in such a way and to such an extent as is compatible with this judgment's having the form of practical knowledge—through being brought under this concept of the simply good.”²¹

However, making the capacity to confer reasonableness on desire the sole preserve of practical reason will not suffice for meeting the challenges raised above.

§ 5.1. Conceptual Challenge

If the capacity to feel pleasure and pain is associated with the lower faculty of desire and presupposes no involvement of spontaneity, how could its operation provide representations that are so much as able to suggest possible objects for practical cognition? In attempting to address this concern Engstrom argues that although the faculty of sensible desire forms an intelligible system independent of the exercise of rational powers, its form is intelligible to practical reason because the form of the agreeable bears an affinity to the form of the good, namely, by displaying the structure of an end. In this picture reason finds the agreeable determined by a form of its own and in a practical judgment informs sensibility by giving sensible matter “*new form*.”²²

This raises the following concerns: First, what is the source of the “form” of the agreeable? Secondly, how can the practical thinker access this form, so as to then re-form it in a further step?

A desire, the sort of thing I am ordinarily said to have acted on (and sometimes said to resist), is not merely a matter of having the practical equivalent of a “pure” sense-impression. In

²¹ Engstrom, *The Form of Practical Knowledge*, 70.

²² Engstrom, “Kant on the Agreeable and the Good,” in *Moral Psychology*, ed. Sergio Tenenbaum, Poznan Studies in the Philosophy of the Sciences and the Humanities, vol. 94, no. 1 (Amsterdam: Rodopi, 2007), 152.

the primary case desire is for intelligible ends. If by the practical equivalent of a mere sense-impression we mean something momentary and singular, as we must think of the merely agreeable, the desire to do something cannot be just a matter of the momentary and singular affections of sensory consciousness that are afforded by the sensation of pleasure. Rather, the idea of an object of desire, i.e., an end, is the idea of something that outstrips the momentary pull of a brute “this” or “more.” An end of an action is something that, for instance, can be pursued in different ways, which are systematically related to one another. An end displays a certain kind of unity, namely a teleological one. If desire is to be for ends, it must contain a consciousness of this unity. It must be part of the description of the desire that what it aims at is an end, which can be achieved by the activity guided by means-ends reasoning. For instance, when I see a cup of tea in front of me, there is a sense in which my desire is directed to the particular cup in front of me. But if what I want is a cup of tea, the description of what I want cannot be restricted to the particular sensible object in front of me, one which I have been empirically habituated to expect would afford me some pleasure. The end of my desire is hydration, or sociability, or, indeed, pleasure via the experience of warmth or a certain flavor that I appreciate, any of which I may fulfill in alternative ways using alternative means. And this is, at least implicitly, part of my desiring consciousness. In desiring the cup of tea in front of me I am aware of desiring an end that I may obtain in another way.

Desires exhibit, in other words, a certain kind of unity and generality, that enables them to be articulated in the expression of intentions, or in self-conscious practical knowledge of what I am doing. What is crucial is that this unity cannot be simply the function of the impressions the mind receives. It stands over and above the sensible impressions themselves, which are the result

of the mind being affected. This feature, the unity or form that characterizes desire, cannot be itself a product of the mind's being affected.

This renders opaque the claim that sensible desire and its object, the agreeable, bear a quasi-teleological form of their own that is intelligible without reference to its possible (re)forming by spontaneity. For the source of this form remains mysterious.

Assuming that sensible desire thus characterized can enter into an intelligible relation with a rational capacity, whether or not the object can receive this form is dependent on the desire's sensible matter standing in agreement with the form of practical law. This agreement would be of course accidental: it would be neither constitutive of nor of any normative significance to sensibility itself.

These considerations raise the following concern: if sensible desire is understood as a faculty whose form is utterly heterogeneous to reason why would practical reason judge the object of sensible desire to be the object of the will? Why would it, in other words, confer upon it a reason-supplying status?

§ 5.2. Exegesis: Kant's Characterization of Sensible Desire

To substantiate his claim Engstrom points to the following passage:

The agreeable, which as such represents the object solely in relation to the senses, must first be brought under principles of reason through the concept of an *end* in order to be, as object of the will, called good. (*KU* 208)

This involves, according to Kant, the exercise of reason "which looks beyond to the consequences" (*KU* 208) and considers the object independently of its appeal to the senses but with a view only to the object's possible contribution to ends (or means to ends) that are set by

reason itself. Engstrom reads the passage to justify both moments of subsumption that I introduced above: the concrete and the general.

However, Kant's characterization of the agreeable raises concrete challenges to the claim that his talk of bringing the agreeable under principles of reason through the concept of an end is to be understood in the terms of a *general* practical subsumption of the agreeable under the concept of the good.

According to Kant, the imperatives of practical reason have for their object the good, while sensible inclinations have for their object the agreeable. Kant's respective characterization of the agreeable and the good renders it difficult to see how the former could be deemed worthy of pursuit, and not just because it is difficult to see how the desire to pursue the end could be taken for a reason to pursue it. Kant's characterization of sensible desire and its object also makes it difficult to see how the exercise of sensible desire could yield anything that could be experienced as an intelligible desire to step back from—let alone to deem reasonable in a second step.

A minimal requirement for rational pursuit is an understanding of the course of action undertaken. Thus Kant writes, "In order to find something good, I must always know what sort of thing the object is supposed to be, i.e., I must have a concept of it" (*KU* 207). However, the agreeable, according to Kant, "represents the object solely in relation to sense" (*KU* 5:208), and "does not serve for any cognition at all, not even that by which the subject cognizes itself" (*KU* 5:206). In particular, Kant denies that "any judgment about the constitution of the object belongs...to that which is agreeable." It is hard to see how the object of the agreeable, as such, or

on particular occasions, could ever be deemed reasonable to pursue in a second step. This is true for two reasons.

First, the motivation for pursuing the agreeable end will remain completely divorced from anything that might render that end worthy of pursuit. According to Kant, the agreeable indicates nothing but the matter of the faculty of sensible desire (*KpV* 5:24). The satisfaction in the agreeable consists in a feeling “implanted in us” (*KpV* 5:24), and, accordingly, sensible desire’s “material determining ground can be cognized only empirically by the subject” (*KpV* 5:25).

Second, the relation of practical judgment to purely sensible desire relies on sensible desire providing the subject with some sort of intelligible awareness of a possible end. This awareness, Engstrom claims, is of a peculiar kind:

The characterization of pleasure as the consciousness of a representation’s causality should not be understood as implying that pleasure itself is a type of cognition or involves the application of concepts (e.g., of cause, or representation), for as a type of feeling, pleasure belongs to receptivity, not spontaneity; the consciousness in which pleasure is said to consist is better regarded as the sensible indication, or inward manifestation, of the self-sustaining character of a state of mind.²³

To describe the kind of awareness by which the object of the agreeable is purportedly made available to us, Engstrom uses notions like “sensible indication” and “inward manifestation.” The relation of these apparently *sui generis* forms of awareness to familiar forms of awareness, namely forms of self-conscious awareness, grounded in the spontaneous unity of apperception, is left undetermined. As such it is obscure how this kind of awareness is supposed to be accessible from the standpoint of a practical judgment.

²³ Engstrom, “Kant on the Agreeable,” 142, footnote 25.

It moreover raises the question whether the account of animal sentience can be arrived at independently of a consideration of rational cognition. Presumably our (formal) concept of representation is one that is available to us on account of the self-conscious character of the faculty of representation in us. The idea of representation (better, presentation) that is not self-conscious is arrived at by abstraction in thought of the character of this cognitive faculty as self-conscious.

This challenge rehearses the challenge Engstrom identifies and answers for theoretical cognition: it is not clear how the power of reason can take up the material provided by wholly receptive sensation, and this threatens to render the idea of the involvement of content in the exercise of practical reason in us unintelligible.

It appears difficult to preserve both Kant's characterization of sensible desire and its object and his claim that the agreeable must be subsumed under principles of reason in order to become an object of the will. The difficulty is resolved if we allow ourselves to read Kant here as speaking not of a general subsumption of all agreeable objects, insofar as they are fit objects of willing, under the concept of an object of the will, but simply of a putative agreeable object. Kant means that in order for an agreeable object to be the object of practical judgment, that object, which happens to be agreeable, must be referred to an end. In other words, it is not qua agreeable (not even qua agreeable and otherwise fit to serve as the object of the will) that I pursue the object. Kant's examples substantiate the interpretation.

First, he is not concerned with the possibility of pursuing ends but with different ways of justifying our actions, namely, by referring them to pleasure mediately, through reason or, immediately, through the senses. For example:

Of a dish that stimulates the taste through spices and other flavorings one may say without hesitation that it is agreeable and yet at the same time concede that it is not good; because while it immediately appeals to the senses, considered mediately, i.e., by reason, which looks beyond to the consequences, it displeases. (*KU* 5:208)

Kant's final example, that of the pursuit of happiness, is especially illuminating for our purposes:

In respect to happiness, finally, everyone believes that the greatest sum (in terms of number as well as duration) of the agreeableness of life can be called a true good, indeed even the highest good. But reason also balks at this. Agreeableness is enjoyment. But if this were all that is at stake, then it would be foolish to be scrupulous with regard to the means for providing ourselves with it, that is, whether it is obtained passively, from the generosity of nature, or through self-activity and our own effort. But that the existence of a human being who lives merely for enjoyment (however busy he might be in this respect) should have a value in itself, even if as a means to this he was as helpful as possible to others who were likewise concerned only with enjoyment, because he participated in all gratification through sympathy: of this reason could never be persuaded. (*KU* 5:208)

An account where the wish for happiness is adopted under condition of goodness, whereby our practical efforts aim at the gratification of sensible inclinations that please not mediately by reason but wholly immediately through the senses, albeit only insofar as they conform to rational form, amounts to a life of permissible enjoyment. Kant continues,

Only through that which he does without regard to enjoyment, in full freedom and independently of that which nature could passively provide for him, does he give his being as the existence of a person an absolute value; and happiness, in all the fullness of its agreeableness, is far from being an unconditional good. (*KU* 5:208–9)

Recall that in Engstrom's picture the object of sensible desire, the agreeable in general, is set in the primitive, fundamental, person-constituting act of practical judgment as the will's own end. It

is by this primitive act of practical judgment that a person constitutes herself as such.²⁴ As I have argued, the awareness afforded by the feeling of pleasure and displeasure associated with the agreeable is an awareness of a *blind* disposition. One which I can *cognize* only empirically, something I discover in experience, and that as something that is implanted in me. It is difficult to see why, from the standpoint of pure reason, the agreeable should become a legitimate concern for it at all. How can we understand practical cognition determining *itself* to pursue these ends?

Engstrom explicitly proposes one possible answer, offering an account of why practical judgment might require recourse to the agreeable in his account of how judgments of choice *require* the judgments of wish. Judgments of wish make possible “comparisons” between the jointly practicable goods, comparisons that become necessary once the empirically specified limits of practicability are taken into account. But the formal requirement for a standard of arbitration does not address the question: Why is the standard of comparison between different possible choices, possible courses of action in the world, to be supplied by the agreeable?

It is tempting here to draw on our theoretical knowledge that physical pains and pleasures, and precisely those implanted in us, indicate “natural” purposes, and thus indicate, at least sometimes, things that can be recognized as good or bad for us from the standpoint of reason (our health, for example, according to Kant, can be an object referred by reason to a practical end). This is not in dispute and indeed any account of practical cognition must explain how our awareness of our bodily needs and any other kinds of needs imposed upon us by our finitude and material vulnerability inform our practical lives. However, a theoretical understanding of the possible function of the feelings one happens to experience would be purely

²⁴ Ibid., 81.

theoretical. Thus, this is not an understanding that is essentially available to practical reason so as to guide the formation of *the* primitive act of practical judgment. (I leave to one side the further concern that from the Kantian perspective a judgment concerning natural ends would be a teleological judgment, and thus a reflective, and not a determining theoretical judgment.)

The supposition that the sensible desire is intelligible independently of the involvement of rationality in its determination therefore makes it difficult to see *how* the good will can ever hear the voice of sensible desire, and *why* it should ever try to do so. These considerations begin, I hope, to give credence to the claim that Kant himself was not propounding a so-called additivist conception of rationality. Another consideration that speaks in favor of this claim is its consistency with (an interpretation of) Kant's underlying commitments in his theoretical philosophy.

§6. Alternative Reading of Kant's Statements in the *Religion*

To substantiate the rejection of Allison's influential reading of Kant, i.e., to reject what he names "Kant's Incorporation Thesis" as a valid interpretation of Kant, requires offering an alternative reading of those of Kant's statements in the *Religion* on which Allison bases his textual argument.

To recall, the key passage reads as follows: "the power of choice [*Willkür*] has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*] *except insofar as the human being has admitted the incentive [*Triebfeder*] into his maxim* (has made this a universal rule for himself, according to which he wills to conduct himself)" (R 6:23–4).

Attending to the context in which the question is raised suggests that Allison's interpretation is misguided. This context makes clear that Kant is not speaking of the subject's ability to step back from particular desires and assess whether they happen to provide adequate reasons for action, i.e., whether they can be made into objects of the will. This is because in the relevant section, Kant is speaking of admitting not a particular desire into one's particular maxim, but of admitting one of two *kinds of incentives* [*Triebfedern*] (namely, either sensible or moral) into a maxim of choice of *the utmost generality*.

The passage and surrounding discussion occur in the "Comment" to the introductory remarks of the "First Piece" of the *Religion*. The question that Kant is answering is decidedly different than the one that Allison puts forth IT as answering. What is at stake is a defense of a way of thinking that Kant identifies as "rigorism," and consists in the claim that "*The human being is (by nature) either morally good or morally evil*" (R 6:22). From this it follows that the human being is not "both simultaneously, namely in some points good, in others evil." Kant admits that experience appears to "confirm this intermediate position between the two extremes" and proceeds to offer a dual justification for the claim. First, he offers a general methodological defense whereby, he claims, the doctrine of morals resists "as long as this is possible," admitting moral intermediates, whether in actions or human characters, because "with such an ambiguity all maxims run the risk of losing their determinateness and stability" (R 6:22). Second, Kant offers a strictly conceptual defense of the rigoristic doctrine. He begins with the claim that Allison singles out, to the effect that the power of choice can be determined to an action by any incentive [*Triebfeder*] only insofar as the subject has admitted this incentive [*Triebfeder*] into her maxim. Kant proceeds to claim that if someone acts such that the law does not determine her

power of choice, then an incentive [*Triebfeder*] must have influenced her power of choice, something possible only if one admits the incentive [*Triebfeder*] into her maxim. From this he is able to immediately conclude that, in that case, the subject's "attitude [*Gesinnung*] in regard to the moral law is never indifferent (never neither of the two, neither good nor evil)" (R 6:24).

The *Gesinnung*, as Kant characterizes it in the *Religion*, is however not a particular maxim or fleeting attitude; it is instead the single rule that the power of choice "makes for itself." This is a maxim which is the "first basis [...] for the adoption of good maxims or the adoption of evil (unlawful) ones, and this moreover universally, as a human being" (R 6:21). Indeed, as Kant affirms later in the "Comment" itself, "The attitude [*Gesinnung*], i.e., the first subjective basis for the adoption of maxims, can only be one, and it applies universally to the entire use of freedom" (R 6:25).

According to Kant's considered view in the *Religion* evaluating an action requires attending to its principle, and to evaluate a particular principle requires evaluating the fundamental principle of choice of the utmost generality—"the subjective basis for the adoption of maxims"—that the individual has adopted in a free act. This fundamental principle of choice determines a person's moral character, their *Gesinnung*.

This suggests that in speaking of admitting an incentive [*Triebfeder*] into one's maxim, in the passage Allison quotes, Kant is speaking of admitting an incentive [*Triebfeder*] into one's most general maxim. It is because of this that Kant is able to draw the conclusion that the individual cannot be at some points good and at some points evil:

For if he is good in one point, then he has admitted the moral law into his maxim; thus if in another point he were to be simultaneously evil, then, because the moral law of compliance with duty as such is only one and is universal, the maxim referred to it would be universal, while simultaneously being only a particular maxim—which

is contradictory. (R 6:24)

Goodness, the goodness of a particular action, Kant claims, must be understood as grounded in the recognition of the necessity and *universality* of the moral law. To recognize the necessity and universality of the moral law is to admit the moral incentive [*Triebfeder*] into one's most general maxim of choice, and thereby to be motivated by the moral law. A recognition of the necessity and universality of the law is not consistent with the occasional admission of the sensible incentive [*Triebfeder*] as ground of motivation. Since one is good if and only if one has recognized the necessary and universal validity of the law, one cannot be good at one time and evil at another.

What is important for our purposes here is not to elucidate or defend this difficult claim (to which I will return later in the essay). I mean only to substantiate the interpretative claim that in the passage that Allison singles out Kant does not appear to have in mind the ability to step back from concrete desires, recognize them as desires for concrete ends, reflect upon those ends, and choose whether or not it would be reasonable to pursue them. I will return to elucidate the meaning of Kant's remark in detail later in the essay (Chapters Three and Four).

§7. Kant's Critique of "Additivism" in his Theoretical Philosophy

Before concluding, it might be worthwhile to briefly note that interpreters of Kant's account of theoretical cognition face a challenge that bears an analogous structure to the one I have been engaging thus far in Kant's practical philosophy. The debate is characterized as one between conceptualists and non-conceptualists, although the criteria of application for these

labels are themselves in dispute.²⁵ Broadly speaking parties to the debate disagree about whether or not Kant held that the exercise of sensibility, the capacity for intuition, requires, in any sense, the exercise of the understanding, the capacity for judgment. Conceptualists affirm the claim; non-conceptualists deny it. Abstracting from the details of the various accounts, the considerations against the non-conceptualist are similar to the anti-additivist concerns I raised above.²⁶

Conceptualists differ with regard to the question of how to conceive of the involvement of rational capacities in the exercise of sensibility. They agree, however, that Kant's account of theoretical cognition is grounded in an argument against so-called additive accounts of cognition; and that the account of the relation between sensibility and understanding in the case of theoretical cognition secures the exercise of rational capacities in experience. I find persuasive the claim that it does so by securing the exercise of rational capacities in the exercise of our receptive capacities of inner and outer sense as the source of the unity that characterizes the exercise of those capacities, namely categorical unity. This is not a claim that I can defend here at any length.

However, the involvement of rationality in the exercise of sensibility is conceived, since conceptualists are committed to the claim that Kant was sensitive to anti-additive considerations

²⁵ One confusion that plagues the application of the "conceptualist" labels is the failure to distinguish different possible senses of the term "concept," in particular to fail to distinguish the radical difference between empirical concepts and the pure concepts of the understanding, i.e., the categories, and so to fail to distinguish the question of whether and how empirical concepts are involved in the intuition of objects from the question of whether and how the pure concepts of the understanding, i.e., the categories, are involved in the intuition of objects.

²⁶ For a perspicuous overview of the debate see Thomas Land, "Kantian Conceptualism," in *Rethinking Epistemology*, eds. Günter Abel and James Conant (Berlin: De Gruyter, 2011), 197–239.

in his account of theoretical cognition, they should hesitate before saddling him with an account of that shape in his account of practical cognition, if it can be avoided.

§8. Our Options Going Forward

Two options present themselves moving forward:

1. *Dualism*: Admit both faculties of desire, the lower one being intelligible independently of the activity of the higher one, as being part of our own practical-affective life, and reject the notion that there is any interaction between them and therefore of a non-accidentally unified practical life.
2. *A Non-Additivist (Monist) Reading*: Reject the supposition that reference to purely receptive sensible desire and feeling is meant to indicate a capacity that is present *in us* as an independently intelligible element functioning in the way we desire. The power of rationality is not something that could be “tacked on” to a mind that already forms an intelligible system apart from this addition. Rational form pervades our entire life, including *all* our sentient powers.²⁷

On the face of it, both options outlined seem unattractive. The first, often derisively attributed to Kant, denies the possibility of conceiving of our life as a unified practical life: doing the good and finding satisfaction come together only accidentally. It would, additionally, render mysterious how the exercise of practical reason ever comes to be directed towards concrete objects in the world and how could we be held responsible for acting on inclination.

²⁷ There are, strictly speaking, two more options. Both allege that that there is only one faculty of desire: either a lower or higher one.

The second suggestion, whereby our desiring is itself understood as an actualization of an essentially rational power, seems to fly in the face of Kant's insistence on the distinction between the autonomous good will and heterogeneous inclinations.

§9. Kant on Reason-Determined Feeling

Yet despite the apparent foreignness to Kant of a non-additive account of the relation between practical reason, desire, and sensibility I would like to remind the reader of one moment where Kant seems to appeal to the idea of reason-determined feeling, namely, in his treatment of moral respect.

These are passages from texts spanning the entire critical philosophy in which Kant appears to admit the possibility, indeed necessity, of supposing that reason can determine sensibility not only insofar as it makes the world available to theoretical cognition, but also insofar as it provides the materials for practical cognition in feeling, i.e., where he claims that reason can determine our feelings, namely our feeling of moral respect.

In the *Groundwork* Kant says that the human will is subjectively determined by pure respect for the practical law, which is a feeling. But moral respect is not "received by means of influence" but a feeling "self-wrought" by means of a rational concept of reason (*G* 4:401n). One can feel not just in accordance with duty but from duty, so to speak.

Later in the *Groundwork* Kant claims,

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 *induce 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. (*G* 4:460)

Kant goes on immediately to claim that this “causality” is “a special kind of causality” that cannot be understood on the lines of empirical causality, which can only relate objects of experience. As we shall see below, in the third *Critique* he calls this the causality of freedom.

In the *Critique of Practical Reason* Kant speaks of moral respect as a feeling produced solely by reason and which serves as the incentive [*Triebfeder*] to make the law the will’s maxim. Respect “is a feeling which is directed only to the practical and which depends on the representation of a law *only as to its form* and not on account of any object of the law” (*KpV* 5:80). This isn’t an “antecedent feeling” (5:75), attuned mysteriously to morality in the way feeling may be attuned to hunger, but is *determined by reason*.

In the *Metaphysics of Morals* Kant claims:

Every determination of choice proceeds from the representation of the possible action, through the feeling of pleasure or displeasure in taking an interest in the action or in its effect, to the deed; and here the *sensitive condition* (the affection of the internal sense) is *either a pathological or a moral feeling*. The former is that feeling which precedes the representation of the law; the latter is that which can only follow the representation of the law. (*MS* 6:399)

In the *Critique of the Power of Judgment* Kant speaks freely of the good as that which pleases “by means of reason alone, through the mere concept” (*KU* 5:207).

And finally in §12 of the third *Critique* he writes:

[I]n the critique of practical reason *we actually derived the feeling of respect (as a special and peculiar modification of this feeling, which will not coincide exactly either with the pleasure or with the displeasure that we obtain from empirical objects) from universal moral concepts a priori*. [...] [T]here we could also step beyond the bounds of experience and *appeal to a causality that rests on a supersensible property of the subject, namely that of freedom*. But even there we did not actually derive this feeling from the idea of the moral as a cause, rather it was merely the determination of the will that was derived from the latter. The state of mind of a will determined by something, however, *is in itself already a feeling of pleasure and is identical with it, thus it does not follow from it as an effect*. (*KU* 5:221–2)

These passages leave undetermined, of course, whether Kant was committed to an additive account of practical cognition, namely, whether was he committed to the claim that purely pathological feeling is possible in us, and thus that on some level there is something that is *merely* sentient in the finite rational being, something that threatens to render mysterious how thought ‘interacts’ with it.

But that he admits that reason *can* determine feeling, indicates that the idea of receptivity determined through and through by practical reason was not in principle unthinkable by him. In what follows I will argue that a non-additive reading of Kant’s practical philosophy is both exegetically defensible and philosophically valuable.

§10. An Aside to Assuage Some Worries

I will conclude by trying to assuage a worry that may arise at this point, and simultaneously indicate a final way in which so-called additive interpretations miss the mark, namely by failing to secure the kind of claims that may motivate additive accounts in general but that certainly motivate additive interpretations of Kant.

The worry is that that the aims of the project at I have outlined them will render unintelligible a key observation which rightfully grounded Kant’s moral project: we are such as to be susceptible to an experience of desiderative and in particular, moral, conflict. So much so that we can experience being assailed by desires as strange, as *alien* to our “true selves.” We find ourselves with desires of which we say, and not without no justification: “that’s not me.” Any account of practical cognition must be able to say something about these topics, certainly an account that purports to offer an interpretation of Kant’s moral philosophy. After establishing an account of moral respect as the human faculty of feeling, I will in Chapter Three offer an account

of the constitution of moral character. In the fourth and final chapter I will offer a new reading of how to understand evil in Kant and it is there that I will be able to address the question of motivational conflict in detail.

However, for now I will limit myself to making the following suggestion: the intelligibility of the experience of the strangeness of a feeling or desire presupposes at least the possibility of feeling and desire from which I am not alienated. The exclamation: “I’m sorry, that’s not me, I don’t know what took over me” presupposes that one can, as a subject of desires of feelings, take these to be one’s own. Not only in the sense that they are merely in line with her considered commitments, but in the sense that she takes them to be grounded in her commitments. For example, in the primary case, a practical subject does by luck happen to be angry when wronged, nor is her anger efficiently caused by her being wronged; in the primary case, her being wronged is the reason and only thereby the cause of her anger. It is in this sense that a subject can be said to be *rightfully* angry because she is wronged. Additive theories can account neither for the possibility of feeling and desire being a subject’s own, nor for the priority of this kind of feeling and desire. As such they cannot account for the phenomenon of conflict either.

Chapter 2

Moral Respect

§1. Introduction

Kant grounded his moral system in the thought that to do what is right is to do what reason prescribes, whether or not it serves your personal interests, whether or not you happen to feel like it. According to the dominant reception of Kant, advanced by his sympathizers and critics alike, this principle and the system to which it gives rise are grounded in an irreducible dualism between, on the one hand, our rationality and freedom and, on the other, our animality and feeling. In this picture, we are essentially rational, free beings capable of self-determination, *and* subject to the influence of our feelings—forces produced by nature and habituation, whose presence is beyond our rational control.

According to this traditional interpretation, Kant does violence to the most fundamental aspects of human experience. First, in our common ways of self-understanding we do not encounter an in-principle unbridgeable gap between our emotional states and our rational mental activity. Of course we find ourselves often questioning whether a particular emotion is reasonable or justified, but it is precisely this act of questioning which presupposes the possibility of emotional experience that is reflective of our rational commitments. For example, we consider it *reasonable* and indeed *appropriate* to feel indignation in response to moral injustice or to feel regret at the realization of wrongdoing.

Second, in this interpretation—one that sees emotions as divorced from our essence as rational and free—what we most deeply *care* about is rendered irrelevant from the standpoint of

the practical agent, which is characterized as the standpoint of universal legislation, i.e., the impartial standpoint of the moral law. Thus it seems that Kant undermines the legitimacy of our passionate concern with the objects of our affective lives—what we yearn to have and dread losing.

Third, a dualism of rationality and feeling threatens the internal coherence of the Kantian account itself. Specifically, it would render it difficult for Kant to provide a convincing account of how we are *motivated* by the moral demands he labors to articulate. To be capable of motivation by a moral demand is to be capable of doing something just because one has recognized that doing so is the right thing to do. In other words, it is to *be moved* to do the right thing in light of a recognition of an action's goodness. This requires our capacity to *feel* to respond to the claims of reason. Kant insists on the point himself and claims that a feeling (namely, moral respect) is necessary for the performance of moral action. A picture where our emotional lives are divorced from our lives as free and rational cannot support the idea that feeling could come to reliably manifest the demands of reason, and since it renders unintelligible the idea that a concern with the right thing to do could ever come to manifest itself affectively, it can hardly support the idea that that concern could move us to act.

Scholars sympathetic to Kant's account of morality, which promises to secure objective universal moral claims, have tried to resist these implications by rehabilitating Kant's treatment of human feeling. In particular, commentators have attended to his account of the feeling of moral respect—which Kant defines as the necessary effect of reason on sensibility in the determination to moral action—as a key to understanding the relation of reason and feeling in Kant generally and securing the role of feeling in moral motivation specifically. Nevertheless, a

coherent reading of the role of feeling in moral motivation—and with it Kant’s profound insight into the distinctive role feeling plays in the life of a rational animal—has so far eluded commentators.

I argue that Kant, often taken to be a classical source of the opposition of freedom and feeling, is in fact one of the most insightful guides toward a coherent non-dualistic view of this relationship. Specifically, I argue that implicit in Kant’s account of the feeling of moral respect is the idea that human feeling is a unique mode of self-consciousness—self-revelation and self-understanding—which discloses the subject to herself as rational, embodied and capable of freely determining herself to act in the world. Human emotion, the distinctive human capacity for feeling, emerges as the form of self-consciousness *constitutive of practical agency*, i.e., of freedom. This interpretation not only solves the stubborn interpretive puzzle of moral respect in Kant (thereby showing the Kantian account to be coherent on its own terms), but reveals Kant’s account to be profoundly relevant to a philosophical understanding of human feeling.

§2. Argument Overview

I begin by introducing the interpretative problem. Kant insists that feeling must play a role in moral motivation. However, the involvement of feeling in determination to action threatens to jeopardize autonomy, and thereby the moral worth of actions so determined (§3). I next turn to the interpretative debate. I identify the common interpretative assumption shared by parties to the debate, most importantly the supposition that when Kant speaks of feeling, moral or otherwise, he is speaking of various effects on a single “faculty”: a sensible, receptive

capacity for feeling whose exercises we recognize as pleasure and pain. In understanding the relationship between this capacity for feeling and moral action, commentators are divided into two camps. Intellectualists deny any role to feeling in the determination to moral action (against the letter of the Kantian account, and as I've suggested, against our intuitive understanding of the relation between our feelings and our rational activity). Affectivists, in an attempt to secure for moral respect a unique status that would legitimize its role in moral motivation, appeal to Kant's distinction between "pathological" and "non-pathological" feeling and claim that only pathological feeling poses a threat to freedom, while non-pathological feeling, the kind of feeling that is moral respect, can happily play "a motivating role" in moral action (§4). However, the characterization of moral respect as "non-pathological" must be adequate to fund the critique that Kant himself offers of sentimentalist accounts of the role of feeling in moral motivation. Despite their efforts, affectivists fail to offer a characterization of "non-pathological" feeling that's adequate to account for Kant's critique against the involvement of pathological feeling in moral motivation. Their account leaves the sensible aspects of the feeling of moral respect as a merely accidental feature of our non-rational nature, and therefore, by Kant's lights, unsuitable to play the requisite motivating role in a satisfying account of moral action (§5). Moreover, the sort of account of non-pathological feeling that I here criticize is one that seeks to account for the rationality of respect by divorcing its sentient and sapient aspects from one another. The sharp distinction that these commentators deploy between the intellectual and affective moments of respect is not only textually unsupported but issues in a distinctively unattractive philosophical picture (§6).

Against this, my reading draws attention to Kant's oft-overlooked *general* characterization of feeling [*Gefühl*], which, according to Kant, is the awareness of how something affects a power of the mind, an awareness of whether it promotes its operation or hinders it (§7). In particular, I draw attention to the fact that Kant's notion of feeling is defined in strict relation to the notion of a "faculty of desire," the ability of a being to act for ends. According to this account, pain marks the awareness that something hinders me from acting, pleasure the fact that something benefits and promotes my efforts to act. However, as is well recognized, two different kinds of faculties of desire are in play in Kant's account: a lower, sensible faculty of desire of the kind that belongs to subrational animals, and a higher, rational faculty of desire. Our *human* faculty of desire is of course the latter: a will, a faculty of *rational* desire. Previous commentators have failed to appreciate the bearing of this distinction between sensible animal desire and *human* rational desire on the entire shape of Kant's moral psychology. They are thereby forced to ultimately classify all feeling, including moral respect, as the awareness of promotion or hindrance of the *non-rational, sensible faculty of desire*. By disputing and correcting this assumption, I conclude that we cannot do justice to Kant's own conception of moral respect—to the feeling that he describes as one that is "self-wrought by reason" (*G* 4:400n)—if we simply conceive of it as an awareness of effects on a *sensible* animal faculty of desire. Or, in other words, in order to do justice to this feeling it is not enough to just classify it as one feeling among others, albeit a very special one; rather, we must come to terms with the way in which this feeling differs *in kind* from non-rational feeling: it marks effects not on a sensible faculty of desire but on a rational one. The human capacity to feel is the mode of awareness of how things affect our *rational desire*, that is, our "will." Therefore, human feeling

is the awareness of how things promote or hinder my capacity to act, *insofar as* I am rational, embodied, and free, capable of acting for ends I set for myself and in accordance with principle. Thus, I conclude, moral respect names for Kant the fundamental structure of the distinctly *human* capacity for feeling (§8).

In this interpretative framework Kant's account of moral respect yields further insight into the essential nature of human feeling. Kant claims that the object of the feeling of moral respect is the moral law. This does not, however, mean that the object of this feeling has some sort of fully determinate content on its own; for the moral law, on Kant's account, is nothing but the *form* of our will. This means that the moral law is the *fundamental principle* of our rational, free activities. If the feeling of moral respect has for its object not a particular end or object but this principle, then it is the mode of awareness or the form of one's activity. It is, in other words, the mode of *self*-consciousness that characterizes moral being: moral respect, the human faculty of feeling, is the way human beings are conscious of *themselves* as free and efficacious agents. Moreover, the moral law is not only the principle of *my* individual activity as free and rational, but also the principle of the activity of every other person—in fact it is all-pervasive, it is the fundamental principle of everything touched, so to speak, by practical rationality. It follows that human feeling is not only the mode of awareness of ourselves as practical agents, but equally the mode of awareness of other persons, as well as to our shared forms of activity (§9). Thereby, moral respect, as the mode of self-consciousness which grounds our awareness of ourselves and others as moral beings, is constitutive of practical agency as such.

§3. The Interpretative Problem

Kant's account of the feeling of moral respect poses an apparently insoluble exegetical and philosophical puzzle.¹ On the one hand, moral action is supposed to be autonomous—in acting from the moral law I set my own ends in accordance to principle—and as such must be independent of any external determination. Feelings, according to the conventional understanding of Kant, are reflective of our contingent, non-rational animal nature and empirical habituation. We can only come to know them empirically, i.e., observe them in experience. Therefore, in particular, in acting from the moral law feelings cannot determine or influence the will to perform an action that has moral worth.² On the other hand, Kant claims that the “feeling of moral respect” is a necessary moment of acting from the moral law: “[i]mmediate determination of the will by means of the law and consciousness of this is called respect” (*G* 4:400n).³ This seems to threaten Kant's characterization of morally worthy action as autonomous. If, in order to act from the moral law one's consciousness of the moral law as a law binding on oneself is not sufficient, but *in addition* a particular feeling is necessary for the action to be performed, then one's acting morally would be explained not *only* by one's consciousness of this very law but *also* by the presence of a feeling, over which the subject cannot exercise any rational control. We face the following impasse: morally good action is at once autonomous and heteronomous.

¹ So much so that Robert Paul Wolff famously charged that “the introduction of the emotion of reverence [respect] is contradictory to the entire thrust of Kant's argument.” See Wolff, *The Autonomy of Reason* (New York: Harper & Row, 1973), 83.

² Onora O'Neill provides a clear statement of the worry about attributing any role, motivating or otherwise, to the moral feeling of respect or reverence: “To act ‘out of reverence [respect] for the law,’ is not to act with any peculiar *feeling* of reverence or awe. [...] Pathology, as Kant would have it—psychology, as we would say—is irrelevant to the moral worth of acts.” See O'Neill, *Acting on Principle* (New York: Columbia University Press, 1975), 111.

³ See also *G* 4:401, 401n, 459, 460; *KpV* 5:73, 79, 80, 116–119; *MS* 6:211, 211n, 399, 436, 449, 464, *KU* 20:207, 5:257, 5:266, *R* 6:23, 27–28).

§4. The Interpretative Debate

Parties to the interpretative debate customarily share the following assumptions:

1. “Respect” is necessary for the performance of morally worthy, i.e., free action.
2. There is a distinction between two aspects of respect: a “purely intellectual recognition of the supreme authority of the moral law” (defined negatively by having nothing to do with feeling) and “a peculiar moral feeling of respect for law.”⁴
3. When Kant speaks of feeling, moral or otherwise, he is speaking of various effects on a single “faculty”: a sensible, receptive capacity for feeling whose exercises we recognize as pleasure and pain.

Two camps emerge in the interpretative debate: affectivists and intellectualists. The distinction between them, following Richard McCarty, turns on the question “whether the affective component of respect plays any role in the mechanism of moral motivation.”⁵ That is, does *feeling* play a role in the determination of morally worthy action?

Intellectualists deny any role to the *feeling* aspect in the determination to moral action.⁶

Andrews Reath, for example, writes, “It is the practical aspect [of respect] that is active in

⁴ Richard McCarty, “Kantian Moral Motivation and the Feeling of Respect,” *Journal of the History of Philosophy* 31, no. 3 (1993), 421.

⁵ *Ibid.*, 430.

⁶ The paradigmatic proponent of the intellectualist position is Andrews Reath. See Reath, “Kant’s Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination,” *Kant-Studien* 80, no. 1–4 (1989), 284–302. Jens Timmermann holds that the cultivation of moral feelings is an indirect duty concerning the acquisition of instrumental means for implementing moral ends. Moral feelings are thus not valuable in themselves, but they make it easier to act morally. Timmermann, “Kant on Conscience, ‘Indirect’ Duty, and Moral Error,” *International Philosophical Quarterly* 46, no. 3 (2006), 298–302. This view is also defended by Henry Allison in *Idealism and Freedom: Essay on Kant’s Theoretical and Practical Philosophy* (Cambridge: Cambridge University Press, 1996), 123, and Pablo Muchnik in “The Heart as Locus of Moral Struggle in Religion,” in *Kant on Emotion and Value*, ed. Alix Cohen (London: Palgrave MacMillan, 2014), 233–4. Marcia Baron claims emotions cannot have a motivational function because of impurity concerns and argues that feelings only have a supportive epistemic function: they “help to direct ... our attention to the needs of particular others and to ways we might help.” Baron,

motivating moral conduct, while the affective side, or *feeling* of respect, is its effect on certain sensible tendencies.”⁷ In this picture, a “practical [intellectual] aspect of respect” is wholly responsible for acting as one should, and no feeling is required to explain why the agent acted as she did when she acted from respect for the moral law. Feelings merely accompany this moral activity as a side effect.

The challenge to the intellectualists is exegetical. Whereas the affectivist can easily point out to many passages that suggest that the “feeling aspect” of respect is necessary to motivate moral conduct, intellectualists have tried to dismiss these passages (for example Kant’s pronouncements reported by his students and published in the *Lectures on Ethics* (*MPC*, *AA* 27:1428), or even the *Groundwork* (*G* 4:400–1), as expressions of an early, abandoned view.⁸ Yet this intellectualist suggestion becomes impossible to sustain in the face of passages from Kant’s later writings (e.g., *KpV* 5:75, 5:79). In dismissing these passages, we will see, the intellectualists do not only ignore the letter of Kant’s text, but miss Kant’s deep insight into the distinctive role feeling plays in the lives of free, rational beings.

Affectivists, in an attempt to secure for moral respect a unique status that would legitimize its role in moral motivation, appeal to Kant’s distinction between “pathological” and

Kantian Ethics Almost without Apology (Ithaca, NY: Cornell University Press, 1995), 220. Nancy Sherman claims moral feelings are not necessary for moral action but are a “layer of character that can ... best support moral motivation,” they “positively promote ... our duty motive.” In addition to this subsidiary motivational role, there is also an epistemic function of moral emotions: they are “modes of attention that help us to track what is morally salient ... in our circumstances.” Sherman, “Kantian Virtue,” in *Making a Necessity of Virtue: Aristotle and Kant on Virtue* (Cambridge: Cambridge University Press, 1997), 144–146. A similar view is held by Anne Margaret Baxley in *Kant’s Theory of Virtue: The Value of Autocracy* (Cambridge: Cambridge University Press, 2010: 124, 135, 136, 145, 164). See also: Onora O’Neill, *Acting on Principle*; Paul Guyer, *Kant and the Experience of Freedom: Essays on Aesthetics and Morality* (Cambridge: Cambridge University Press, 1993); Stephen Engstrom, “The Triebfeder of Pure Practical Reason,” in *Kant’s Critique of Practical Reason*, eds. Andrews Reath and Jens Timmerman (Cambridge: Cambridge University Press, 2010).

⁷ Reath, “Kant’s Theory of Moral Sensibility,” 287.

⁸ See O’Neill, *Acting on Principle*, and Wolff, *The Autonomy of Reason*, 83.

“non-pathological” feeling and claim that only pathological feeling poses a threat to freedom, while non-pathological feeling, the kind of feeling that is operative in “moral respect,” can happily play “a motivating role” in moral action.

That is, affectivists claim that the affect produced when the moral law determines the will, i.e., the feeling aspect of respect, is responsible for the *motivation* to moral action.⁹ The affectivist, in McCarty’s characterization, “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 the law, one consequent to the initial recognition or moral judgment the intellectualists emphasize exclusively.”¹⁰ The affectivist account of moral respect thus identifies two necessary steps for the performance of morally worthy action: intellectual recognition followed by moral feeling. It is the latter that explains why the agent, having “recognized” the moral law, was moved to act.

How do the affectivists respond to the intellectualists’ worry of heteronomy? McCarty, on behalf of affectivists, identifies an implicit assumption that the intellectualists erroneously hold and which gives rise to this undue worry: the “classification of all feelings as

⁹ A paradigmatic proponent is Richard McCarty (see “Kantian Moral Motivation”). In later work McCarty conceives of moral emotion as a “psychologically forceful incentive” that aids agents to do what duty requires by outweighing other motives. Moral feelings are pleasures and displeasures, which allows us to say that “the maxim incorporating the motivationally stronger incentive” prevails in cases of conflict. McCarty, *Kant’s Theory of Action* (Oxford: Oxford University Press, 2009), 167, 182. Christine Korsgaard claims that our capacity for respect names the rational will’s capacity to provide “not only the ground of choice but *also* the incentive to act in accordance with that ground” and goes on to demonstrate how the moral law could provide painful and pleasurable feelings that could serve as this incentive. Korsgaard, “From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action,” in *The Constitution of Agency* (Oxford: Oxford University Press, 2008), 187 (emphasis mine). See also: Owen Ware, “Kant on Moral Sensibility and Moral Motivation,” in *Journal of the History of Philosophy* 52, no. 4 (2014), 727–46; Larry Herrera, “Kant on the Moral Triebfeder,” *Kant-Studien* 91, no. 4 (2000), 395–410; Ido Geiger “Rational Feelings and Moral Agency,” *Kantian Review* 16, no. 2 (2011), 283–308; Jeanine Grenberg, “Making Sense of the Relationship of Reason and Sensibility in Kant’s Ethics,” *Kantian Review* 16, no. 3 (2011), 461–72.

¹⁰ McCarty, “Kantian Moral Motivation,” 423.

pathological.”¹¹ The affectivists charge the intellectualists with mistakenly supposing that because respect for the moral law must be capable of motivating action independently of the typical, empirical, contingent, or as Kant calls them, “pathological” motivational resources of human agency, commentators are forced to claim that “respect for the moral law motivates independently of any feeling or affections whatsoever.”¹² Embracing the letter of the text, namely Kant’s insistence that respect in its role in motivation is a feeling, seems therefore as easy as denying that all feeling is pathological and acknowledging that, specifically, respect is not. Indeed, the claim that not all feeling is pathological and the further claim that thereby the mystery of moral respect can be put to rest is a conclusion reached with increasing frequency by readers of Kant, who take it as trivially obvious.¹³

At this point, the affectivist position faces two challenges. (1) She must provide a characterization of non-pathological feeling that withstands Kant’s critique of the involvement of feeling in moral motivation; (2) The affectivist faces both philosophical and textual challenges to her reliance on a *distinction* between intellectual recognition and an affective response, which is required for her interpretation. Identifying and addressing these two challenges will lead us towards an alternative interpretation.

§5. First Challenge: Characterizing Non-Pathological Feeling

¹¹ Ibid., 424.

¹² Ibid.

¹³ Ido Geiger claims that “the feelings Kant claims are necessary conditions on moral agency are not what he calls inclinations. They belong to distinct category of affects, namely, rational moral feelings.” Geiger, “Rational Feelings and Moral Agency,” 284. Owen Ware claims that respect must be non-pathological which he glosses as “not based on sensibility” but provides no further elucidation. Ware, “Kant on Moral Sensibility,” 731. Allen Wood characterizes moral feeling as the pleasure which rational desire “produces in our sensibility.” Wood, *Kantian Ethics* (Cambridge: Cambridge University Press, 2008), 35.

Affectivists try to overcome a primary deficiency of the intellectualists' reading by pointing to a distinction between pathological and non-pathological feeling and claiming that respect is a non-pathological feeling. It is only pathological feeling, the kind characteristic of our empirical, contingent nature, that is a threat to autonomy. Since respect is non-pathological, it can motivate action without jeopardizing freedom.

Affectivists must, however, characterize non-pathological feeling in such a way that it does not collapse into empirical and contingent feeling. Recognizing this necessity, affectivists point to moral respect's allegedly intellectual and non-natural (i.e., not empirically cognizable) cause. Larry Herrera, for example, writes, "we all know that for Kant there is a fundamental difference between respect and sensuous feelings. The former is effected by pure reason alone; the latter, sensuously so. Thus, although all feeling is sensuous, not all feeling is pathological."¹⁴ In this account, respect is not pathological *by virtue of the source of the cause that effects it*. Specifically, in non-pathological feeling the representation which affects our faculty of feeling and induces the feeling of moral respect is a representation of the moral law, it is a not representation of an object of the senses, but *a representation of reason*.

However, this interpretative suggestion, broadly adopted and unchallenged, faces a major exegetical challenge that reveals in turn a deep philosophical problem. Kant vehemently objects to the idea that the source of the representation can be used to distinguish between feelings as far as the moral worth of the actions they motivate is concerned. Early and prominently in the second *Critique* (Part I, §3) Kant explicitly denies that the origin of a representation in reason could secure for it a legitimate motivating role in the determination to moral action.

¹⁴ Larry Herrera, "Kant on the Moral Triebfeder," 401.

However dissimilar representations of objects may be—they may be representations of the understanding *or even of reason*, in contrast to representations of sense—the *feeling of pleasure by which alone they properly constitute the determining ground of the will* (the agreeableness, the gratification expected from the object, which impels activity to produce it) *is nevertheless of one and the same kind not only insofar as it can always be cognized only empirically but also insofar as it affects one and the same vital force that is manifested in the faculty of desire*, and in this respect can differ only in degree from any other determining ground. (*KpV* 5:22–3, emphasis mine).

Here Kant explicitly claims that *no feeling* is to serve as the determining ground of the will, even if it is caused by a representation of reason. Therefore, the affectivists' attempt to distinguish moral feeling from other types of feeling by virtue of the affecting representation's origin in reason is insufficient to secure its legitimate role in motivating moral action. Even if doing the right things would in fact happen to feel good, that promise of pleasure cannot be what motivates the agent to act.

Since in the affectivists' picture the faculty of feeling is conceived as an ordinary, empirical, and sensible faculty of feeling, the only way in which the thought of the moral law could induce a feeling would be for our faculty of feeling to be so constituted as to *of itself* respond to a thought of a certain kind, in this case thought “of the moral law,” with pleasure and thereby a positive incitement to action. Reason, in this picture, is “responsible” for our recognition of the moral law, and particular moral demands, itself, but it is a feature of our contingent, psychological constitution that this recognition produces a certain feeling. In order to do the right thing we would have to have just the right kind of non-rational sensible faculty, one that would be able to distinguish and privilege, of its own, the source of a representation in reason and thereby respond to inducements by commandments of the moral law. It is not clear whether the idea of a wholly non-rational sensible faculty whose exercise bears no *essential*

connection to the exercise of rational capacities that is reliably disposed to respond to pronouncement of reason with pleasure is intelligible.¹⁵

But even supposing that a sensible faculty could *of itself* respond to representations of reason, in the affectivists' picture the claim that reason is able to motivate us to moral action amounts to the supposition of a certain kind of empirical sensibility. Rejecting this possibility is at the heart of Kant's moral system and his critique of sentimentalism:

There is here no *antecedent* feeling in the subject that would be attuned to morality: that is impossible, [...] the incentive [*Triebfeder*] of the moral disposition must be free from any sensible condition. (*KpV* 5:75)

The affectivist does precisely what Kant warns against: she assumes an antecedent capacity for feeling in the subject that is "attuned" (i.e., attuned of its own) to the claims of the moral law. In addition to recognizing the law, i.e., recognizing what it would be good to do, in order to *act* on this understanding, the independent cooperation of a sensible non-rational faculty is required. In the case of a human agent doing the right thing, the full explanation of why she acted as she did would have to be: I recognized what the right thing to do was, and, luckily, felt like doing it.

So, while affectivists promisingly direct our attention to Kant's characterization of respect as a non-pathological feeling as the key to resolving the puzzle of feeling's role in moral action, they miss the deeper import of the intellectualists' worry. Whatever the feeling of moral respect is, its role in the determination to moral action cannot be a matter of inducing in a non-rational faculty, by whatever means, a feeling that motivates action. Without recovering an account of non-pathological feeling that is not grounded in a distinction *between the sources of representations that affect the feeling faculty*, the affectivist cannot appeal to the

¹⁵ An account of this structure would be "additive" in the sense I have introduced in the previous chapters, following Matthew Boyle in "Additive Theories of Rationality."

pathological/non-pathological distinction to secure the feeling of moral respect a role in the determination to moral action.

A clue toward an alternative ground for the distinction between pathological and non-pathological feeling will emerge from a consideration of a second challenge to the affectivist, namely, to her reliance on a two-stage account grounded in a distinction between the “intellectual recognition” of the moral law and the “feeling” of moral respect responsible for being motivated by the moral law. A two-stage account, I will go on to argue, misses Kant’s radical claim about the unity of the recognition of the law and the feeling of moral respect: on Kant’s account, the feeling of moral respect is *nothing but* the mode of recognition of the moral law.

§6. Second Challenge: The Distinction Between Recognition and Feeling

As we’ve seen, the affectivist interpretation shares with the intellectualist the distinction between an affective, feeling mode of response to the moral law and another primary mode of recognition, which they call intellectual or practical. The affectivist differs from the intellectualist in claiming that recognition of the law in the intellectual or practical sense is followed by a feeling that is responsible for the *motivation* to action. For example, McCarty claims that “the moral law determines the will directly, and *then follows* the feeling of respect, which subsequently determines the choice to act accordingly” (emphasis mine).¹⁶

Kant’s own characterization of the feeling of moral respect is however subtly and importantly different:

¹⁶ McCarty, “Kantian Moral Motivation,” 428.

What I cognize immediately as a law for me I cognize with respect, which signifies merely consciousness of the subordination of my will to a law without the mediation of other influences on my sense. Immediate determination of the will by means of the law and consciousness of this is called respect. (*G* 4:401n)

To recognize the law by the subject of the law as a law for the subject is just “to cognize the law *with respect*.” The act of recognition of the law *as a law* for the subject, i.e., the act of recognizing that one is bound by the law, *is itself* the feeling of respect. Thus, respect is the way we recognize the law as a law for us, and this recognition leaves no remainder: nothing further is necessary for the subject to act.

With Kant’s characterization of moral respect, Kant is denying the claim that recognition of the law, i.e., rational recognition of demand, and feeling are fundamentally distinct. If you are not moved to act in the manner you’ve recognized you ought to, nothing is left of the idea of recognition of a demand. A *practical* demand is something you *must* respond to not with “intellectual” assent but *with action*. To recognize a law as issuing a demand is to find that demand necessary and binding, i.e., to act in accordance with it and because of it. (In the same way that to recognize a proposition as true is to come to believe it).

If feeling does not play a secondary motivating role in addition to intellectual recognition, in insisting that feeling is necessary for the recognition of the law, Kant can only mean that the feeling of respect is itself the form of recognition of the moral demand: the feeling of moral respect is nothing but the form of recognition of ourselves as bound by moral considerations.

The claim that determination of the will, the recognition of the law as a law for the subject, is itself the *feeling* of respect is reiterated by Kant in the *Critique of Judgment*. The following (sorely overlooked) passage is particularly illuminating. Here, Kant reviews his own treatment of moral respect in the second *Critique* and straightforwardly denies that the feeling of

respect is an *effect* of the determination of the will, and explicitly insists that it is instead *identical* with it:

[I]n the critique of practical reason we actually derived the feeling of respect [...] from universal moral concepts a priori. [...] there *we could also step beyond the bounds of experience and appeal to a causality that rests on a supersensible property of the subject, namely that of freedom*. [...] The state of mind of a will determined by something, however, *is in itself already a feeling of pleasure and is identical with it, thus it does not follow from it as an effect*. (KU 5: 221–2, emphasis mine)

Thus, for Kant, the determination of the will by the law, the recognition of the law as a law for the subject, does not effect respect at all, but *is identical* to the feeling of respect for the law.¹⁷

In order to give content to the idea of the identity of feeling and determination of the will, we must turn to Kant's *general* characterization of feeling. We will then be in a position to elucidate the precise sense in which respect is a “non-pathological” feeling.

§7. Kant's General Account of Feeling

What could Kant's assertion of *identity* between feeling and the state of mind in willing—an exercise of the capacity to act for the sake of ends and in accordance with principles—mean? To answer this, we need to examine and reconstruct Kant's *general* account of feeling.¹⁸

¹⁷ See, G 4:401n, 459, KpV 5:79, 88, 116–119, MS 6:211, 399, KU 5:222. Paul Guyer, who holds a similar two-stage reading, curiously claims that there is an “absence of any explicit characterization of [moral respect] in the *Groundwork*” and speaks of the “the introduction of the feeling of respect in the *CPrR*.” Guyer, *Kant and the Experience of Freedom*, 359.

¹⁸ Kant's general account of feeling receives scant attention in the context of interpretive debates about moral respect. This may partly be due to the fact that the account is scattered in remote corners of Kant's writings and is, as we shall see in a moment, somewhat technical and opaque. It may also be partly due to the fact that this general account of feeling does not easily cohere with any of the available interpretations

I reconstruct an account of feeling from Kant's sporadic and brief discussions of feeling in the first, second and third Critiques as well as the *Metaphysics of Morals*. An analysis of Kant's account of feeling will reveal that in speaking of feeling Kant does not have in mind *a single faculty of feeling at all*. Feeling is not another power of the mind: it is not one more ability we have in addition to others, such as the ability to perceive objects around us, to form empirical judgments, or to act according to ends we set for ourselves. It is instead an awareness of how things stand with the moral subject vis-à-vis her own activities and features of her environment.

For Kant, feeling is a kind of sensibility.¹⁹ Sensibility divides into two aspects of susceptibility to representation (1) sensibility as sense and (2) sensibility as feeling:

of moral respect. In any case, a proper understanding of Kant's conception of feeling is essential to a proper understanding of moral respect and the curious claim that feeling is identical to the state of mind characteristic of willing. Owen Ware is an exception but his treatment is brief: he notes the self-reflexive character of feeling but stops short of investigating the relationship of feeling and willing and particular, Ware, "Kant on Moral Sensibility," 733–734.

¹⁹ To say that feeling is a kind of sensibility is to say that it is a "receptive" capacity, not a purely spontaneous one. *Exercise* of a receptive faculty depends on conditions external to it. Significantly, the fact that a capacity is receptive does not preclude the involvement of spontaneous, i.e., rational capacities in its exercise; the exercise for our sensible capacities in empirical experience, for example, requires the exercise of functions of the understanding. In order to make sense of how we come to form judgments based on our perceptual experiences, we must understand perceptual experience as itself unified in particular ways. The source of these forms of unity, e.g., cause and effect, part and whole, is the understanding. Thus, although perception in a rational being is receptive—it depends on conditions outside the perceiving subject—perceiving is no less so an exercise a rational capacity. To be receptive, to see objects, to hear sounds, is, too, a kind of rational activity. *This is all to say that the very fact that feeling is receptive, i.e., that its exercise can depend on the existence of things outside the experiencing subject, does not thereby prevent it from being a kind of rational activity.* Kant claims in the B-edition of the transcendental deduction that the synthetic unity belonging to the a priori sensible representations of space and time, with which every particular spatiotemporal representation must agree, "can be none other than [the unity] of the combination of the manifold of a given *intuition in general* in an original consciousness, in agreement with the categories, only applied to our *sensible intuition*" (B161). So the a priori representations of space and time that determine the forms of theoretical sensibility have unity only insofar as a sensibly given manifold, in this case, the manifold of pure space and time, is combined in accordance with pure concepts *of the understanding*. In a famous footnote to the B161 passage, Kant concedes that it may not always have been clear that the form of sensibility form is in part determined by the activity of the understanding: "in the Aesthetic I ascribed this unity merely to sensibility, only in order to note that it precedes all concepts, though to be sure it presupposes a synthesis, which does not belong to the senses but through which all concepts of space and time first become possible" (B161n).

1. *Sensibility as Sense*: Sensibility as sense is the susceptibility to “objective” representation of the senses (*KU* 5:206). The representations of sense can (but do not have to) be referred by the understanding to an object for the cognition of an object: the subject can form judgments on the basis of the sensible representations she has. When she does this, the subject thereby *thinks* something by means of these representations.

2. *Sensibility as Feeling*: Feeling, Kant claims, is the “subjective aspect” of our enjoying any representations in general (*MS* 6:211). Feelings, Kant claims, “*cannot be explained by themselves at all*” (*KU* 20:231–2, emphasis mine). The “representations” of sensibility as feeling consist only in the awareness of the *relation* of representations (of whatever sort) to the subject. Therefore, “they can be only inadequately explained through the influence that a representation has on the activity of the powers of the mind” (*ibid.*). This influence is “the effect of a representation (that may be either sensible or intellectual) upon a subject,” (*MS* 6:211). This influence upon the subject Kant defines as the “causality of a representation for maintaining (pleasure) and hindering (displeasure) a state of the subject” (*KU* 5:220). When a representation is causally efficacious with respect to the power of the mind that is the subject’s faculty of desire, her capacity to act for the sake of ends, feeling reveals “the causality of a representation for producing its object” (*KU* 20:230–2).

We are now in a position to identify the basic structure of feeling: it is a relation of the subject to a representation, and through it to the object of the representation. Taking pleasure is an awareness of a representation as beneficial for a certain power of the subject and so—since the subject is a unity of various mental powers—it is an awareness of the subject as benefiting, *qua* subject of that power. Finding something to be painful is an awareness of a representation as

harmful to a certain power of the subject, and so an awareness of the subject, qua subject of that power, as harmed.

Two implications should be pointed out here. First, it follows from this characterization that in Kant's account of feeling there is no *distinct faculty* of feeling at all. There is no standalone ability to feel, whose own laws determine what it is the subject will find pleasurable and what painful. Instead feeling reflects and reveals the form and purpose of the various conscious activities in which we may engage, theoretical or practical.

Second, since what benefits or harms a particular power of mind depends on the constitution of each power, feeling does not have its own principles but, in revealing the different powers of the mind as benefited or harmed, reflects and reveals the constituting principles of the different powers of mind themselves—the understanding, imagination, the will. *Because an awareness of its different powers*, feeling is therefore *essentially an awareness of self*.

§8. Feeling and the Will

In the second *Critique* Kant applies his general understanding of feeling to the relation between feeling and the essential power of mind that is the faculty of desire—the power to act for the sake of ends:

The faculty of desire is a being's *faculty to be by means of its representations the cause of the reality of the objects of these representations*. Pleasure is the *representation of the agreement of an object or of an action with the subjective conditions of life*, i.e., with the faculty of the *causality of a representation with respect to the reality of its object* (or with respect to the determination of the powers of the subject to action in order to produce the object). (*KpV* 5:8n)

Feeling, in its relation to desire, our capacity to act for the sake of ends, is the representation of agreement or disagreement with, i.e., promotion or hindrance of a subject's exercise of a faculty of desire. However, the faculty of desire in rational agents in general and human beings in particular is not merely an animal faculty of desire, but "a specially constituted faculty of desire" (G 4:428), "distinct from a mere faculty of desire" by being a "faculty of determining itself to action as an intelligence and hence in accordance with laws of reason independently of natural instincts" (G 4:459). By saying that the will is not a "mere" power of desire, Kant is of course not denying that the will is a power of *desire* but claims that our power of desire is a special species of the genus "faculty of desire." The *human*, rational faculty of desire is a faculty of desire, the power to act for the sake of ends, but it is not merely that: it is the capacity to act for the sake of ends *one sets for oneself and in accordance with principle*, therefore in awareness of one's freedom from external determination.²⁰

In a being in whom reason is practical, a being with a will and not a power of "mere desire," feeling will reveal how objects and actions benefit or harm the subject's power of willing, i.e., her capacity to act freely.

²⁰ This idea requires a little clarification. Rational desire is the capacity to act for ends that one sets for oneself, in light of an understanding of what it is good to pursue. Other subrational animals are able of course to act for ends: for example, a lion can run *in order to* catch his prey, but these ends are not "set by" the lion. The ends belong to the lion essentially not as an autonomous individual, but as a member of the kind of living species that the lion belongs to. Lions hunt. This is what being a lion is, and this is what explains why the lion is hunting when it is. Rational beings act for ends that they set for themselves according to principle—I may run to catch prey, but it does not essentially belong to my "species," the human, to hunt. In the ordinary case, if I'm running after prey, I've chosen to do so, under some understanding that this, hunting, is what I should do for the sake of procuring sustenance, entertainment, or upholding tradition and under recognition that procuring sustenance, seeking entertainment, or upholding tradition is what is good for me to do. What explains why I am hunting is my having set a particular end to myself and having recognized a principle of action according to which I should act in order to secure this end.

We are now able to elucidate the claim that respect is a non-pathological feeling. Moral respect is distinguished as non-pathological feeling not by virtue of the representation (the moral law) which affects a putative “faculty of sensibility” having its source in reason. Instead, non-pathological feeling is the awareness of how things promote or hinder our *rational desire*, or how they promote or hinder *us* as free, rational and embodied beings.

We can also clarify the precise sense the feeling of moral respect is nothing but the conscious determination of the human will. Since the principle that is the moral law is the form of will, the form of our practical reason,²¹ the state of a mind of a will determined by the moral law will necessarily consist in the awareness of this *self-agreement*, and this awareness is nothing but the primary mode of the feeling of moral respect. If instead the practical subject determines herself to deviate in her actions from the moral law, to prioritize the gratification of non-moral incentives [*Triebfedern*], then the subject will be in a state of awareness of self-disagreement. In either case moral respect will consist in essentially an awareness of self, whereby the subject will be revealed to herself in her moral character, i.e., in agreement with or in opposition to the moral law which commands necessarily and universally.

§9. Moral Respect as Practical Self-Consciousness

Thus we see that the intellectualist is right in insisting that no special feeling itself motivates the agent to perform morally worthy action, but she is wrong to deny feeling room in the account of moral motivation. By securing a distinction between pathological and non-

²¹ For the definitive articulation of this idea see Engstrom, *The Form of Practical Knowledge*.

pathological feeling that is grounded in the distinction between sensible and rational desire, we are able to secure feeling a role in an account of practical cognition while avoiding the intellectualist's worries of heteronomy.

Feeling is revealed as the fundamental mode of awareness of ourselves as moral beings, subject to moral demands, an awareness that is necessary for the performance of morally evaluable actions. Furthermore, Kant claims that the object of the feeling of moral respect is the moral law. This does not, however, mean that the object of this feeling has some sort of fully determinate content on its own; for the moral law, on Kant's account, is nothing but the *form* of our will. This means that the moral law is the *fundamental principle* of our rational, free activities. If the feeling of moral respect has for its object not a particular end or object but this principle, then it is the mode of awareness or the form of one's activity. It is, in other words, the mode of *self*-consciousness that characterizes moral being.

In the remaining sections I will demonstrate this claim by arguing that a particular way of being self-conscious is required to for the performance of free action.

§9.1. Moral Self-Consciousness is Necessary for Morally Evaluable Action

According to Kant the essence of the rational self lies in self-consciousness. The determination of the self qua self are representations. The subject has them by thinking them, and its thinking them is constituted by its awareness of itself as thinking them. The exercise of mental capacities is grounded in a consciousness of the self that is not a consciousness of another accompanying representation. The consciousness of the self is the self-awareness of the subject *as* the ground of rational activity: in self-consciousness the subject knows herself as capable of

engaging in rational activity and as the ground of unity of the indeterminately various possible actualizations of its capacities. For example, in performing an inference it is my consciousness of myself as the ground of the various judgments and thereby of myself as the ground of their unity that warrants me in drawing the conclusion.

The exercise of the will is the exercise of practical reason. This is the capacity to determine oneself to act for ends one sets oneself and in accordance with the representations of principles. It must therefore itself be a self-conscious activity. It follows that if I am to act freely I must know myself as capable of acting freely, and in acting freely must recognize myself as the ground of the capacity's various actualizations. However, the form of self-consciousness required for the exercise of practical reason cannot be reduced to either of the two chartered forms of self-consciousness in Kant's system: the practical mode of self-awareness is not identical with either empirical or pure apperception.

1. In empirical apperception I know myself as capable of affections in time and space, i.e., as capable of affection by objects whose existence is *independent* of my being affected by them. In practical self-consciousness, however, I must know myself as the *cause* of the objects of my representations (of principles). Nothing in the knowledge of myself as capable of affection secures that I am capable of bringing into existence objects as their cause.

2. In pure apperception I know myself as the ground for the spontaneous rational activity which grounds thinking. In pure apperception I am thus revealed to myself as the spontaneous ground of my representations, as free in thought. However, I am not thereby conscious of myself as efficacious. I do not know myself as capable of bringing into existence the objects of my representation, objects that I may subsequently know in theoretical cognition as independent of

my being affected by them or representing them in judgment. Thus I do not know myself as capable of action.

The self-awareness required for the exercise of practical reason must make manifest the self as ground of action and therefore make manifest the self in its character as spontaneously efficacious, as capable of bringing into existence the objects of my representations of principles. Moreover, for the action to be morally evaluable, one must know oneself as responsible for it, and therefore as free. This requires knowing oneself as subject to a principle one has set for oneself, in freedom from external determination. This is precisely what is provided for by Kant's characterization of the feeling moral respect:

What I cognize immediately as a law for me I cognize *with* respect, which *signifies merely consciousness of the subordination of my will to a law without the mediation of other influences on my sense.* (G 4:401n)

Thus, moral respect, the conscious recognition of the moral law as the law for me, is nothing but the constitution and making manifest of the self as a rationally desiring being, i.e. efficacious and free.

§10. Other People and Particular Feelings: The Revelation of Practical Demands on Moral Agents

Moral respect, Kant says, is for the moral law. The moral law, we've seen, is not an object in the world to be encountered on some determinate occasions and not others, nor is it a representation of an idea of reason that we bring to consciousness at will. The moral law is the form of our practical reason—the *fundamental principle* of our rational, free activities. If the

object of moral respect is the moral law, moral respect's object is the form, the essence and principle, of our own selves in our most essential determination, that of moral beings.²²

However, the moral law is not only the principle of *my* individual activity as free and rational, but also the principle of the activity of every other person—all-pervasive, it is the fundamental principle of everything touched by practical rationality. It follows that human feeling is not only the mode of awareness of ourselves as practical agents—the constitution and making manifest of the self as a rationally desiring being, i.e., embodied, free, and efficacious—but equally the mode of awareness of other persons, as well as to our shared forms of activity. Thereby, moral respect, as the mode of self-consciousness which grounds our awareness of ourselves and others as moral beings.

We can further see how all particular characteristically human feelings are manifestations of this unique mode of *self*-consciousness. The capacity to sustain oneself as a practical agent, i.e., as a free, embodied and efficacious, requires intervening in the world in particular ways. This in turn requires assessing the *practical* status and relevance of the features of our environment. Therefore, there is a need for a *receptive* capacity, which reveals the demands of reason on us, as embodied in the world around us, *on particular occasions*. This calls for a capacity of feeling, for as we've seen, Kant holds that feeling is the mode of awareness which reveals how the objects of our representations agree or disagree with our essential form and therefore reveal how those objects would promote or hinder us from freely determining ourselves to act. Since this receptive capacity must be grounded in a self-understanding of oneself and others as free and therefore subject to moral demands, this capacity could be none other than that to which Kant refers to as the feeling of moral respect.

²² See *Religion* 6:26.

For example, if I recognize a fellow human being has been wronged and feel indignation, this would be a reflection and revelation of the moral fact that the situation ought to be remedied and moreover that it is perhaps up to me to do so. This feeling would reflect and reveal my understanding of myself as a moral being that is concerned with what is the right thing to do, and as one who is responsible for her actions. The fundamental mode of self-consciousness that is moral respect is thus in turn the basis of all particular feelings, concrete exercises of a rational receptive capacity, which reveal the demands of reason on us *in concreto*.

§11. A Brief Remark on Other Feelings

To claim that moral respect names the fundamental structure of human feeling requires saying something about how to account for feelings that are not obviously moral. I cannot develop this part of the account in any detail but I hope the following short comment can provide a sense of what direction such a development might take. To be aware of how conditions outside us would promote or hinder us from freely determining ourselves to act, we need likewise to be aware of, but not thereby be determined by, conditions outside us that might affect us qua being embodied, not self-sufficient, and vulnerable. The feelings that reveal our dependence and vulnerability, are to be understood as revealed to us in relation to our moral being. Therefore, these too will be revealed *through* moral respect, the fundamental revelation of the self to itself as a practical agent.

§12. Conclusion

I have argued that as long as an account of the role of feeling in moral motivation seeks to legitimize the role of feeling in moral motivation without recognizing that the very idea of the capacity to feel must be conceived as an integral part of the exercise of practical reason, the involvement of feeling will always remain an extra-rational condition for the performance of moral action, and therefore contradictory to Kant's account of the moral worth of our actions.

Moral respect should be understood as the distinctive *human* capacity for feeling. This should not be understood as a standalone faculty of feeling, whose operations are intelligible on their own. Instead, feeling in a practical agent is to be understood as a way of being morally self-consciousness: the way in which we know ourselves as free, embodied and efficacious. This is, in turn, a mode of receptive, moral, awareness of a rational agent whereby he becomes aware of how her own activity as well as the activities of others stand with respect to her essential end: to act for ends she sets for herself and in accordance with principle. Thus, moral respect, the distinctive human capacity for feeling, emerges as the form of self-consciousness *constitutive of practical agency*.²³

²³ Carla Bagnoli has similarly argued that respect is “the emotional attitude that is constitutive of rational agency.” Bagnoli, “Emotions and the Categorical Authority of Moral Reasons,” in *Morality and the Emotions*, ed. Carla Bagnoli (New York: Oxford University Press, 2011), 33. Likewise, Oliver Thorndike, attending to the systematic role played by moral respect in Kant's account, has claimed that “moral feelings are *essential* to autonomous agency—not merely epistemological or motivational means to moral ends. [...] Moral feelings are dispositions that should be cultivated for their *own* sake (their cultivation is *morally* obligatory), because they are essential—not merely supportive—to autonomous agency. They are not optional instrumental means that facilitate moral action or help to ward off temptations to trespass the moral law.” Thorndike, “Kant's Transition Project” (paper, Third Biennial North American Kant Society Meeting, Emory University, Atlanta, GA, May 2016). I take my account to be a further determination of these insightful interpretive suggestions.

Chapter 3

Moral Character

§1. Introduction

In the first chapter, we began our investigation by critiquing Henry Allison's widely accepted reading of a passage in Kant's *Religion within the Bounds of Bare Reason*, a reading that he titled "Kant's Incorporation Thesis." According to the so-called Incorporation Thesis, a practical subject can freely act on a sensible desire, since it is in an act of spontaneity that a sensible desire is deemed an appropriate basis of action. The reflective mind, as Christine Korsgaard claims, must endorse the desire before it can act on it: "It must say to itself that the desire is a reason." I directed my critique against the assumption, shared by all the proponents of IT, that in Kant's account the involvement of receptivity in our desiring involves attributing an end-suggesting role to deliverances of a purely sensible faculty, one whose exercise does not require the exercise of rational capacities. I argued that this is philosophically unsatisfying and therefore that Kant would have erred to hold such commitment. I also argued that no such commitment is espoused in the passage on which Allison bases his reading. In the previous chapter I offered an alternative account of how to understand the involvement of receptivity in our desiring. This leaves us with an outstanding debt: a positive reading of Kant's claim that if the practical subject is to be determined to an action by a *Triebfeder*, that *Triebfeder* must first be admitted into the subject's maxim.

To recall, the key passage reads as follows: "The power of choice [*Willkür*] has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*]

except insofar as the human being has admitted the incentive [Triebfeder] into his maxim (has made this a universal rule for himself, according to which he wills to conduct himself)” (R 6:23–4).

I showed that the statement’s surrounding context makes clear that Kant is not speaking of the subject’s ability to step back from particular desires and assess whether they happen to provide adequate reasons for action (whether they can be made into objects of the will). Kant is not speaking of admitting a particular desire into one’s particular maxim, but of admitting one of two *kinds* of *Triebfeder* (namely, either sensible or moral) into the *Gesinnung*-constituting maxim, a maxim of choice of *the utmost generality*.

This maxim of utmost generality grounds the agent’s moral attitude, her *Gesinnung*. The *Gesinnung* is not a particular maxim or fleeting attitude, it is instead a single rule that the power of choice “makes for itself” which guides and unifies its determinate exercise in the adoption of particular good or evil (unlawful) maxims: “The attitude [*Gesinnung*], i.e., the first subjective basis for the adoption of maxims, can only be one, and it applies universally to the entire use of freedom” (R 6:25). Specifically, the supreme maxim in *conformity* with the law is the “formal basis” for good actions, the supreme maxim in *opposition* to the law is the “formal basis” for any unlawful deed or any which is lawful but “done from incentives [*Triebfedern*] that do not consist in the law itself” (R 6:31).

Since all actions are evaluated morally in terms of their maxims and all maxims are evaluated in terms of the basis of their adoption, the *Gesinnung*, this maxim of utmost generality must itself be adopted in a free act. This fundamental principle of choice determines a person’s moral character, their *Gesinnung*. Thus in speaking of admitting an *Triebfeder* into one’s maxim

Kant is speaking of admitting a *Triebfeder* into one's most general maxim. In the first chapter, I meant only to substantiate the interpretative claim that the passage Allison singles out does *not* refer to practical subject's ability to step back from concrete desires, recognize them as desires for concrete ends, reflect upon those ends, and choose whether or not they might serve as appropriate bases for action, i.e., whether it would be reasonable to pursue them. I promised I would return to these remarks of Kant's in order to elucidate them in greater detail. In order to do that we must first attend to Kant's characterization of the *Gesinnung* (§2–4). Only then will we be able attend to the *Gesinnung*'s characterization as a principle that admits incentives [*Triebfedern*], moral or sensible, into itself (§5–7).

§2. The Puzzle

The notion of “the first subjective basis for the adoption of maxims,” features in Kant's early work, the *Groundwork for the Metaphysics of Morals*, under the name “the subjective determining ground of the will,” which he says is “the maxim of complying with such a law even if it infringes upon all my inclinations” (*G* 4:401). Very importantly, in the same passage Kant identifies the subjective determining ground of the will with the feeling of pure respect for the practical law (*G* 4:400), thereby implying at least a very close connection—if not an identity—between the maxim of complying with the moral and the feeling of pure respect for the law. I will be able to explore this relation later in the chapter, as well as in the next.¹

¹ For a review of Kant's use of the concept, see G. Felicitas Munzel's extensive discussion in *Kant's Conception of Moral Character: The “Critical” Link of Morality, Anthropology, and Reflective Judgment* (Chicago: University of Chicago Press, 1999), 57–70, 164–174. In the next chapter I will address her interpretation directly, where I will argue that she fails to appreciate the role of feeling in the determination of moral character.

The subjective determining ground of the will in its guise as the *Gesinnung*-constituting maxim receives its most detailed treatment in the *Religion*. There Kant claims that in order for moral evil (or good) to be imputable to the person, the adoption of the *Gesinnung*-constituting maxim, the act whereby the power of choice makes for itself, for the use of its freedom, a rule, i.e., the act by which it constitutes itself as good or evil, must *itself* be free, imputable to a person as its originator. Kant's characterization of this act, however, makes it difficult to see how could this rule-adoption be considered an imputable act at all.

On the one hand, a practical subject's *Gesinnung*, the subjective basis of the use of her freedom, precedes any deed [*That*] that strikes the senses (*R* 6:20). On the other hand, as imputable, the *Gesinnung* must itself be thought of as a free act.

This subjective basis itself, however, must always in turn be an act [*Actus*] of freedom (for otherwise the use or abuse of the human being's power of choice in regard to the moral law could not be imputed to him, and the good or evil in him could not be called moral). (*R* 6:21)

This act is different than any other, for it is itself baseless: "Now, concerning this maxim it must not be possible to go on asking what is the subjective basis, in a human being, for the adoption of this maxim rather than of the opposite one" (*R* 6:21).

Like any other free act this one cannot be based in natural laws: "For if ultimately this basis itself were not a maxim any more but a mere natural impulse, then the use of freedom could be reduced entirely to determination by natural causes—which, however, contradicts freedom" (*R* 6:21). But its basis is not another maxim, either, for otherwise a regress would ensue: "The subjective basis or the cause of this adoption cannot again be cognized (although

inquiring about it is unavoidable), because otherwise one would in turn have to adduce a maxim into which this attitude had been adopted, and this maxim must likewise have its basis in turn.”²

Inexplicable in terms of natural causes or other maxims, Kant claims that “the human being himself is the originator of this character” (R 6:21). He adds: “It is laid at the basis (in earliest youth, back at the point of birth) prior to any use of freedom that is given in experience, and thus is conceived as present in the human being simultaneously with birth—though not exactly as having birth as its cause” (R 6:21–2). A little later Kant repeats that the human being is “the originator” of her *Gesinnung*, though “it has not been procured over time (that one or the other he *has always been, from his youth*)” (R 6:25), and adds “we cannot derive this attitude [*Gesinnung*], or rather its supreme basis, from any first act of the power of choice in time [*irgend einem ersten Zeit-Actus der Willkür*]” (R 6:25).

Kant’s characterization of the founding of moral constitution as our act raises the following questions: How can the principle that grounds our moral constitution, the basis of maxim selection, i.e., action, which is prior to any particular use of freedom, be itself an act? What kind of act is an act whereby one adopts a principle of action not at a time or over time? And, moreover, in what sense is this act anything that a person may intelligibly be said to be the subject of?

§3. Two Significations of Deed: Sensible and Intelligible

Kant raises the question thus: Since “nothing is morally (i.e., imputably) evil but that which is our own *deed*,” the subjective determining ground for the use of our freedom, our

² See also R 21, footnote 31.

Gesinnung, must itself be a deed. But since it “*precedes every deed*” it is “*itself not yet a deed*” (R 6:34). Kant’s solution to the threat of unintelligibility is to introduce a distinction between two significations of “deed,” which can both “be reconciled with the concept of freedom” (R 6:31):

The expression *deed* as such can apply both to that use of freedom through which the supreme maxim is admitted (in conformity with the law or in opposition to it) into the power of choice, and to that use where the actions themselves are (in terms of their matter, i.e., as regards the objects of the power of choice) performed in conformity with that maxim. [...] The first is an intelligible deed, cognizable by bare reason without any time condition; the second deed is sensible, empirical, given in time (*factum phaenomenon*). (R 6:31)

Both kinds of deed are cases of a “use of freedom,” but while one is given in time, the other lacks temporal conditions. This is one characterization of the *Gesinnung*’s atemporality.³ The second (although not on that account distinct) characterization of the *Gesinnung*’s atemporality is that it must be thought of as an absolute unity, and, in particular, not itself divisible in time. Kant writes,

The morally subjective principle of the *attitude* [*Gesinnung*], by which his [a person’s] life must be judged, is (as something suprasensible) of such a kind that its existence cannot be thought as divisible into time segments but can be thought only as an absolute unity; and since we can infer the attitude only from the actions (as appearances thereof), life will, for the sake of this assessment, be considered only as a *unity of time*, i.e., as a *whole*. (R 6:70 fn127)

How this distinction is meant to resolve the philosophical tension is far from obvious.

Matthew Caswell summarized the worry as follows:

To many readers, this surely seems like Kant attempting to avoid paradox by falling back on grossly unjustified metaphysical speculation. With the notion of a purely intelligible act of timeless legislation, has Kant abandoned the modest conceptual analysis of moral concepts, in order to have things both ways?⁴

³ Cf. *KrV* A539/B567: “Now this acting subject, in its intelligible character, would not stand under any conditions of time, for time is only the condition of appearances but not of things in themselves.”

⁴ Matthew Caswell, “Kant’s Conception of the Highest Good,” *Kant-Studien*, 97, no. 2 (2006): 199.

For the distinction to be useful we must examine whether Kant can answer the following three interrelated questions (or whether we can on his behalf):

1. *In what sense are both kinds of deed “acts”?*

The distinction between kinds of deed, the sensible and the intelligible, is supposed to map onto the distinction between the act of the adoption of the subjective determining basis for the selection of maxims, i.e., the *Gesinnung*, and the particular deeds which it precedes. In particular, the question requires attending to the following:

1.1. *In what sense is the intelligible deed an act? How can an act be performed neither in nor over time?*

1.2. *In what sense is the sensible deed an act?*

It would not do to say that sensible deeds are merely happenings explicable in terms of natural causal laws, for then they would not be acts at all. In the passage quoted before last (*R* 6:31) Kant speaks of the sensible deeds as “actions,” whose matter are the objects of the power of choice. Objects of the power of choice are ends, and to move in order to procure objects of the power of choice is to act. (Mere happenings of nature are not performed through the power of choice, and their matter are not objects of the power of choice, i.e., particular ends.) This is not, however, a verbal slip on Kant’s part. Throughout the text, Kant repeatedly claims that although sensible acts cannot be evaluated in terms of their moral worth, they can be evaluated in terms of their legality or lawfulness, (they may be “legally good (lawful)” [*gesetzliche gute (legale)*], *R* 6:30; See also, for example, *R* 6:20, 36, 47–48, 99). This means that one can determine whether or not a sensible act has been performed in accordance with the moral law (although not whether or not it was performed from a consciousness of the moral law). From this it follows that sensible

deeds are, first, not mere changes of location in time, and second, that they are actions, movements for the sake of ends. (The latter implies the former but it is worthwhile to address them separately and in this order.) Mere movement does not constitute a subject of legal evaluation. A mere happening cannot succeed or fail to accord with a principle that is the moral law. Since empirical causal principles are laws for how things *are*, an event that fails to accord with an empirical causal principle is not “unlawful,” but one to which that law is simply not applicable. In order for something to fail to accord with a law such that it may be considered “unlawful,” that law must apply to it such that its failure to accord with the law does not simply render the law inapplicable to it. The law must be a normative law, a law of how things, in some sense, should be. Thus, although sensible deeds are opaque with regards to their moral worth, that they can be either lawful or unlawful, implies that they are normatively evaluable and are therefore not mere happenings. Moreover, the moral law is a law that applies to principles of action. The only thing that can accord or fail to accord with the moral law is therefore something that stands under a principle of action, i.e., an action.

2. *How can both kinds of act: the intelligible and the sensible, be, as Kant claims, “reconciled with the concept of freedom”?*

This question is related to, but, for our purposes here at least, not reducible to, the former question. My way of formulating the first question leaves undetermined (dialectically) whether in acting I am setting ends for myself or am motivated by a brute desire that I seek to gratify by exercising instrumental reasoning (seeking means to a desired end). For Kant, however, human action is free and therefore the will is never simply determined by an incentive [*Triebfeder*] from the outside, so to speak. In the case of morally worthy action the moral law itself serves as the

incentive [*Triebfeder*] for action and in the case of heterogeneously motivated action, an incentive [*Triebfeder*] cannot determine a moral agent to act if she has not herself admitted it into her *Gesinnung*-constituting maxim. This means that both significations of deed must be intelligibly conceived of as free. Indeed, Kant refers to the sensible deed as a “use of freedom that is given in experience” (*R* 6:22), he repeats, in the passage quoted before last (*R* 6:31), that both uses of “deed” are performed through a “use of freedom,” and claims that *both* notions must be “reconciled with the concept of freedom.”

3. *How are the two acts related? In particular, how are sensible acts grounded in the intelligible act? How are acts in time grounded in the act that is not subject to time conditions?*

These two kinds of act must not only both be intelligible as acts and consistent with freedom but must be related to one another such that the intelligible deed can serve as the “formal basis” (*R* 6:31) of the “sensible deed.” Kant claims that the maxim of utmost generality “lies at the basis” of the actions whose performance is observable in time; it is, recall, a rule adopted by the capacity for choice for its own exercise, by serving as a rule “for the adoption of good maxims or the adoption of evil (unlawful) ones” and thereby the basis for the performance of any action in time (*R* 6:21). What sense is there to be made of this characterization of the relation between the two significations of deed?⁵

⁵ My presentation of the three questions and in particular my drawing attention to Kant’s characterization of a sensible deed as a free action invites the question of the relation between this distinction and Kant’s distinction between phenomena and noumena in his earlier critical works. It is beyond the scope of this essay to address the latter question. My goal is to elucidate the distinction between the two significations of deed, intelligible and sensible, as Kant introduces them in the *Religion*. I hope to address the implications of the apparent tension between this characterization and his earlier commitments in future work. Briefly, I aim to explore how Kant’s elucidations in his mature practical works might help us reevaluate the distinction as it was presented in the theoretical works.

In what follows I will attempt to answer these questions in a way that also addresses the question of how we can secure the intelligible imputation of actions, in both significations, to the practical subject, and in particular, the imputation of the *Gesinnung*—the maxim of utmost generality which functions as the rule for the adoption of all determinate maxims.

§4. Aristotelian Elucidations

In what follows I will introduce a characterization of rational activity that is expounded by Aristotle in the *Nicomachean Ethics* and the *Metaphysics*. Aristotle's treatment of the notion of *energeia* provides a logical framework for thinking about rational activity that is not subject to time conditions, that is unified and whole in itself, but that can play the role of grounding activities in time. To see this, we must attend to the contrast between *energeia*, actuality or activity, and *kinesis*, change or movement (§4.2). To put ourselves in the position to do so, however, we must briefly attend to another distinction in Aristotle, that between *dynamis* as potentiality or being-in-capacity and *energeia* as actuality or being-in-activity (§4.1).⁶

This characterization will begin to shed light on the Kantian distinction between the two significations of deed, and in particular on Kant's characterization of the intelligible activity that is not subject to time conditions. I note several advantages of this approach.⁷ I then introduce the question of what kind of intelligible deed is the founding of a *Gesinnung*—the subordination of two kinds of *Triebfeder*, moral or sensible, to one another (§5). At this point I offer a new

⁶ I am following in a tradition that returns to Aristotle for insight into the logical character of action. Anthony Kenny was one of the first to notice that verbs of action exhibit logical features of their own that had gone largely unrecognized. Kenny, *Action, Emotion and Will*, 2nd ed. (1963; repr., London: Routledge, 2003).

⁷ I do not seek to stake interpretative claims with regard to Aristotle's works.

interpretation of Kant's notion of the "sensible *Triebfeder*" as logically dependent not on the idea of our non-rational, so-called "animal" self, but as dependent on the idea of the sensible deed (as *kinetic* activity) as I elucidate it in §4. Finally, I argue that the claim that the intelligible deed that is the constitution of a *Gesinnung* consists in the subordination of two kinds of *Triebfeder*, moral or sensible, to one another, implies that the *Gesinnung*-constituting maxim, in the case of a good will, determines the subject's subordination of all kinetic activity to energetic activity grounded in the consciousness of the law (§7).

§4.1. *Energeia* and *Dynamis*

In his recent monograph on Aristotle's *Metaphysics Theta*, Jonathan Beere argues that Aristotle's *energeia* and its contrasting term, *dynamis*, stand for "two ways of being": being-in-capacity (*dynamis*) and being-in-activity (*energeia*). These ways of being cut across the different ways of being that correspond to the different categories of predication (quality, quantity, time, place, etc.). Generally speaking, however a thing can be, it can be that way in one of two ways: it can be what it is "in capacity" or it can be what it is "in activity." A being may have the capacity to be a particular way and either not exercise that power or else exercise it and be that particular way. Something can be F (say, a bearer of a particular quality) in capacity or in activity. Something is F in capacity by virtue of the capacities it possesses. To be an F in activity, however, is not simply a matter of possessing or bearing capacities that go with Fs but is a matter of exercising these capacities.⁸ For example, as she is watching a football game in the stadium, a dance instructor is a dance instructor in capacity, however, when she is teaching a class of

⁸ Jonathan Beere, *Doing and Being: An Interpretation of Aristotle's Metaphysics Theta*, Oxford Aristotle Studies (Oxford: Oxford University Press, 2009), 155–218.

students to dance, the dance instructor is a dance instructor in activity. These two ways of being are interdependent: activities are essentially identified, qua activities, as exercises of capacities to perform those activities, whereas capacities are identified, qua capacities, as capacities to perform activities. Moreover, each determinate capacity is identified, qua the capacity that it is, in terms of the determinate characteristic activity of which it is a capacity. We will return to this distinction explicitly in what follows but for now is enough to enable us to introduce Aristotle's apparently orthogonal distinction between two kinds of activities: *kinetic* and *energetic*.

§4.2. *Energeia* and *Kinesis*

For Aristotle, activities and their associated capacities can be separated into two kinds: *kinetic* and *energetic*. Kinetic capacities are capacities whose activity is a change: a kinetic capacity is the capacity of a thing to be at one time in a certain way that is different than the way it is at another. When a *kinesis* is in activity, a change takes place over time that is directed toward an end. The exercise of the capacity is the bringing about of the change. When the end is reached, the change is completed, and the capacity to change in that particular way is no more. Energetic activities do not aim at ends that are outside of themselves, and therefore their capacities are not capacities for undergoing change. An *energeia* is the activity of maintaining a form of activity. When a capacity for a certain *energeia* is in activity, the end of the activity is contained within it by virtue of its very exercise and as such an *energeia*'s end is already reached at every instance of its exercise, and is not thereby extinguished.

This characterization is underscored by a grammatical distinction. Kinetic activity, whose concept includes a contrast between being under way and being complete, is signified by a verb

that is predicated under the contrast of aspect: perfective and progressive.⁹ If you are now F-ing, you have not already F-ed. For example, returning borrowed money is a case of *kinesis*. As long as I am busy procuring the funds or delivering them to you, I have not yet returned what I owe. If you have F-ed, you are F-ing no more. The capacity to return that particular sum of borrowed money is exhausted in handing the money over to its rightful owner. I can't pay you the same debt twice.

Grammatically, while the end of kinetic activity is signified by a verb predicated under the contrast of perfective and progressive aspect, the end of an energetic activity is signified by a verb that is not predicated under a contrast of aspect, but represents the act as always already complete.¹⁰ If F-ing is an activity, you are in activity insofar as you have at every moment of F-ing already F-ed. For example, seeing is an activity. If I am seeing, exercising my capacity for sight now, I have already seen. Likewise, being true to one's word is an activity. If I am being true to my word now, I have been true to my word.

Now, Aristotle's distinction between energetic activity and kinetic activity provides a framework within which we can understand Kant's characterization of the two significations of deed—sensible and intelligible—and the relation between them. Sensible and intelligible deeds are not acts occurring on separate metaphysical (or theological) planes, but activities that bear unique logical forms, that of *kinesis* and *energeia*.

The exercise of a *kinetic* capacity is subject to time conditions: the ends of kinetic capacities specify conditions for “getting started” and “being done.” The successful exercise of a kinetic capacity essentially includes exercising it in the appropriate circumstances. To say that I

⁹ Sebastian Rödl, “The Form of the Will,” in *Desire, Practical Reason, and the Good*, ed. Sergio Tenenbaum (Oxford: Oxford University Press, 2010), 147.

¹⁰ *Ibid.*, 148.

know how to use a linguistic expression means that I am able to use it appropriately, on the right occasion: to an audience of speakers of the language, in the correct order in a sentence, as a response to particular circumstances, for a given purpose, etc.

To see how the exercise of *energetic* capacities is not subject to time conditions, we must consider how capacities for *kinetic* activity can be grounded in the performance of *energetic* activities. Consider the following example. My having returned two particular debts on different occasions might be a coincidence: explained each time by some particular end that I happen to desire to which returning the debt was a means (e.g., on one occasion I wished to impress a colleague, on another I feared retribution). If my doing the same is not a coincidence, both actions must be explained by one principle of action. That principle of action can be a desire for a particular end or a law. If the principle of the action is a single desire for an end (e.g. gaining the confidence of a particular individual), it is no accident that I performed both actions in the sense that both are explained by one desire that I happen to have. But that I performed identical actions is accidental insofar as it is a contingent fact that both actions happened to advance that desire. It is not internal to my desire to gain someone's confidence that I hold true to my word. If on the whole gaining their confidence would be better served by my breaking a promise, I will do so. (If I don't, the explanation would have to appeal to another principle of action.)

However, a principle of action could be a law according to which I act: "repay one's debts" or, more generally, "keep one's promises." If this is the case, then my having repaid the two debts would be explained by my commitment to the principle of debt repayment or promise keeping. I repaid the particular debt because I remain true to my word. This is the sense in which the exercise of an *energetic* capacity is not subject to time conditions. In performing an *energeia*

I am committed to a law. A law is general with regard to time: it has no temporal limit.¹¹ Therefore, the ends of energetic capacities do not specify any conditions for “getting started” or “being done.” Put another way, being committed to being true to my word is something one does always or never. If I’m intentionally only true to my word on occasion, the principle I am, strictly speaking, *committed to* is not “keep one’s promises” but “keep one’s promises unless C,” where C stands for a set of conditions in which I am not committed to keeping my word. If I’m only true to my word on occasion I am committed to keeping my word only conditionally, and therefore, strictly speaking, not at all.¹²

Thus, an energetic activity is not simply the sum of those concrete kinetic activities: it rather gives them unity. As such the logical characterization of energetic activities helps elucidate Kant’s claim that the intelligible deed that is the *Gesinnung* is an absolute indivisible whole. A law is a time general principle, the capacity which is exercised in maintaining a commitment to this principle may be exercised in indeterminately many possible circumstances. Abstracting from those particular exercises we can say of the subject practically committed to the law that she has unconditionally maintained a unified commitment to the law. Thus energetic activities are instantiated in kinetic activities, i.e., particular actions. For example, the activity of being true to one’s word is instantiated in particular instances of keeping particular promises and the specific measures one takes to make that possible.¹³

¹¹ For a discussion of these issues, see Sebastian Rödl, *Self-Consciousness* (Cambridge, MA: Harvard University Press, 2007), chap. 2.

¹² For a discussion of the generality of maxims and the *Gesinnung*-constituting maxim, see Matthew Caswell, “Kant’s Conception of the Highest Good, the *Gesinnung*, and the Theory of Radical Evil,” 193.

¹³ Versions of the claim that maxims stand to one another in a hierarchical system have been defended by Lewis White Beck in *A Commentary on Kant’s Critique of Practical Reason*, 2nd ed. (Chicago: University of Chicago Press, 1996), 118; Christine Korsgaard, *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996), 58; Onora O’Neill, *Constructions of Reason*:

Moreover, the formal distinction between *energetic* and *kinetic* activity provides us with a new perspective on Kant's claim that it is the intelligible deed that is the *Gesinnung* which is the ultimate subject of moral assessment. Kant, recall, claims that sensible deeds cannot be evaluated for moral worth, only legality. The moral assessment of any action principle adopted by the agent must ultimately be referred to the moral status of the maxim of utmost generality, the *Gesinnung*. We see how this claim is not a dogmatic commitment to the moral opacity of practical agents, grounded in epistemic caution (or skepticism) or moral humility. For this claim follows with necessity from the manner in which kinetic activities depend for their full determination on the energetic activities which they instantiate. Sensible deeds do not wear their moral worth on their sleeve because a sensible deed is a *kinetic* activity, performed locally for the sake a particular end. But for Kant, it is not particular ends that determine an action's moral worth. The same particular end, considered in abstraction from the context in which it is pursued, could be pursued in either a morally good or bad action, i.e., its ultimate principle might have been adopted in recognition of the authority of the moral law, or in deviation from it. As we have seen, repaying a particular owed debt, considered in abstraction from the context in which the action is performed, might be explained by a commitment to the law: "repay one's debts" or "keep one's promises" or by fear of retribution. Without knowing which principle one ultimately instantiates in performing the particular, kinetic or sensible deeds, their moral worth remains indeterminate.

Moreover, without knowing which principle one instantiates in exercising one's capacity for kinetic activity, the local end pursued remains indeterminate. *How* I should perform the

Explorations of Kant's Practical Philosophy (Cambridge: Cambridge University Press, 1989), 83–85; Henry Allison, *Kant's Theory of Freedom*, 91–94.

action that we're calling "repay the debt" is determined by what it is I'm doing in repaying it. If I'm seeking to impress a colleague I will repay the debt in a manner that would be visible to the colleague and only to the extent required to impress her (in effect, repaying the debt only accidentally). If I'm following a time-general principle, I will be concerned with repaying the debt in full and in mutual recognition. So it is not just that sensible deeds' ends are not morally evaluable, they are indeterminate in terms of their content. The determinate content of the end is given by the ends which that particular end is itself meant to promote, i.e., the principle of a particular action depends for its full determination on the more general principles of action which it instantiates.

Now, recall that for Kant the *Gesinnung* is a single rule that the power of choice makes for itself for the adoption of maxims, good or bad. Thus every action is grounded ultimately in the principle of utmost generality that unifies the adoption of all other principles. The *Gesinnung* is therefore the time-general principle of action that holds together all other time-general principles of action and thereby all concrete activities that are performed in consciousness of the commitment to those time-general principles. To determine *what* it is a subject has done therefore requires determining what the ultimate principle of her action is. For Kant, moral worth depends on whether an action was performed from recognition of the universal authority of the moral law. This is the same as saying that it requires determining how the action was ultimately motivated. Thus the *Gesinnung* is a principle of grounds of motivation.

§5. What kind of activity is the *Gesinnung* activity?

Aristotle's account of action helps to elucidate the distinction between intelligible and sensible deeds by providing a conceptual framework for thinking about the interrelation of two forms of activity, *kinesis* and *energeia*. In particular, we are able to see how an activity which is the maintenance of a commitment to a law (rule-following) may be considered as an absolute whole and not itself subject to time conditions, while in turn grounding (and unifying) diverse acts in time.

What kind of activity, however, is the activity that is the having of a moral *Gesinnung*? What is the law which one follows or to which one maintains one's commitment? The answer should have been simple. It should be the moral law, the imperative which commands us, in its various formulations in the *Groundwork*, to act only from principles that are fit for the task of universal legislation, or to treat others as ends and never only as means, etc. However, we find Kant claiming in the *Religion* that the grounding principle of the moral *Gesinnung*, is not the moral law or any other apparent principle of principles, but is instead a principle that orders the two kinds of grounds of desire [*Triebfedern*], moral and sensible, whereby compliance with the one is rendered the condition of compliance with the other.

"Ground of desire" translates the German *Triebfeder* (often, and therefore throughout this essay, translated as "incentive"). Kant defines a *Triebfeder* as "the subjective ground of desire," which is distinguished from the "objective" one (*G* 4:427). The objective ground of desire is internal to the desiring capacity as the principle that a proper exercise of the desiring capacity, i.e., acting, would follow in recognition of the authority of that principle's authority. When acting morally well, i.e., from recognition of the authority of the moral law, the law serves as its own *Triebfeder* (which is, as we saw in the previous chapter, is identical to the form of self-

consciousness designated by the feeling of moral respect). Human beings are capable of exercising their capacity of desire, i.e., acting, in ways that involve ignoring the authority of that principle. The subjective ground of desire is an answer to the question whether one acted from recognition of the moral law or otherwise. Thus, in the second *Critique* a *Triebfeder* is “the subjective determining ground of the will of a being whose reason is not already in virtue of its nature necessarily in accordance with the objective law” (*CpV* 5:72). If one acts otherwise than from a recognition of the authority of the law, one’s *Triebfeder* is, Kant says, “sensible,” i.e., one acts not from a recognition that an action agrees with the form of one’s will—one’s being practically rational—but in the expectation of sensible pleasure, an awareness of the action’s agreement with an aspect of one’s being that is not the form of our will.

The rule that the power of choice makes for itself for the use of freedom is not an imperative to follow the moral law or the principle of happiness (as the sum of all inclination). The *Gesinnung*-constituting maxim, the first principle of the power of choice, is a maxim that admits a kind of *Triebfeder* into the power of choice. The power of choice legislates for itself *how* it will be motivated.

The freedom of the power of choice has the quite peculiar characteristic that it cannot be determined to an action by any incentive [*Triebfeder*] *except insofar as the human being has admitted the incentive [*Triebfeder*] into his maxim* (has made this a universal rule for himself, according to which he wills to conduct himself). Only in this way is an incentive [*Triebfeder*], whichever it may be, consistent with the absolute spontaneity of the power of choice (i.e., with freedom) (*R* 6:23–4).

Kant complicates the picture by claiming that the human being, who is receptive to both kinds of *Triebfedern*, moral and sensible, admits both into his maxim:

Through the power of his [...] innocent natural predisposition, [...] he is also attached to the incentives [*Triebfedern*] of sensibility and admits them (in

accordance with the subjective principle of self-love) also into his maxim. (R 6:36)

The moral *Gesinnung*'s quality is determined by which kind of *Triebfeder* is given priority:

The distinction as to whether the human being is good or evil must lie not in the distinction of the incentives [*Triebfedern*] that he admits into his maxim (not in the maxim's matter), but in their *subordination* (in the maxim's form): *which of the two he makes the condition of the other*. [...] the latter [moral *Triebfeder*] should be admitted into the universal maxim of the power of choice as the *supreme condition* of the gratification of the former [sensible *Triebfedern*] and as the sole incentive [*Triebfeder*]. (R 6:36)

The *Gesinnung*-constituting maxim determines which of the two kinds of *Triebfedern* is subordinated to the other. Moral goodness consists in respect for the moral law being admitted “as an incentive [*Triebfeder*], sufficient by itself, of the power of choice” (R 6:27). Moral evil consists in the latter, “sensible,” *Triebfedern* being admitted into the *Gesinnung*-constituting maxim “as by themselves sufficient for determining the power of choice” (R 6:36).

What does this characterization of moral *Gesinnung*—a power of choice which makes a rule for itself in which it *subordinates* one *Triebfeder* to another—amount to? In particular, in what sense are both *Triebfedern* admitted into the *Gesinnung*-constituting maxim? My own interpretation of Kant's account will emerge most clearly through consideration of a reading to which the text may seem to lend itself to but that is ultimately unsustainable.

The reading I will reject hinges on understanding *Triebfeder*, encouraged by the habit of translations, as a kind of “incentive.” Accordingly, a practical agent is aware of two kinds of possible incentives to act, and on every occasion chooses among them in accordance with the rule she has set for herself. She experiences, so to speak, two kinds of temptation. One kind arises directly from her being a possessor of practical reason: a course of action will promote her capacity to act for ends she sets for herself and in accordance with principle. Another kind of

temptation arises not from her being a practically rational agent but because a course of action agrees with her qua possessor of capacities that abstract from her being a practically rational agent (i.e., a being in whom reason can set ends, not only be used to calculate means to obtain the objects of non-rational desire). The moral law announces itself invariably to all practical agents; the sensible *Triebfeder* depends on a combination of innate and acquired dispositions. For example, hunger provides the agent with the incentive to eat, ambition with the incentive to exert herself in competition and moral respect with the incentive to do the right thing, e.g., repay an outstanding debt. What the agent is aware of as beckoning her is determined independently of the constitution of her *Gesinnung*. The individual agent determines which kind of temptation to gratify first: moral or sensible. The agent satisfies the other incentive only if it happens to coincide with the demands of the other.

I say “happens to coincide” and not “merely agree” since it is not obvious that the absence of conflict would ever be sufficient to justify the gratification of the secondary kind of *Triebfeder*. In the case of a morally good agent, the moral law would command imperfect duties to help others and develop one’s talents to the highest extent possible. Since those ends are indeterminate, one can never exhaust them. Insofar as gratifying the call of a sensible *Triebfeder* would come at the expense of the further satisfaction of an imperfect duty, the prioritization of the moral *Triebfeder* over the sensible one would preclude the satisfaction of the latter. Similarly, in the case of an evil agent, considering that a sensible being’s interest are best served by its conservation of unnecessary efforts, the pursuit of any end that is not sensible would stand in conflict with the goal of maximizing the satisfaction of sensible incentives and is therefore precluded.

Therefore, on this interpretation the admission of *both* kinds of incentives would amount to an empty claim.

However, this reading is falsified by Kant's account of the constitution of a *Triebfeder*. It is, recall, the subjective ground of desire. Desire is the capacity to be a cause through one's representations. It is not the capacity to suggest, or to tempt one to courses of action, but the capacity to act. Thus, *Triebfeder* is essentially defined by its role in action. A *Triebfeder* does not stand in causal relations to action; it names the ways in which an agent may be a cause in the world. *If* a subject fails to act from the moral law, it would be misleading to say that the moral *Triebfeder* failed to move the subject. Instead we should say that the authority of the moral law is not acknowledged and that therefore the *Triebfeder* for the act performed is not the moral law.

Kant's statements about the power of choice admitting *Triebfedern* into its *Gesinnung*-constituting maxim should thus *not* be heard as describing: (i) the adoption of incentives whose constitution is independent of whether they end up motivating the agent and (ii) the activity of conferring authority on some incentives over others. Instead, to admit an *Triebfeder* into a maxim is to *determine a structure of motivation*.

Kant confirms as much in his account of the fundamental determinations of the human being. In the first section of the first part of the *Religion*, Kant briefly presents three "elements of the determination of the human being" (R 6:26). He clarifies the significance of these predispositions thus: "By the predispositions of a being we mean the constituents, as well as the forms of their combination, that are required in order to be such a being" (R 6:28). He adds that these are "original," meaning that they "belong to the possibility of such a being necessarily" and

that all three “refer directly to our power of desire and the use of our power of choice” (R 6:28).

They are:

1. The predisposition [*Anlage*] to the *animality* of the human being as a *living* being;
2. To the *humanity* of him as a living and at the same time *rational* being;
3. To his *personality* as that of a being who is rational and at the same time *capable of imputation* [*Zurechnung fähigen*] [of actions to him] (R 6:26).

Kant goes on to define these in relation to the involvement of reason in their constitution: 1.

Animality is “a kind of self-love for which reason is not required” it is a “physical and merely mechanical self-love.”; 2. Humanity is rooted in reason “but only as subservient to other incentives [*Triebfedern*],” i.e. instrumental reason; 3. In personality alone is “reason practical on its own, i.e., legislative unconditionally.”

The predisposition to personality is “the receptivity to respect for the moral law, as an incentive [*Triebfeder*], sufficient by itself, of the power of choice” (R 6:27):

This receptivity [*Empfänglichkeit*] to mere respect for the moral law within us would be the moral feeling, which by itself does not yet amount to a purpose of the natural predisposition, but amounts to such a purpose only insofar as it is an incentive [*Triebfeder*] of the power of choice. [...] this becomes possible solely through the free power of choice’s admitting the moral feeling into its maxim, the constitution of such a power of choice is a good character. (R 6:27–8)

To be a person, i.e., a human being, is to be receptive to respect. As we saw in the second chapter, this amounts to the capacity to recognize oneself as capable of acting for ends one sets for oneself and in accordance to principle. Good character consists in actualizing this capacity. *Receptivity* to mere respect for the moral law can be considered a *Triebfeder* only if “the free the power of choice admits the moral feeling into its maxim.”

In the case of good character, the awareness of agreement and disagreement with one's being qua moral agent, i.e., a person, is sufficient for action. In the case of bad character, that awareness of agreement and disagreement with one's personality is not sufficient for action.

Note however that the separation of feeling and that feeling being admitted into the *Gesinnung*-constituting maxim and thereby becoming a *Triebfeder* does not mark two stages of motivation (first feeling, then motivation). It instead marks the fact that a morally evil person does not lose the capacity to feel moral respect, i.e., the awareness of agreement and disagreement with one's personality. It can, for example, manifest itself as a determination of the feeling of moral respect in feelings like guilt or regret. Thus, Kant says, "No human being is entirely without moral feeling, for were he completely lacking in receptivity to it he would be morally dead" (*MM* 6:400).

The activity of admitting *Triebfedern* into one's *Gesinnung*-constituting maxim is not primarily a matter of exercising a capacity to evaluate desires before acting on them, or a decision to gratify the pull of some incentives over others, but is instead the constitution of a structure of motivation. A person's *Gesinnung* answers the question: What will the person find a sufficient reason to act? (A capacity that, in the logically primary case, operates without reflection.)

What is there then to the idea of subordinating *Triebfedern* if *Triebfedern* are what *in fact* explains how you were moved to act? How can one act on two kinds of *Triebfedern* simultaneously, with one subordinated to the other? To answer this question we must attend to the constitution of the sensible *Triebfeder*. In doing so, we will see that to admit a *Triebfedern*

into a maxim is to *determine* a structure of motivation by determining one's structure of pursuing ends with respect to their kind: *kinetic* or *energetic*.

§6. The Sensible *Triebfeder*: the *Triebfeder* of the Sensible Deed

Thus far I have left the constitution of the sensible *Triebfeder* opaque, only characterizing it negatively: one acts on a sensible *Triebfeder* when one acts on ends that are not set by reason. We must say more if we are to understand what the subordination of a *Triebfeder* in a person's determination of a structure of motivation consists in. In particular, we ask, where do ends which are not set by reason come from? In Chapter One I offered a critique of an interpretation of Kant's account of the admission of *Triebfedern* into one's maxim, proposed by Henry Allison, and referred to it as Kant's Incorporation Thesis, on the grounds that it attributed a role to a capacity for receptivity which did not require for its intelligibility recourse to the exercise of rational capacities. This is a prohibition that an account of the sensible *Triebfeder* must respect: the ends of the sensible *Triebfeder* cannot be understood simply as the ends of our non-rational nature.

I propose elucidating the sensible *Triebfeder* without reference to any putatively non-rational elements in the constitution of our mental powers, in particular, without reference to a separate faculty of desire, namely, sensible desire, whose constitution is intelligible without reference to reason. Instead I propose elucidating the idea of a sensible *Triebfeder* with reference to Kant's account of the "sensible deed" which I introduced and elucidated in §3. Drawing on Aristotelian distinctions between *energeia* and *kinesis* I suggested that we should understand Kant's distinction between intelligible and sensible deeds along the lines of Aristotle's distinction between activities that contain their end in themselves and activities that aim at the

accomplishment of an end which is not attained in the activity itself, but which terminates it.

Using this distinction we can logically demarcate the character of the two kinds of pleasure that attend the exercise of the capacities for both kinds of activity.

Recall Kant's account of feeling as introduced in Chapter Two. Feeling is the awareness of agreement (pleasure) or disagreement (displeasure) of a representation with the activity of the subject. It is a relation of the subject to a representation, and through it to the object of the representation. Feeling, in its relation to desire in general—the capacity to act for the sake of ends one represents—is the representation of agreement or disagreement with, i.e., promotion or hindrance of a subject's exercise of a faculty of desire, her capacity to act. In order to elucidate the sense in which the feeling of moral respect is non-pathological, I focused in Chapter Two on Kant's claim that the faculty of desire in rational agents in general and human beings in particular is not merely an animal faculty of desire, but "a specially constituted faculty of desire" (*G* 4:428), "distinct from a mere faculty of desire" by being a "faculty of determining itself to action as an intelligence and hence in accordance with laws of reason independently of natural instincts" (*G* 4:459). I emphasized there that Kant is not denying that the will is a power of *desire* but claims that our power of desire is a special species of the genus "faculty of desire." The *human*, rational faculty of desire is a faculty of desire, the power to act for the sake of ends—but it is not merely that: it is the capacity to act for the sake of ends *one sets for oneself and in accordance with principle*, therefore in awareness of one's freedom from external determination.

But in attending to the distinction between sensible and intelligible deeds in terms of the distinction between *kinetic* and *energetic* activities we see that there are two ways in which we

could consider the exercise of a rational faculty of desire: as a capacity to act from time general laws, and therefore for ends that are contained in their activities, and the capacity to act on particular maxims, and therefore aiming at an accomplishment of an end which is not attained in the activity itself, but terminates it. This is the distinction, in Kant, between the capacity to act on hypothetical imperatives and the capacity to act on categorical imperatives (*G* 4:414–6).

Therefore, we can distinguish in this context between two kinds of pleasure, i.e., two kinds of awareness of agreement with oneself: awareness of agreement with an intelligible deed/energetic activity and awareness of agreement with a sensible deed/kinetic activity. The actualization of a capacity to act is the actualization of an end. Since the end of an energetic activity is contained in every instance of its exercise, the subject's awareness of agreement ("pleasure") of an intelligible deed with herself will attend upon every instance of the exercise of the energetic capacity. ("As the bloom of youth does on those in the flower of their youth," Aristotle says in the *Nicomachean Ethics*, 1174b 33). Since the end of a kinetic activity is not contained in every instance of its exercise but is that at which the activity aims, the subject's awareness of agreement of a *kinetic* deed with herself will attend upon that end's attainment and the activity's completion.

Of course, while exercising a capacity for a *kinetic* activity the subject may experience feelings. These can be either the pleasure or pain that attends the *energeia* of exercising a capacity in general, e.g., I can enjoy the experience of exercising a capacity I have well, as such, abstracting from the pleasure to be had in attaining the aim at which it aims. (While exercising a capacity for *kinetic* activity, I am energetically engaged qua bearer of the capacity itself. For example, a runner aiming at the finish line is engaged in a *kinetic* activity insofar as she is aiming

at the finish line in particular circumstances, but she is engaged in *energetic* activity insofar as she is being a runner in-activity). This pleasure would not stand in a relation to the particular end at which the *kinetic* activity aims. Alternatively, I may experience pleasure or pain that are grounded in the expectation of success or failure to attain the particular end at which the *kinetic* activity aims. Thus, although pleasure related to a particular end can be experienced while exercising a capacity aiming at that particular end, that pleasure is present in the course of the exercise of the kinetic capacity as an awareness of *expected* gratification.

This characterization of sensible pleasure is confirmed by Kant's rejection of sentimentalism to which we attended in Chapter Two. It moreover allows us to see the relation between sensible pleasure and the sensible *Triebfeder*:

However dissimilar representations of objects may be—they may be representations of the understanding or even of reason, in contrast to representations of sense—the *feeling of pleasure by which alone they properly constitute the determining ground of the will (the agreeableness, the gratification expected from the object, which impels activity to produce it) is nevertheless of one and the same kind not only insofar as it can always be cognized only empirically but also insofar as it affects one and the same vital force that is manifested in the faculty of desire, and in this respect can differ only in degree from any other determining ground. (KpV 5:22–3, emphasis mine).*

I have claimed above that to admit a *Triebfeder* into a maxim is to *determine a structure of motivation*, i.e., it is to determine certain kinds of ends as being *Triebfedern*. The two possible kinds of ends available to the agent to act on are *kinetic* and *energetic*. Pleasure and displeasure, insofar as they relate to desire, are the awareness of a representation being causally efficacious with respect to the agent. The pleasure in a kinetic act is the awareness of a representation of a kinetic end, i.e., one at which the action aims, an action that is not complete until that aim is achieved. Thus, the determining ground of the will in insofar as the agent is working towards a particular (*kinetic*) end will be, as Kant puts it here, “the gratification expected from the object,

which impels activity to produce it.” That is, the determining ground of the will is the expectation of consciousness of agreement in the accomplishment of the end. This is why Kant can say in this passage that sensible pleasure can be taken in representations that are not “of sense,” but belong to “the understanding” or “even of reason.” An activity is sensibly motivated as long as the logical structure of the activity the activity is *kinetic*, i.e., insofar as it is a “sensible deed.”

The pleasure in an energetic act is the awareness of a representation of an energetic end, i.e., one which the action accomplishes in every instance of its exercise. The performance of an *energetic* act is not performed in an expectation of gratification, because it contains its end in itself. This is why Kant is able to say, as we have seen in Chapter Two, that the feeling of moral respect does not follow from the determined will (from an action) like an effect, i.e., it is not such as to be, as we have seen him say of sensible pleasure, “cognized only empirically.” The feeling of moral respect can instead can be derived “*a priori*” (*KU* 5:221–2). This, Kant claimed, required stepping beyond “the bounds of experience and appealing to a causality that rests on a supersensible property of the subject: that of freedom.”

But even there we did not actually derive this feeling from the idea of the moral as a cause, rather it was merely the determination of the will that was derived from the latter. The state of mind of a will determined by something, however, is in itself already a feeling of pleasure and is identical with it, thus it does not follow from it as an effect. (*KU* 5:221–2).

The determination of a will, i.e., reason in its practical employment, is the determination not of a capacity for instrumental rationality, not just a capacity to act on hypothetical imperatives, but a capacity of reason “to determine the power of choice unconditionally” (*R* 6:26 fn67). To say that we can act not on the expectation of gratification but from the recognition of the authority of a

law of reason is the same as to say that being determined by the moral law is identical to a feeling of pleasure.

To be motivated by the moral law not by expecting gratification in the performance of a moral action—this, for Kant, is freedom. In other words, while the determination to *kinetic* activity is rational (one acts in consciousness of a hypothetical imperative, and therefore in an exercise of capacities of theoretical cognition), that activity is not, considered in itself, free. Whether or not it is free depends on whether or not it is grounded in an *energetic* activity that is free, i.e., that is performed from consciousness of the necessity of a universal law.

Notice however that the interpretation of *Triebfeder* with reference to the kind of ends that belong to rational activities, provides us with a sense in which *Triebfeder* can be subordinated to each other, namely, in terms of their ends. Thus, that a sensible *Triebfeder* can be subordinated to a moral one means that the pursuit of a particular end can be subordinated to activity that contains its end in itself, activity from the recognition of law. Thus both *kinetic* pleasure (or a sensible *Triebfeder*) can be identified in a morally worthy activity, without vitiating the moral worth of the action. Any kinetic activity, whether it instantiates an energetic activity performed in recognition of the authority of the moral law or not, will aim at an end, and thereby be determined, to that extent, by an expectation of awareness of agreement of end and self, the pleasure which will be attained when the end is accomplished.

This interpretation gains support from Kant's introductory remarks in the "Preface to the First Edition" of the *Religion*. Here, Kant casts the distinction between good and bad action as turning on the question of whether a particular purpose is required in order to cognize one's duty or observe it. The laws of pure practical reason, Kant writes, "bind through the bare form of

universal lawfulness of the maxims that are to be adopted in accordance with this form as the supreme (itself unconditioned) condition of all purposes” (R 6:3) and therefore “morality needs throughout no material determining basis of the free power of choice at all, i.e., no purpose [*Zweck*], neither for cognizing what one’s duty is, nor for impelling one to its performance” (R 6:4). In order to know whether I ought to do the right thing (Kant considers as examples giving true testimony and acting in good faith when property is put in one’s trust) “there is no need to inquire after a purpose [*Zweck*] that I might—in my explication—perhaps propose to myself to bring about” (R 6:4). The determination to act is indifferent to what it is that I will bring about and whether it serves any interest that I may have besides my recognition of the authority of the law.

However, Kant goes on to claim, although morality does not require a “presentation of a purpose which would have to precede the determination of the will,” it is not thereby the case that the good practical agent cannot act for purposes, in fact her actions will have a “necessary reference” to a purpose.

Namely not as the basis of the maxims adopted in conformity with those laws but as these maxims’ necessary consequences. —For without any reference to a purpose, no determination of the will can take place in the human being at all, because such determination cannot be without any effect; and the presentation of the effect, even if not as the determining basis of the power of choice and as a purpose that precedes in the intention, must yet be capable of being admitted as the consequence of that power’s determination to a purpose by the law (*finis in consequentiam veniens*). (R 6:4)

The power of choice must be instructed not only as to “how it is to operate” but also “toward what.” The purpose that a good will adds is an “objectively determined object,” an object that it “should have,” and it “emerges from morality,” simply insofar as a rational will “cannot possibly be indifferent [...] to how the answer to the question might turn out, *what it is that will result*

from this right action of ours” (R 6:5). The idea of this object is the idea “of all purposes as we ought to have them” and “everything that is harmonious therewith and is conditioned, in all those purposes that we do have.”

I take these statements to confirm the claims I have presented above. Namely, that moral goodness consists in never acting for a concrete end, but acting for a concrete purpose only insofar as it instantiates acting in recognition of the moral law. We may add ‘or in a way that harmonizes with it’ but I would emphasize (though I cannot argue for it here) that to harmonize with a set of given purposes means not only to act in a way that does not conflict, i.e., obviously hinder, those purposes, but in a way that promotes that set of purposes.

§7. Ordering *Triebfedern*: Subordinating Sensible Activity (*Kinesis*) to Intelligible Activity (*Energieia*)

According to the above, the constitution of a structure of motivation, one’s moral *Gesinnung*, consists in acting from the consciousness of the moral law. To act from the moral law is for that subject’s reason to have the ability “to determine the power of choice unconditionally through the mere presentation of the qualification of its maxims for universal legislation” (R 6:26 fn67).

In the constituting maxim of a morally good *Gesinnung* one subordinates the performance of actions whose *Triebfeder* is sensible to the performance of actions whose *Triebfeder* is moral, i.e., one subordinates kinetic activities to energetic ones, and in particular to the energetic activity that is performed from consciousness of the time-general moral law. This means that one subordinates the performance of actions considered as sensible deeds, actions

considered as having their ends in “*objects* of the power of choice” (R 6:31, emphasis mine) to the performance of intelligible deeds, actions considered as having their ends within themselves, whose principles are not the desire for a particular end, but a law, namely the moral law.

*

I have hereby offered an account of the constitution of a morally good character in Kant. However, no account of moral character can be complete without attending to three apparent ways in which a character may fall short of being good: (1) it can be evil; (2) it can be indeterminate, as in the case of a child acquiring a moral character; (3) it can be intermediate, a notion apparently necessary to account for the kind of morally significant differences which are not exhausted by a division into good and evil character, most importantly, the kind of morally significant differences exhibited in the process of moral progress.

To address (1), I will attend to the *Gesinnung*-constituting maxim of a morally evil agent and address the question of the sense in which a person with evil character can be said to be acting on a *Gesinnung*-constituting maxim at all. There I will offer an interpretation of Kant’s account of the three grades of evil. In the course of that interpretation I will elucidate Kant’s idea of the acquisition of moral character—and thereby the idea of *immaturity* of character. This account will render intelligible both the possibility of mixed motivations and the evaluation of moral progress.

Chapter 4

Moral Education and Moral Evil

“Can you tell me, Socrates—can virtue be taught, or is it rather to be acquired by practice? Or is it neither to be practiced nor to be learned but something that comes to men by nature or in some other way?”

—Plato, *Meno*

§1. Introduction: The Apparent Conflict between Rigorism and the Levels [*Stufen*] of Evil

In Chapter Three, I began elucidating Kant’s account of moral character. Moral character is a *Gesinnung*, a freely determined moral attitude. This moral attitude is identified by a maxim (or principle of activity) of utmost generality, which serves as the fundamental rule for the exercise of the free power of choice. It consists in the admission of *Triebfedern*, kinds of grounds of desire, into the maxim of utmost generality. In Chapter Three I grappled with the difficulty of accounting for the way in which this moral attitude is *freely* determined. Kant claims that unlike ordinary free actions, which are “sensible deeds”, the determination of the moral attitude constitutive of character, takes place in an “intelligible deed.” I argued that we should consider interpreting the distinction between intelligible and sensible deeds against the Aristotelian distinction between kinetic and energetic activities, activities that aim at the attainment of an end and activities that contain their end in themselves. I further argued that to admit a *Triebfeder*, a ground of desire, into one’s character-constituting maxim is to determine a structure of motivation. The faculty of desire is one’s capacity to act from the representation of ends. Accordingly to determine grounds for one’s faculty of desire is to determine certain kinds of ends to be grounds of activity.

I further proposed that Kant's distinction between sensible and intelligible deeds, or activities, should guide our interpretation of the two kinds of *Triebfedern*, grounds of desire, that are admitted into the character-constituting maxim: sensible and moral. To act on a sensible ground is to act for a *kinetic* end, an end at which the action aims and that is accomplished when the action is completed. To act on a moral ground is to act for an *energetic* end, an end accomplished in every instance of the exercise of the capacity for *energetic* activity. Specifically, it is to act in recognition of the moral law. Finally I suggested a reading of Kant's claim that in the *Gesinnung*-constituting maxim the two kinds of *Triebfedern* are subordinated to one another, as a claim concerning the constitution of the structure of motivation whereby the pursuit of one kind of end—sensible-*kinetic* or moral-*energetic*—is subordinated to the pursuit of the other.

Good and bad characters reflect which way these grounds are subordinated to one another. That is, good and bad characters reflect which kind of ground is made supreme—which kind of end is made the condition of pursuing of the other. Moral goodness consists in respect for the moral law being admitted “as an incentive [*Triebfeder*], sufficient by itself, of the power of choice” (*R* 6:27). Moral evil consists in the sensible *Triebfedern* being admitted into the *Gesinnung*-constituting maxim “as by themselves sufficient for determining the power of choice” (*R* 6:36). I suggested that moral goodness should therefore be understood as activity grounded in a motivational structure whereby the pursuit of particular, sensible-*kinetic* ends is subordinated to the pursuit of the moral-*energetic* end, namely, the activity whose principle is the moral law. Moral evil should be understood, on the contrary, as activity grounded in giving, within one's character-constituting maxim, the sensible *Triebfedern*, i.e., the pursuit of *kinetic*, particular ends, *absolute* priority in action.

I mentioned briefly in Chapter One that Kant takes his account of the *Gesinnung* to justify a commitment to “rigorism,” the claim that “*The human being is [...] either morally good or morally evil*” (R 6:22), i.e., that the human being is not “both simultaneously, namely in some points good, in others evil” (*ibid.*) Kant holds that since the power of choice can be determined to an action by a certain kind of *Triebfeder* only insofar as the subject has admitted this *Triebfeder* into her *Gesinnung*-constituting maxim as sufficient for the determination of one’s will, if someone acts such that the law does not determine her power of choice, then a non-moral *Triebfeder* must have influenced her power of choice, i.e., must have been admitted as a *Triebfeder*, a ground of desire, into her *Gesinnung*-constituting maxim. He also holds that morally evaluating any action requires attending to its maxim, its principle, and, further, that to evaluate a particular principle of choice requires evaluating the fundamental principle of choice in which all of a subject’s particular principles of choice are grounded, i.e. “the subjective basis for the adoption of maxims,” that the individual has adopted in a free act. This subjective basis for the adoption of maxims is identified with the *Gesinnung*. The *Gesinnung*, Kant claims, “can only be one, and it applies universally to the entire use of freedom” (R 6:25). Since both *Triebfedern* are admitted and must be ordered, one way or another, Kant concludes that the “attitude [*Gesinnung*] in regard to the moral law is never indifferent (never neither of the two, neither good nor evil)” (R 6:24). Since any morally evaluable action implies a *Gesinnung* and a *Gesinnung* is determined with regard to the moral law, a morally responsible agent is always either good or evil.

Moreover, the *same* agent is not only at any moment either good or evil, but, Kant claims, the individual cannot be at some points good and at some points evil:

For if he is good in one point, then he has admitted the moral law into his maxim; thus if in another point he were to be simultaneously evil, then, because the moral law of compliance with duty as such is only one and is universal, the maxim referred to it would be universal, while simultaneously being only a particular maxim—which is contradictory. (*R* 6:24)

Goodness, the goodness of a particular action, Kant claims, must be understood as grounded in the recognition of the absolute necessity and therefore *universality* of the moral law. To recognize the necessity and universality of the moral law is to admit the moral *Triebfeder* into one's most general maxim of choice, and thereby to be motivated by the moral law. A recognition of the necessity and universality of the law is not consistent with the occasional admission of the sensible *Triebfeder* as ground of motivation. If one gives priority to the pursuit of a particular end, a purpose, in one's determination to action, one cannot be said to have adopted the good *Gesinnung*, for the latter would preclude the prioritization of a sensible *Triebfeder*. Since one is good if and only if one has recognized the necessary and universal validity of the law, one cannot be good at one time and evil at another.

However, it may seem that this account of good and evil, as stable states of character grounded in a single overarching maxim—a single, stable structure of motivation—leaves no room for Kant's own account of the levels (*Stufen*) of evil: frailty, impurity, and wickedness. If the levels of evil are meant to be species of the genus of evil, and evil *generally* consists in the reversal of the *Triebfedern* of the power of choice, every one of its species ought to be a species of this general perverse reversal. In other words, if evil character is the stable perverse ordering of *Triebfedern*, each of the levels of evil should instantiate this structure and offer a further determination of it. Frailty, impurity, and wickedness ought all to be grounded in the perverse ordering of *Triebfedern*; the difference in the levels should mark a differentia, a further

determination of the general structure. But Kant's account of the levels of evil, which appears but a few pages after his brief defense of rigorism, defies this expectation.

Firstly, the definition of wickedness, the third and highest level of evil, is *identical* to the general definition of evil (wickedness "reverses the moral order in regard to the incentives [*Triebfedern*] of *free* power of choice" (R 6:30). If wickedness is a species of the genus, why characterize it in terms of the general definition of evil while refusing that characterization of the two former levels? Secondly, the former two levels of evil are not only not characterized in terms of the general definition but appear to conflict with it.

Frailty (Kantian "weakness of will") is not a stable disposition. Kant introduces it as "the heart's weakness in complying with adopted maxims," (R 6:29) and defines it as follows:

I admit the good (the law) into the maxim of my power of choice; but this good, which objectively, in the idea (*in thesi*) is an insurmountable incentive [*Triebfeder*], is subjectively (*in hypothesi*) the weaker (by comparison with inclination) when the maxim ought to be complied with. (R 6:29)

This characterization seems to contradict Kant's general account of evil. According to the previous argument, if one has admitted the good into one's *Gesinnung*-constituting maxim one cannot consistently be said to fail to act from it. Or, conversely, if one has acted on a sensible *Triebfeder*, that sensible *Triebfeder* must be taken as itself to have been admitted as sufficient for the determination of action in the *Gesinnung*-constituting maxim. A sensible *Triebfeder* cannot, of itself, overcome the determination to act in recognition of the authority of the moral law or else the action would be no longer imputable to the agent.

Impurity, a state where the performance of good action requires the cooperation of sensible *Triebfedern* but does not simply prioritize them, likewise fits awkwardly with the general definition of evil. Kant introduces it as "the propensity to mix immoral incentives

[*Triebfedern*] with the moral ones (even if this were done with good intention and under maxims of the good)” (R 6:29), and defines it as follows:

although in terms of its object (the intended compliance with the law) the maxim is indeed good, and perhaps even powerful enough for performance, it is not purely moral, i.e., it has not, as should be the case, admitted the law *alone* into itself as *sufficient* incentive [*Triebfeder*], but usually (perhaps always) still needs other incentives [*Triebfedern*] besides that law in order thereby to determine the power of choice to what duty demands. In other words, that impurity consists in this: that actions conforming to duty are not done purely from duty. (R 6:30)

That an action conforming to duty is not done from duty is no problem for Kant’s general account of evil, for an agent of bad character may accidentally do that which the law commands when it happens to coincide with the concrete ends (purposes) that the agent happens to find herself attached to. Cases of this sort would easily fall under the general definition of evil, and therefore wickedness, the third level of evil. However, as Kant presents impurity, the performance of the duty-conforming action is not accidental. The order of *Triebfedern* is not simply perversely reversed, the *Triebfedern* are instead “mixed.” Compliance with the law is “intended”; the maxims are maxims “of the good.” This characterization recalls Kant’s description in the Preface to the First Edition, which I have mentioned in Chapter Three, of a person who “although his avowal is legitimately being demanded, still finds it necessary to look around for some purpose” that he may bring about in doing that which he knows he ought (R 6:4). He seems to have something similar in mind when, in the second part of the *Critique of Practical Reason*, the “Doctrine of the method of pure practical reason,” Kant describes a moral failure of the following sort: “When we bring any flattering thought of merit into our action, then the incentive [*Triebfeder*] is already somewhat mixed with self-love and thus has some assistance from the side of sensibility” (*KpV* 5:159). In the description of impurity as well as his

brief discussions of examples, Kant appears to have in mind an action that is not performed simply from a non-moral interest. Instead, the descriptions suggest that had the action not conformed to duty, the agent would not have performed it on the basis those sensible *Triebfedern* she in fact looked for upon recognizing it as her duty. The agent knows what the right thing to do is, and she knows it not accidentally but from consciousness of the moral law. However, she requires sensible *Triebfedern* in order to prop herself up. She compensates herself for the sacrifice of some particular ends (purposes) by identifying other particular ends that her duty-conforming action may promote.

Actions grounded both in frailty and impurity seem not to be cases of action stemming from evil character. How then are they cases of evil at all? It seems that one of the two accounts—that of evil as grounded in the single *Gesinnung*-constituting maxim or that of the first two levels of evil—must be jettisoned.

Dismissing the account of the *Gesinnung*-constituting maxim would leave Kant with no resources to explain how evil actions determined by a non-moral *Triebfeder* are imputable to the subject, and therefore with no resources to explain how can they be morally evaluable at all. In particular, admitting frailty and impurity appears to require giving up on the idea that moral goodness and moral evil are fundamentally grounded in a stable attitude toward the moral law, the *Gesinnung*, constituted by a single, motivation-structure determining maxim. Without an account of how we may be understood to be the originators of our structure of motivation (how is our moral character imputable to us), any failure in our motivation appears to be beyond the reach of our free, rational capacity to determine ourselves.

Dismissing frailty and impurity would likewise come at a high philosophical cost. First, frailty and impurity apparently constitute fundamental phenomena of moral life: doing something wrong with some apparent awareness that it is the wrong thing to do and doing the right thing with some apparent awareness that it is anyway in one's interest to do it. Second, frailty and impurity are apparently necessary in order to make sense of the distinctions between characters that are not wholly good, but that are not wholly and stably bad, either. To deny the possibility of evil actions as occasional deviations from the good would commit one to the claim that, for example, if a person has ever acted badly, even once, it must be the case that they are in fact wholly committed to the satisfaction of their particular purposes, that they prioritize these above all else, and that their otherwise consistent performance of actions in accordance with duty—which for all they can tell are performed in the recognition of duty as such—are in fact all merely a coincidental consequence of their committed pursuit of particular ends.

Moreover, even if one is willing to grant that goodness as a state of perfection is singular—and thus that any deviations from the good constitute failure to have a good character—still, denying any gradation in states of imperfection renders it impossible to account for moral differences between moral subjects and, in particular, to recognize the kind of moral differences necessary to account for the possibility of moral development.

It seems therefore that Kant's rigorism and his account of the levels of evil are in irresolvable tension with one another. Retaining the idea of a *Gesinnung* as the fundamental ground of action and moral worth as well as the notions of frailty and impurity appears to require supplementing Kant's account with more complex *Gesinnung* constitutions. Kant's argument for rigorism, however, shows that since the moral law is necessary and universal, the idea of

admitting the moral law as a determining ground of the will but “with exceptions” is not coherent. In what follows I will argue that we should understand frailty and impurity neither as alternative states of moral *Gesinnung*, nor as inexplicable exceptions to the account. Instead, I propose thinking of them as cases of moral *immaturity*, a condition that *precedes* the constitution of a stable moral *Gesinnung*.

The challenge to proposal is twofold: first, since an immature person does not have a moral *Gesinnung*, and since Kant argues that an action’s moral worth is grounded in the moral valance of her *Gesinnung*, it may seem that actions grounded in states of moral immaturity cannot be subject to moral evaluation. I will argue that a state of immaturity, although not yet a moral *Gesinnung*, is essentially linked to a *Gesinnung* in virtue of being a partial actualization of the practical-moral capacity: immaturity is a moment on the path to acquiring a *Gesinnung*. We can therefore morally evaluate subjects in a state of immaturity in accordance with the direction of their moral development: whether they are moving towards or away from virtue (a disjunction which, for Kant, is exclusive).

A second and more serious challenge is that Kant is typically taken to reject out of hand the idea of acquiring rational capacities in general and moral capacities in particular. I argue that, on the contrary, Kant provides the resources through which we can comprehend the acquisition of rational capacities and moral character; the latter essentially involving the cultivation of a receptive capacity (feeling) (§3).

§2. Is a *Gesinnung* Acquired? A Puzzle.

How it is that we come to be the bearers of a *Gesinnung* is a question that receives sporadic and apparently conflicting treatment in Kant. Several lines of argument in the *Religion*

appear to decisively preclude the possibility of the gradual acquisition of a *Gesinnung*, while passages from the the same text, as well as from the *Critique of Practical Reason*, seem explicitly to presuppose just such an account.

In our preliminary treatment of the *Gesinnung* in Chapter Three, I examined Kant's characterization of the act by which an agent is "the originator" of her *Gesinnung*, as one that is not in time, but did so abstractly. I focused on Kant's claim that the constitution of a *Gesinnung* is an "intelligible deed, cognizable by bare reason without any time condition," in contrast with the sensible deed, which is "empirical, given in time (*factum phaenomenon*)" (R 6:31). I attempted to dispel some of the worries concerning this characterization by drawing attention to the logical form of activity I referred to, following Aristotle, as *energetic* activity, and in particular *energetic* activity performed in the recognition of the authority of the law. I suggested that we should consider the logical form of *energetic* activity as an interpretative model through which to understand Kant's *Gesinnung*-constituting intelligible deed.

However, Kant's own characterization of the way in which the *Gesinnung* constituting act is not subject to time conditions goes beyond the abstract characterization I offered. In particular, it seems to hinder any intelligible interpretation of the idea of acquiring a moral *Gesinnung*. The act by which a person comes to adopt a *Gesinnung*-constituting maxim is not only not an act in time, "we cannot derive this attitude [*Gesinnung*], or rather its supreme basis, from any first act of the power of choice in time [*irgend einem ersten Zeit-Actus der Willkür*]" (R 6:25), but "it has not been procured over time (that one or the other he *has always been, from his youth*)" (R 6:25). Elsewhere Kant goes as far as to claim that the *Gesinnung* "is laid at the basis (in earliest youth, back at the point of birth) prior to any use of freedom that is given in

experience, and thus is conceived as present in the human being simultaneously with birth—though not exactly as having birth as its cause” (R 6:21–2).

Kant’s claim that a person’s *Gesinnung* is laid in earliest youth, and is conceived as present in the human being simultaneously with birth, seem to leave no room for the idea of acquiring a moral *Gesinnung*, in particular by leaving no room for any *moral* state that is not a good or evil *Gesinnung* that might precede it.

Kant likewise seems hostile to the idea of morally viable states other than a *Gesinnung* in another set of claims he makes while discussing the idea of moral improvement:

However, that someone should become a human being who is not merely *legally* but *morally* good [...], i.e., virtuous in terms of intelligible character (*virtus noumenon*)—who, when he cognizes something as a duty, requires no other incentive [*Triebfeder*] beyond this presentation of duty itself—this cannot be brought about, so long as the foundation of the maxims remains impure, through gradual *reform*, but must be brought about through a *revolution* in the attitude in the human being (a transition to the maxim of the attitude’s holiness); and he can become a new human being only through a kind of rebirth, as if through a new creation [...] and a change of heart. (R 6:47)

Here Kant seems committed to the claim that no gradual reform in moral character is possible at all, only an about-face turn to the good.

However, Kant likewise claims in the *Religion* that character must be acquired. He makes the claim in the central moment where he introduces the three elements of the human being’s essential determination that we encountered in Chapter Three. To recap, Kant claims that personality, the constituent of “human nature” by virtue of which reason can determine the power of choice unconditionally, i.e., in whom reason is practical, is “the receptivity to respect for the moral law, as an incentive [*Triebfeder*], sufficient by itself, of the power of choice” (R 6:27).

This receptivity to mere respect for the moral law within us [*die Empfänglichkeit der Achtung für das moralische Gesetz*] would be the moral feeling, which by itself does not yet amount to a purpose of the natural predisposition, but amounts to such a purpose only insofar as it is an incentive [*Triebfeder*] of the power of choice. [...] [T]his becomes possible solely through the free power of choice's admitting the moral feeling into its maxim, the constitution of such a power of choice is a good character. (R 6:27)

He adds,

Such a character, as in general every character of the free power of choice, is something that can only be *acquired*, but for the possibility of which there must nonetheless be present in our nature a predisposition on which absolutely nothing evil can be grafted [*so ist Beschaffenheit einer solchen Willkür der gute Charakter, welcher wie überhaupt jeder Charakter der freien Willkür etwas ist, das nur erworben werden kann*]. (R 6:27, emphasis mine)

Here Kant claims somewhat opaquely that the character of the power of choice, what he later calls the *Gesinnung*, is grounded in the feeling of moral respect—the capacity to take pleasure in one's performing actions from consciousness of the law—becoming the ground of one's action. He likewise claims that character, i.e., the capacity for moral feeling as the ground of desire, is acquired. What is involved in the acquisition of a character in which the capacity to feel respect, i.e., recognize the law as a law for me, is actualized?

This claim is not unique to the *Religion*. The *Critique of Practical Reason* concludes with a short second part, “Doctrine of the method of pure practical reason,” dedicated to the treatment of the question how one, i.e., the moral teacher, can provide “the laws of pure practical reason with access to the human mind and *influence* on its maxim, that is, the way in which one can make objectively practical reason *subjectively* practical as well” (*KpV* 5:151). Kant explicitly claims that only the “feeling” of the “pure moral motive,” “can ground a moral character” (*KpV* 5:151-152).

Prefiguring his definition of personality, Kant claims that in order to bring forth a “moral disposition” [*Gesinnung*] the constitution of “human nature” must be such that a human being can to prefer the law “from pure respect for it.” In the few pages that follow Kant proceeds to propose a method by which someone in the position of a moral instructor could help bring about the constitution of a moral character through a process of actualizing receptivity to respect:

It certainly cannot be denied that in order to bring either a mind that is still uncultivated or one that is degraded onto the track of the morally good in the first place, some preparatory guidance is needed to attract it by means of its own advantage or to alarm it by fear of harm; but as soon as this machinery, these leading strings have had even some effect, the pure moral motive [*der reine moralische Bewegungsgrund*] must be brought to bear on the soul, the motive which—not only because it is the only one that can ground a character [*Charakter*] (a consistent practical cast of mind in accordance with unchangeable maxims) but also because it teaches the human being to feel his own dignity—gives his mind power, unexpected even by himself, to tear himself away from all sensible attachments so far as they want to rule over him and to find a rich compensation for the sacrifice he makes in the independence of his rational nature and the greatness of soul to which he sees that he is called. We will therefore show, by observations anyone can make, that this property of our minds, this receptivity to a pure moral interest [*Empfänglichkeit eines reinen moralischen Interesse*] and hence the moving force of the pure representation of virtue [*die bewegende Kraft der reinen Vorstellung der Tugend*], when it is duly brought to bear on the human heart is the only incentive to the good [*einzigste Triebfeder zum Guten*] and the only one when an enduring and meticulous observance of moral maxims is in question. (*KpV* 5:152, translation modified)

Note the parallel invocation of receptivity [*Empfänglichkeit*] to a pure moral interest [*Empfänglichkeit eines reinen moralischen Interesse*] in the second *Critique*, and receptivity to mere respect for the moral law within us [*die Empfänglichkeit der Achtung für das moralische Gesetz*] in the *Religion*. In the second *Critique*, Kant says that *when* this receptivity “is brought to bear on the human heart” it becomes the “*Triebfeder* to the good,” and likewise in the *Religion* he claims this receptivity to respect is not yet the full accomplishment of the end associated with the determination to personality, an end that is accomplished only when the receptivity to respect

is admitted as the *Triebfeder* of the power of choice. In other words, Kant is consistent in his claim that the full actualization of practical reason in us consists in the receptivity to the feeling of respect for the moral law that comes to serve as the ground of our activity. Drawing on the results of Chapter Two, we may say that the full actualization of practical reason consists in the actualization of our capacity to be consciousness of ourselves as rational, free, and efficacious.

Moreover, in both the *Critique* (*KpV* 5:152–61) and the *Religion* (*R* 6:48–53) Kant provides similar accounts of “moral education” [*die moralische Bildung des Menschen*] (*R* 6:48). In the second *Critique* Kant contemplates what sort of presentation and treatment of examples of moral conduct could promote the “consciousness of the law as also an incentive [*Triebfeder*]” (*KpV* 5:159). He notes the propensity of people to evaluate the moral motives and worth of other agents, and suggests that here lies the key to the promotion of a consciousness of one’s moral being. The student gradually comes to “the consciousness of his freedom,” whereby he overcomes the “initial feeling of pain” which attends upon the renunciation of actions where the incentives [*Triebfedern*] of inclination have “any influence on it as determining ground” (*KpV* 5:160) and thereby “his mind is made receptive [*empfänglich*] to the feeling of satisfaction from other resources” [*das Gemüth für die Empfindung der Zufriedenheit aus anderen Quellen empfänglich gemacht wird*] (*KpV* 5:161).

In the *Religion* Kant speaks again of children’s predisposition to discover spurious incentives [*Triebfedern*] mixed in with the performance of actions that conform to duty, in which case “the action instantly loses all moral worth for them” (*R* 6:48). This is an appearance of their predisposition to the good [*Anlage zum Guten*], i.e., a manifestation of their personality, which however must be “cultivated,” once again through the presentation of examples of actions that

conform to the law and by inviting “moral apprentices” [*moralischen Lehrlinge*] to discern possible impurities in the agents’ maxims. This predisposition to the good gradually [*allmählig*] passes over into the apprentices’ way of thinking [*Denkungsart*], “so that duty just by itself starts to acquire in their hearts a noticeable weight” (R 6:48). The “repeated arousing of this feeling of the sublimity of one’s moral vocation” is a means for awakening the moral disposition [*Mittel der Erweckung sittlicher Gesinnungen*] and ought therefore be encouraged (R 6:50).

In both accounts, therefore, moral education, the actualization of the predisposition to the good that is personality (practical reason), requires the *gradual* cultivation of the moral apprentice’s capacity to feel respect for the moral law, i.e., to act in consciousness of her freedom, as a means for the constitution of a moral character [*Gesinnung*].

§3. *Denkungsart* and *Sinnesart*

How can we reconcile Kant’s claim that the moral *Gesinnung* is not procured over time but is laid in earliest youth—present in the human being simultaneously with birth—with his descriptions of the gradual cultivation of a capacity for moral feeling that is apparently necessary for the *Gesinnung*? To do so we must attend to Kant’s distinction between *Denkungsart* and *Sinnesart*, the way of thinking and the way of sensing. Kant qualifies his requirement of a revolution in the way of thinking with a gradual reform in the way of sensing in the following terms:

But if the human being is corrupted at the basis of his maxims, how is it possible for him to achieve this revolution through his own powers and to become on his own a good human being? And yet duty commands us to be good, and it commands us nothing but what we can do. This cannot be reconciled except by maintaining that the revolution is necessary for the way of thinking [*Denkungsart*], but the gradual reform for the way of sensing [*Sinnesart*] (which

opposes the former way with obstacles); and hence [each] must also be possible for the human being. (R 6:47)

He introduces the morally relevant notion of a *Denkungsart* inconspicuously, when, in characterizing the third level of evil, wickedness, which “reverses the moral order in regard to the incentives [*Triebfedern*] of the free power of choice,” Kant notes that although wickedness is compatible with the performance of legally good actions, “yet the way of thinking is thereby corrupted in its root (as far as the moral attitude is concerned)” (R 6:30). Shortly after, Kant adds that whatever is not done from duty, i.e., whenever the law is not by itself sufficient as incentive [*Triebfeder*], is sin “according to the way of thinking” (R 6:30). The multiple other uses of “*Dekunsgart*” are consistent with this.¹ The *Denkungsart* therefore appears to be simply identical to intelligible *Gesinnung*.²

¹ However, Kant does not use *Denkungsart*, which appears close to two dozen times in the *Religion*, exclusively as a term of art. For example, he first uses “*Denkungsart*” to characterize a general philosophical commitment, introducing the notion of rigorism thus, “Those who are attached to this strict way of thinking [*Denkungsart*] are commonly called *rigorists*” (R 6:22).

² G. F. Munzel argues that the notion of character in Kant is exhausted by his account of the moral *Denkungsart*. Her critical analysis of interpretative attempts to account for moral character in Kant is insightful. She identifies the two central interpretative positions. First, an interpretation that identifies character with the will. This renders the notion of character entirely idle, and leaves Kant with nothing to offer regarding the notion of character as it is traditionally understood, i.e., as grounded in habits of acting and feeling. The second interpretation holds that there is an Aristotelian-style virtue ethics defense of moral character in Kant. This interpretative strategy locates in Kant’s writing evidence for a concern with “empirical character habituation” that straps a theory of virtue onto Kant’s deontology. For example, Barbara Herman claims that a case for moral development and character can be understood in a Kantian sense, but must be “added” to the account Kant has given us. Herman, “Making Room for Character,” in *Aristotle, Kant and the Stoics: Rethinking Happiness and Duty*, eds. Stephen Engstrom and Jennifer Whiting (Cambridge: Cambridge University Press, 1996), 37, 40, 44. Nancy Sherman likewise claims that in Kant’s account “we are to develop our talents and emotional capacities as part of virtue (and so conceive of virtue along the ancient model of an empirical project of character habituation)” and adds that “responsiveness to morality, as rooted in the rational nature of persons, flourishes best in someone who has cultivated emotional capacities.” Sherman, *Making a Necessity of Virtue*, 143–144. Munzel rightly rejects the idea that “empirical character habituation” is adequate to Kant’s notion of moral (intelligible) character. She goes on to develop an alternative “cognitive” conception of character, grounded in the notion of *Denkungsart*. However, Munzel, who like most Kant interpreters, takes practical sensibility to be limited to the idea of a purely non-rational capacity for feeling, can find no room for moral feeling in

Sinnesart, the way of sensing, appears in the *Religion* only once, in the passage I quoted last, where Kant introduces the notion of a revolution in one's moral attitude. In the quoted passage as well as those following, the revolution in the *Denkungsart* is contrasted with "gradual reform." In particular, as we have seen, Kant claims that one cannot become a morally good human being "so long as the foundation of the maxims remains impure, through gradual *reform*." This sort of "gradual reform" is coextensive with Kant's notion of phenomenal virtue: permanent performance of "lawful actions," which he says is "acquired little by little, and means to some a long habituation (in observing the law)," and which involves "gradual reform of his conduct and stabilization of his maxims" (*R* 6:47). This latter sort of change, Kant says, requires not a "change of heart, but only a change of mores" (*ibid.*) It is therefore tempting to dismiss talk of the necessary "gradual reform for the way of sensing [*Sinnesart*]" which accompanies the revolution in the way of thinking as identical to the gradual reform of conduct that goes with the long habituation in performing lawful actions. Attending to what immediately follows suggests that this conclusion would be hasty. For when Kant claims, that for moral reform there needs to be a revolution in the way of thinking and a gradual reform in the way of sensing, the gradual reform in the *Sinnesart* is not identical to a gradual reform in mores.

Kant glosses the duality of revolution and gradual progress in the following way. First, Kant prioritizes the revolution in the *Denkungsart* as a precondition of the constitution of good character:

That is, if through a single immutable decision the human being reverses the supreme basis of his maxims whereby he was an evil human being (and thereby puts on a new human being), then he is, to this extent, in terms of the principle and the way of thinking, a subject receptive to the good [*so ist er so fern dem*

the constitution of moral character. Her account is therefore vulnerable to the form of critique I levied against the intellectualists in Chapter Two. (See Munzel, *Kant's Conception of Moral Character*.)

Princip und der Denkungsart nach ein fürs Gute empfängliches Subject]. (R 6:47–8)

A so-called revolution in the *Denkungsart*, in the way of thinking, is necessary for the constitution of a good human being. In such a revolution a subject renders herself “a subject receptive to the good.” However, the *Gute empfängliches Subject* is not yet the good human being. As in the definition of personality, “moral feeling” considered “by itself,” i.e., in abstraction, does not yet amount to the purpose of the predisposition to personality, so here, the adoption of a principle, by itself, i.e., in abstraction, does not yet amount to moral goodness.

Kant continues,

but he is a good human being only in continual acting and becoming, i.e., he can hope that with such purity in the principle that he has adopted as the supreme maxim of his power of choice; and with the stability of this principle, he finds himself on the good (though narrow) path of constant *progress* from the bad to the better. (R 6:48)

Not every “gradual reform” pertains merely to the legality of actions. Gradual reform can take place in two *distinctly different* conditions: “as the foundation of the maxims remains impure,” or with “purity in the principle,” where the subject “in terms of the principle and the way of thinking,” is “a subject receptive to the good.” This description of course recalls Kant’s description of the moral apprentice whose mind “is made receptive [*empfänglich*] to the feeling of satisfaction from other resources” [*das Gemüth für die Empfindung der Zufriedenheit aus anderen Quellen empfänglich gemacht wird*] (KpV 5:161).

The gradual reform in *Sinnesart* therefore follows upon a revolution in *Denkungsart*. The good *Denkungsart* renders the subject receptive to the good. But the arousal of the moral feeling, is the cultivation of the predisposition to the good, rendering the subject receptive to the good in actuality.

The *Denkungsart* is therefore a necessary but not sufficient condition for the actualization of the predisposition to personality. Full actualization requires the further gradual reform of a *Sinnesart*. This may sound like a reversal of what is customarily taken to be the Aristotelian doctrine of moral education, whereby the acquisition of the virtues of character, as virtues of feeling, is prior to the acquisition of intellectual virtues, and in particular, prior to the acquisition of practical wisdom. While a possessor of virtues of character acts well in the conviction that an action is good to perform, it is only the possessor of practical wisdom who acts well in knowledge of why it is good to perform it. I cannot explore in detail the relation between Aristotle's and Kant's conception of moral virtue, but I would like to suggest that the priority of the constitution of the *Denkungsart* over the gradual reform in the *Sinnesart* (whereby the latter is a condition of the former) should not be understood as necessarily marking a disagreement in doctrine. In particular, I suggest that the constitution of the *Denkungsart* is not a state of knowledge of the good, i.e., knowledge of what courses of action are good, but is instead an attitude *toward the good*. To be oriented toward the good is a precondition for the acquisition of desires to perform actions that are good, as well as a precondition for the acquisition of the knowledge of *why* good actions are good.

A revolution in the *Denkungsart* marks, I propose, the idea that in order to explain moral progress as genuinely moral progress one must conceive of the moral apprentice as oriented toward the good, i.e., self-conscious of their position as a moral apprentice, one who is being guided in their acquisition of a good moral character. Without such a consciousness, a moral apprentice, although they may acquire habits of acting in a certain way, would not be learning to act well. Consider a case of the latter sort, where a child is subject to habituation grounded not in

self-conscious acquisition of a good character, but is instead habituated to act in ways that conform to the moral law merely by appeal to their sensitivity to promises of external reward and punishment (reward and punishment that is not intrinsically linked to the performance of the relevant kind of action). For example, suppose a child can be made to believe that a supernatural being has unlimited observational access to their activity. Someone may appeal to the existence of this supernatural being in order to discourage the child from lying. “Who broke the vase?” a caretaker may ask the child and add, “it would be futile to lie since the supernatural being, say, St. Nick, knows the truth, and will punish you if you do not disclose it yourself.” Having told the truth the child may be rewarded with praise, for example, the caretaker may say, “St. Nick will be happy not to have to punish you, as he is very busy.” The child does not learn thereby to take pleasure in doing the right thing because it is the right thing, but she may develop an emotional aversion to lying, the prospect of which may come to evoke (irrational) discomfort or fear even after she learns the truth about St. Nick. Note that supposing a child has acquired all predilections and behaviors by means of this mechanism of habituation, the child’s capacity to take pleasure in the good for its own sake will remain to that degree underdeveloped.

For praise and blame to have moral effect, i.e., for them to promote the subject’s capacity to take pleasure in actions performed for their own sake, i.e., because they are good, the child must be made aware of the project of moral education and their role as moral apprentice in it. To that degree the child participates in that process, which is a process of moral self-determination. In the child, good moral character is present only as a *Denkungsart* that has not yet been actualized in the gradual development of a *Sinnesart*. The *Denkungsart* has been actualized in the moral teacher’s *Sinnesart* (manifest, for one thing, in her engagement in the activity of moral

education).

§4. Habit, Free and Unfree

Before proceeding we must attend to a broader challenge to the interpretation I am proposing, namely, Kant's critique of habit and its acquisition (habituation) as insufficient for the acquisition of rational and, in particular, moral capacities. Addressing this challenge will help us to determine more precisely the character of the acquisition of moral capacities.

Kant condemns habit in multiple places. In the *Anthropology* Kant characterizes habit [*Angewohnheit*] as "a physical inner necessitation to proceed in the same manner that one has proceeded until now." He proceeds to warn, "It deprives even good actions of their moral worth because it impairs the freedom of the mind and, moreover, leads to thoughtless repetition of the very same act (monotony)" (A 7:149). In handwritten reflections on the *Anthropology*, Kant adds that "habit is never, not even in the case of good actions, to be fully embraced. [...] Even the good stops being virtuous through being habitual" (*Refl* 15:98f.).

Habit's threat is dual. First, it poses a threat to our freedom:

The reason why the habits of another stimulate the arousal of disgust in us is that here the animal in the human being jumps out far too much, and that here one is led instinctively by the rule of habituation, exactly like another (non-human) nature, and so runs the risk of falling into one and the same class with the beast." (A 7:149)

Kant emphasizes that in this sense we cannot explain virtue as "skill [*Fertigkeit*] in free lawful actions, for then it would be a mere mechanism of applying power. Virtue is moral strength in adherence to one's duty, which never should become habit but should always emerge entirely new and original from one's way of thinking" (A 7:147).

Second, habit not only threatens our freedom, but is an unreliable guide for the accomplishment of our purposes. Unless virtue “results from considered, firm, and continually purified principles,” it is “like any other mechanism of technically practical reason, [...] neither armed for all situations nor adequately secured against the changes that new temptations could bring about” (*MS* 6:384).³

To secure the viability of my interpretation, I must demonstrate that the possibility of a kind of gradual acquisition of a rational capacity that is not reducible to “physical inner necessitation” acquired through and resulting in “thoughtless repetition” is compatible with Kant’s account.

Kant touches upon the possibility (and indeed necessity) of acquiring rational capacities through repetition in the *Critique of the Power of Judgment* in a section devoted to the “maxims of common sense” (*KU* 5:294). Maxims of common sense, Kant says, are to serve as “universal rules” which can be derived a priori as guides to good judgment by virtue of their corrective function: they help avoid the illusion resulting from subjective private conditions being held to be objective (*KU* 5:293).

The three maxims of common sense are (1) “to think for oneself,” (2) “to think in the position of everyone else” and finally, (3) “always to think in accord with oneself” [*mit sich selbst einstimmig denken*]. The third, Kant says, is a maxim of a “consistent way of thinking” [*der consequenten Denkungsart*].

³ For a discussion, see Thomas Khurana’s “The Life of Freedom, Form and Actuality of Autonomy” (unpublished manuscript, March 2016), PDF.

Observance of the first two maxims includes, first, overcoming the tendency toward the prejudiced ways of thinking, and especially superstition, and second, they require a “way of thinking” in order “to make a purposive use” of “the faculty of cognition.” By virtue of this way of thinking the person “sets himself apart from the subjective private conditions of the judgment, within which so many others are as if bracketed, and reflects on his own judgment from a universal standpoint (which he can only determine by putting himself into the standpoint of others)” (KU 5:295).

About the third maxim of common sense, to “always to think in accord with oneself,” Kant says that it is “the most difficult to achieve, and can only be achieved through the combination of the first two and after frequent observance of them has made one proficient in them” [*Die dritte Maxime, nämlich die der consequenten Denkungsart, ist am schwersten zu erreichen und kann auch nur durch die Verbindung beider ersten und nach einer zur Fertigkeit gewordenen öfteren Befolgung derselben erreicht werden*] (KU 5:295, emphasis mine). Thus, to develop an aptitude of thinking in self-agreement with oneself, either theoretically and practically, requires regular observance of the maxims of “thinking for oneself” and “thinking in the position of everyone else.” And moreover, this observance must be frequent. *Fertigkeit*, the condition in which frequent observance culminates, means facility, proficiency or skill. The contexts in which one would expect it to appear are those of artistic skill and craftsmanship, where it would suggest the internalization of a technique: the capacity to perform a skill without having to deliberate from means to ends at each stage of the process.

A key to the reconciliation of the critique of habit and aptitude [*Fertigkeit*] (to which Kant seems to deny a role in the acquisition of virtue, “for then it would be a mere mechanism of

applying power”) on the one hand, and the possibility raised here of an aptitude [*Fertigkeit*] in self-agreement acquired through repeated observance, can be found in Kant’s discussion, in the *Metaphysics of Morals*, of two kinds of aptitude [*Fertigkeit*], free and unfree.

An *aptitude (habitus)* is a facility in acting and a subjective perfection of *choice*. But not every such *facility* is a *free* aptitude (*habitus libertatis*); for if it is a *habit (assuetudo)*, that is, a uniformity in action that has become a *necessity* through frequent repetition, it is not one that proceeds from freedom, and therefore not a moral aptitude. Hence virtue cannot be *defined* as an aptitude for free actions in conformity with law unless there is added “to determine oneself to act through the thought of the law,” and then this aptitude is not a property of choice but of the *will*, which is a faculty of desire that, in adopting a rule, also gives it as a universal law. Only such an aptitude can be counted as virtue. (*MS 6:408*)

Kant introduces *Fertigkeit* as a generic category of “facility in acting and a subjective perfection of *choice*.” Kant then divides this generic notion of an aptitude into two kinds: free and not free. Un-free aptitude is “uniformity in action that has become a necessity through frequent repetition.” The free aptitude, “an aptitude for free actions in conformity with the law,” i.e., virtue, is also uniformity in action that has become a necessity through repetition but it “proceeds from freedom.” The distinction lies in the free aptitude being an aptitude not merely to act a certain way (not to lie, for example), but “to determine oneself to act through the thought of the law” (not to lie in recognition of the authority of the law, “be true to your word”). Kant adds that “this aptitude is not a property of choice but of the *will*, which is a faculty of desire that, in adopting a rule, also gives it as a universal law.” By this, I contend, Kant means that it is not an aptitude to obtain certain means to given, particular ends, not an aptitude in obtaining objects of the power of choice, but instead is an aptitude for acting for ends that are good in themselves, i.e., the *energetic* ends of acting in consciousness of the moral law.

By the same account we can see how the notion of a free *Fertigkeit*, a rational aptitude,

avoids the second line of Kant's attack against habit, namely that virtue should "always emerge entirely new and original from one's way of thinking" (*A* 7:147) not only to secure freedom but also because otherwise, unless it "results from considered, firm, and continually purified principles," it is "like any other mechanism of technically practical reason, [...] neither armed for all situations nor adequately secured against the changes that new temptations could bring about" (*MS* 6:384). The free aptitude is an aptitude in acting on principle. As we are the originators of the principle, any action from consciousness of the principle would emerge new and original from one's way of thinking. Moreover, the facility in following a principle consists precisely in the ability to apply it well in unfamiliar circumstances: i.e., to respond well to situations in which the principle is relevant.

The adoption of the principle, as the adoption of the maxim of common sense, is a *Denkungsart*, one acquires the aptitude of its application gradually, in one's *Sinnesart*.

§5. Levels of Evil

We are now in a position to address the challenge with which I opened this chapter, namely, the apparent tension between Kant's general account of evil and his characterization of the first two levels of evil, frailty and impurity. The account of moral education just offered posits an initial determination of an orientation toward the good in one's *Denkungsart*, which one actualizes in gaining aptitude in the exercise of practical reason, or the gradual development of a *Sinnesart*.

Frailty and impurity can therefore be squarely located in the potential gap between the adoption of an attitude towards the good, i.e., one's *Denkungsart*, and the gradual development

of a *Sinnesart*. Note that in the case of both frailty and impurity, the maxim adopted is indisputably good. One's *Denkungsart* is not corrupted, but, in both cases, the moral *Triebfeder* is not sufficient to determine the action. In the one case, the frail agent performs a different action altogether and in the case of impurity, the agent is able to perform the action but only with the aid of motivational propping: the moral *Triebfeder* is necessary but is not sufficient for the performance of the action. What is lacking is a sufficient cultivation of one's *Sinnesart*.⁴

Second, taking these two levels of evil as states that characterize moral immaturity helps us both to explain the way in which a moral evaluation is possible in the absence of a stable, unifying *Gesinnung*, and, conversely, in identifying the role of moral evaluation in the acquisition of a moral character.

In brief, normative evaluation of cases of moral immaturity must be conducted with a view to whether or not the state of immaturity is one that is directed towards the acquisition of a moral character. This allows us to secure the applicability of moral praise and blame in the absence of a stable moral *Gesinnung*.

This solution is viable only if we can intelligibly maintain that every person who is morally immature is moving either towards or away from the acquisition of a good character. In

⁴ My account is therefore fundamentally incompatible with the interpretation of action from frailty that takes it to be a straightforwardly morally evaluable act. I hold that full practical knowledge is sufficient for acting well and deny that this requires explaining frailty by appeal to self-deception. For appeals to self-deception in the explanation of frailty, see Henry Allison, *Kant's Theory of Freedom*, 157, as well as Alexander Broadie and Elizabeth M. Pybus, "Kant and Weakness of Will," *Kant-Studien* 73, no. 1–4 (1982), 406–12; Stephen Engstrom, "Conditioned Autonomy," *Philosophy and Phenomenological Research* 48 no. 3 (March 1988), 435–453. Thomas E. Hill, Jr. mentions the possibility of self-deception as a way of explaining weakness of will in "Kant on Weakness of Will," *Virtue, Rules, and Justice* (New York: Oxford, 2012), 107–128.

The Metaphysics of Morals Kant claims as much. Concerning virtue, Kant avers, “if it is not rising, [it] is unavoidably sinking” (*MS* 6:409). This is because, it is “an ideal and unattainable” and “constant approximation to it is a duty.” Why is it that if someone is not moving towards virtue one is necessarily moving away from it? I suggest that this is because a person who fails to do the right thing, i.e., who does not act in consciousness that she is acting because it is her duty to act and is subsequently not concerned with it—who does not take a practical interest in preventing a similar failure from recurring in the future—is *thereby* adding moral insult to moral injury. This idea gives us the means to answer another question that hovers above Kant’s discussion of the levels of evil. This is the question of the sense in which the levels are ordered in increased severity.

The three levels of evil are meant to mark increasingly worse moral failures. This is despite the fact that it is only in the case of frailty, the first level, that non-compliance with the law is guaranteed by definition. The second level, impurity, is one in which the law is complied with, and this, non-accidentally, for compliance with the law is in fact intended: the maxim is good but its goodness is not sufficient to ground action, concrete purposes must be identified that the action will promote besides. The third level is one in which the maxim is not good but as long as compliance with the law is advantageous (accidentally, or because the civil laws enforced are constructed with a view to the moral law), it will be followed (“legally good [lawful] actions can always still consist with this, yet the way of thinking is thereby corrupted in its root” [*R* 6:30]). It is only in the case of frailty that, although the good (the law) is admitted into the maxim of the power of choice, an inclination is gratified instead.

I suggest that the increasing severity of the levels of evil is an expression of the predisposition's compatibility with a movement towards virtue. Frailty characterizes a subject in moral ascent. By definition, the frail subject fails to comply with a maxim she has adopted. Given that the maxim adopted is good, the subject will be aware of the disagreement between her activity and her principle, i.e., she will experience pain.⁵ This pain can in turn ground the increase of practical interest in preventing a similar failure from recurring in the future.

By contrast, although the state of impurity is one that is characterized by compliance with the law, it is a tendency to search for a sensible *Triebfeder* to justify the performance of one's duty. Therefore it is characterized by a dependence on sensible *Triebfedern* as necessary conditions for the performance of actions from duty, closer to the adoption of a bad *Gesinnung*, whereby the sensible *Triebfedern* are sufficient conditions for determination to action.

§6. Conclusion

The resulting account of feeling and moral character offers an understanding of practical cognition whereby feeling and character grounded in habits of feeling, although not determined by acts of ordinary choice, are not hindrances to reason or obstacles to ethical life. On the

⁵ Interpreters who (typically in an attempt to refute Allison's cognitivist denial of a schism between objective principles and subjective motivation) draw attention to the fact that the frail subject is committed to the objective principles at issue and therefore finds the experience of acting against her representation of the good painful, were initially led by Marcia Baron's "Freedom, Frailty, and Impurity," *Inquiry* 3, no. 4 (1993), 431–441; Richard McCarty follows suit in "Moral Weakness as Self-Deception," *Proceedings of the Eighth International Kant Congress*, vol. 2. Milwaukee: Marquette University Press, 1995, 587–593. The objection has been echoed by Robert N. Johnson, "Weakness Incorporated," *History of Philosophy Quarterly* 15, no. 3 (July 1998), 349–367; and Iain Morriison, "On Kantian Maxims: A Reconciliation of the Incorporation Thesis and Weakness of the Will," *History of Philosophy Quarterly* 22, no. 1 (January 2005), 73–89. See also McCarty, "Kantian Moral Motivation" 31, no. 3 (July 1993), 421–435.

contrary, they are the necessary conditions of freedom.

In doing so, I have attempted to motivate, at least in part, Heidegger's claim concerning the centrality of moral respect to Kant's own philosophical project as well as the depth of the insight that Kant offers with this account. I have sought to clarify the sense in which the moral "I" is constituted essentially by the feeling of respect, which is not only constitutive of moral action but is a way of being self-conscious, indeed, the essentially human way of being self-conscious.

I hope to have thereby done something by way of achieving a clarification of the whole of Kant's practical philosophy—that I have let Kant say what he intended to say. Perhaps, I have also let him say more, a "more" that rests on philosophical interpretation. If so, the task of taking back the "more" that interpretation has added—a task Heidegger claimed could be readily performed on the basis of the philosophical understanding thereby achieved—is one that I must leave for the future.

Bibliography

- Allison, Henry E. *Idealism and Freedom: Essay on Kant's Theoretical and Practical Philosophy*. New York: Cambridge University Press, 1996.
- . *Kant's Theory of Freedom*. New York: Cambridge University Press, 1990.
- Bagnoli, Carla. "Emotions and the Categorical Authority of Moral Reason." In *Morality and the Emotions*, edited by Carla Bagnoli, 62–80. New York: Oxford University Press, 2011.
- Baron, Marcia. "Freedom, Frailty, and Impurity." *Inquiry* 3, no. 4 (1993): 431–441.
- . *Kantian Ethics Almost Without Apology*. Ithaca, NY: Cornell University Press, 1995.
- Baxley, Anne Margaret. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press, 2010.
- Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. 2nd ed. Chicago: University of Chicago Press, 1996.
- Beere, Jonathan. *Doing and Being: An Interpretation of Aristotle's Metaphysics Theta*. Oxford Aristotle Studies. Oxford: Oxford University Press, 2009.
- Boyle, Matthew. "Additive Theories of Rationality: A Critique," *European Journal of Philosophy* 24, no. 2 (June 2012): 527–555.
- Broadie, Alexander and Pybus, Elizabeth M. "Kant and Weakness of Will." *Kant-Studien* 73, no. 1–4 (1982): 406–12.
- Burnyeat, Myles F. "Aristotle on Learning to Be Good." In *Essays on Aristotle's Ethics*, edited by Amélie Oksenberg Rorty, 69–92. Berkeley: University of California Press, 1981.
- Caswell, Matthew. "Kant's Conception of the Highest Good." *Kant-Studien* 97, no. 2 (2006), 184–209.
- Cohen, Alix, ed. *Kant on Emotion and Value*. London: Palgrave Macmillan, 2014.
- Darwall, Stephen L. "Kantian Practical Reason Defended." *Ethics* 96, no. 1 (1985): 89–99.
- . "Two Kinds of Respect." *Ethics* 88, no. 1 (October 1977): 36–49.
- Engstrom, Stephen. "Conditioned Autonomy." *Philosophy and Phenomenological Research* 48, no. 3 (March 1988): 435–453.

- . *The Form of Practical Knowledge: A Study of the Categorical Imperative*. Cambridge, MA: Harvard University Press, 2009.
- . “The Inner Freedom of Virtue.” In *Kant’s Metaphysics of Morals: Interpretative Essays*, edited by Mark Timmons. New York: Oxford University Press, 2002: 289–315.
- . “Kant on the Agreeable and the Good.” In *Moral Psychology*. Poznan Studies in the Philosophy of the Sciences and the Humanities, edited by Sergio Tenenbaum (vol. 94, no. 1). 111–60. Amsterdam: Rodopi, 2007.
- . “Kant’s Distinction Between Theoretical and Practical Knowledge.” *The Harvard Review of Philosophy* 10, no. 1 (2002): 49–63.
- . “The Triebfeder of Pure Practical Reason.” In *Kant’s Critique of Practical Reason*, edited by Andrews Reath and Jens Timmerman. Cambridge: Cambridge University Press, 2010.
- . “Understanding and Sensibility.” *Inquiry* 49, no. 1 (2006): 2–25.
- Geiger, Ido. “Rational Feelings and Moral Agency.” *Kantian Review* 16, no. 2 (June 2011): 283–308.
- Grenberg, Jeanine. “Making Sense of the Relationship of Reason and Sensibility in Kant’s Ethics.” *Kantian Review* 16, no. 3 (September 2011): 461–72.
- Guyer, Paul. *Kant and the Experience of Freedom: Essays on Aesthetics and Morality*. Cambridge: Cambridge University Press, 1993.
- Haase, Matthias. “Life and Mind.” In *The Freedom of Life: Hegelian Perspectives*, edited by Thomas Khurana, 69–109. Berlin: August Verlag, 2013.
- Heidegger, Martin. *The Basic Problems of Phenomenology*. Translated by Albert Hofstadter. Bloomington: Indiana University Press, 1981.
- . *The Essence of Human Freedom: An Introduction to Philosophy*. Translated by Ted Sadler. London: Continuum, 2002.
- . *Kant and the Problem of Metaphysics*. Translated by Richard Taft. Bloomington: Indiana University Press, 1962.

- . *Phenomenological Interpretation of Kant's Critique of Pure Reason*. Translated by Parvis Emad and Kenneth Maly. Bloomington: Indiana University Press, 1997.
- Herman, Barbara. "Making Room for Character." In *Aristotle, Kant, and the Stoics: Rethinking Happiness and Duty*, edited by Stephen Engstrom and Jennifer Whiting, 36–60. Cambridge: Cambridge University Press, 1996.
- Herrera, Larry. "Kant on the Moral Triebfeder." *Kant-Studien* 91, no. 4 (2000): 395–410.
- Hill, Thomas E. Jr. *Dignity and Practical Reason in Kant's Moral Theory*. Cornell University Press, 1992.
- . "Kant on Weakness of Will." In *Virtue, Rules, and Justice*, 107–128. New York: Oxford University Press, 2012.
- Johnson, Robert N. "Weakness Incorporated." *History of Philosophy Quarterly* 15, no. 3 (July 1998): 349–367.
- Kant, Immanuel. *Critique of the Power of Judgment*. Translated by Paul Guyer and Eric Matthews. The Cambridge Edition of the Works of Immanuel Kant, edited by Paul Guyer. Cambridge: Cambridge University Press, 2000.
- . *Critique of Pure Reason*. Translated by Paul Guyer and Allen W. Wood. The Cambridge Edition of the Works of Immanuel Kant, edited by Paul Guyer and Allen W. Wood. Cambridge: Cambridge University Press, 1998.
- . *Notes and Fragments*. Translated by Curtis Bowman and Paul Guyer. The Cambridge Edition of the Works of Immanuel Kant in Translation, edited by Paul Guyer. Cambridge: Cambridge University Press, 2005.
- . *Practical Philosophy*. Translated by Mary J. Gregor. The Cambridge Edition of the Works of Immanuel Kant, edited by Mary J. Gregor. Cambridge: Cambridge University Press, 1996.
- . *Religion within the Bounds of Bare Reason*. Translated by Werner S. Pluhar. Indianapolis: Hackett Publishing Company, 2009.
- Kenny, Anthony. *Action, Emotion, and Will*. 2nd ed. London: Routledge, 2003.
- Khurana, Thomas. "The Life of Freedom, Form and Actuality of Autonomy." Unpublished manuscript, last modified March 2016. PDF.
- Korsgaard, Christine M. *The Constitution of Agency: Essays on Practical Reason and Moral Psychology*. Oxford: Oxford University Press, 2008.

- . *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- . *Self-Constitution: Agency, Identity, and Integrity*. Oxford University Press, 2009.
- . *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- Land, Thomas. “Kantian Conceptualism.” In *Rethinking Epistemology*, edited by Günter Abel and James Conant, 197–239. Berlin: De Gruyter, 2011.
- McCarty, Richard. “Kantian Moral Motivation and the Feeling of Respect.” *Journal of the History of Philosophy* 31, no. 3 (1993): 421–35.
- . *Kant’s Theory of Action*. Oxford: Oxford University Press, 2009.
- . “Moral Weakness as Self-Deception.” *Proceedings of the Eighth International Kant Congress*, vol. 2., 587–593. Memphis, TN . Milwaukee: Marquette University Press, 1995.
- McDowell, John. “Avoiding the Myth of the Given.” In *Having the World in View: Essays on Kant, Hegel, and Sellars*, 256–272. Cambridge, MA: Harvard University Press, 2013.
- Morrisson, Iain. “On Kantian Maxims: A Reconciliation of the Incorporation Thesis and Weakness of the Will.” *History of Philosophy Quarterly*. 22, no. 1 (January 2005), 73–89.
- Muchnik, Pablo. “The Heart as Locus of Moral Struggle in Religion.” In *Kant on Emotion and Value*, edited by Alix Cohen, 224–244. London: Palgrave MacMillan, 2014.
- Munzel, G. Felicitas. *Kant’s Conception of Moral Character: The “Critical” Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press, 1999.
- O’Neill, Onora. *Acting on Principle: An Essay on Kantian Ethics*. New York: Columbia University Press, 1975.
- . *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. Cambridge: Cambridge University Press, 1989.
- Reath, Andrews. “Kant’s Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination.” *Kant-Studien* 80, no. 1–4 (1989): 284–302.
- Rödl, Sebastian. “The Form of the Will.” In *Desire, Practical Reason, and the Good*, edited by Sergio Tenenbaum, 138–159. Oxford: Oxford University Press, 2010.
- . *Self-Consciousness*. Cambridge, MA: Harvard University Press, 2007.

- Schapiro, Tamar. "Foregrounding Desire: A Defense of Kant's Incorporation Thesis." *Journal of Ethics* 15, no. 3 (September 2011): 147–67.
- Sherman, Nancy. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press, 1997.
- Thorndike, Oliver. "Kant's Transition Project." Paper presented at the North American Kant Society Meeting, Emory University, Atlanta, GA, May 2016.
- Timmermann, Jens. "Kant on Conscience, 'Indirect' Duty, and Moral Error." *International Philosophical Quarterly* 46, no. 3 (March 2006): 293–308.
- Ware, Owen. "Kant on Moral Sensibility and Moral Motivation." *Journal of the History of Philosophy* 52, no. 4 (2014): 727–46.
- Wolff, Robert Paul. *The Autonomy of Reason*. New York: Harper & Row, 1973.
- Wood, Allen W. *Kantian Ethics*. Cambridge: Cambridge University Press, 2008.
- . *Kant's Ethical Thought*. Modern European Philosophy. Cambridge: Cambridge University Press, 1999.
- . *Kant's Moral Religion*. Ithaca, NY: Cornell University Press, 1970.