Rethinking Kant
Volume 6

Edited by

Edgar Valdez

Cambridge
Scholars
Publishing

Rethinking Kant Volume 6

Edited by Edgar Valdez

This book first published 2022

Cambridge Scholars Publishing

Lady Stephenson Library, Newcastle upon Tyne, NE6 2PA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN (10): 1-5275-8437-2
ISBN (13): 978-1-5275-8437-2

EVIL OR ONLY IMMATURE? KANT AND THE COMPLEXITY OF MORAL FAILURE

ANASTASIA BERG

I. Introduction

In his late work *Religion within the Bounds of Mere Reason* Kant famously argues that an agent cannot act at times morally well and at others morally badly. The moral quality of all of an agent's actions depends on the moral quality of their moral character. Since their moral character can be either absolutely good or absolutely bad, all of an agent's actions share the same moral quality: good or evil (R 6: 22).[1] This claim, which implies that any agent who is not *wholly* good must therefore be wholly evil, has vexed Kant's readers. Ordinary moral intuitions suggest that differences in moral character come in degrees, and leaving some room for moral variance seems necessary to any account of moral corruption and improvement.[2]

---

[1] Kant refers to this claim as a case of "rigorism." This is not to be confused with the interpretative literature's use of the term "rigorism" to refer to Kant's position with regard to moral duties, namely, his apparent refusal to admit any exceptions to strict moral duties and in particular to the prohibition against lying. On the latter problem, as well as for an exposition and a thorough review of the literature, see Tamar Schapiro, "Kantian Rigorism and Mitigating Circumstances."

[2] As Stephen Engstrom has noted, Kant seems to refer to the possibility of gradual improvement himself. In particular, Engstrom, argues, Kant's argument for the immortality of the soul relies on its assumption (cf. KpV 122-23) (Stephen Engstrom, "Conditioned Autonomy"). In the same piece, Engstrom provides an exposition and a proposed "Kantian" alternative to Kant's own account of moral character rigorism. See also Henry Allison, *Kant's Theory of Freedom,* 146–48. Daniel O'Connor condemns Kant's account for this consequence, namely, its inability to help us "appreciate the indefinite variety of strengths and weaknesses of moral character which human beings manifest" (Daniel O'Connor, "Good and evil disposition", 302). Kelly Coble argues that Kant's rigorism does not have the negative consequences that Engstrom and others find in it. Instead, she argues, "Kant's bipolar view of dispositions must be assessed in light of his dynamic theory

What is not often remarked upon is that Kant's rigorism about moral character stands in apparent tension with his own account, given that in the same work and only a few pages after the exposition of rigorism, of the *Stufen*, levels or grades, of evil: frailty, impurity, and wickedness. When interpreters do note the tension, they take it as license to dismiss Kant's commitment to rigorism.[3] I argue instead that rigorism is grounded in a philosophical insight which we should not give up and show that we can preserve it while making room for the complexity of moral failure if we understand the first two grades of evil, frailty and impurity, as states of moral *immaturity*, a condition that precedes the acquisition of a stable moral character.

To substantiate the 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. Not merely an "empirical" concern, the acquisition of character must be understood as a genuinely rational accomplishment: the development and determination of our uniquely rational capacities for feeling, necessary for the development of stable moral character.

II. The Conflict between Rigorism and the Levels [*Stufen*] of Evil

Kant's account of evil must be read against the background of his account of character, or moral disposition—the moral *Gesinnung*.[4] The

---

of character, his assumption that ethical attitudes are called good or evil with reference to the quality of agents' moral change over time, not their unvarying inner purity or perversity" (Coble, "Kant's Dynamic Theory of Character," 41). I agree with Coble that as a matter of moral practice the only way we have to evaluate moral conduct is on the basis of the quality of the agent's moral change, but I argue in what follows that we can only make sense of this practice if we accept moral character is not simply merely "dynamic" but is acquired over time. The agent whose character is not of a unified quality—neither wholly good nor wholly evil—is morally immature.

[3] See Coble and Engstrom.

[4] I follow Julia Peters ("Kant's *Gesinnung*") in leaving the term *Gesinnung* untranslated. As she notes, rendering *Gesinnung* as 'disposition' (as is the case in both George di Giovanni's translation in the Cambridge Edition of the Works of Immanuel Kant and the earlier translation by Theodore M. Greene and Hoyt H. Hudson) is misleading, insofar as 'disposition' can also refer to a behavioral disposition or tendency, a connotation absent from the German *Gesinnung*. 'Attitude,' Werner Pluhar's preferred translation, is closer to the German meaning.

Gesinnung is a freely determined moral attitude. This moral attitude is identified by a maxim (the principle of activity upon which the subject actually acts) of utmost generality, which serves as the fundamental rule for the exercise of the free power of choice. Kant calls it the "supreme maxim" (R 6: 36) and characterizes it as a unifying rule that the power of choice "makes for itself". It is, he writes the "first basis […] for the adoption of good maxims or the adoption of evil (unlawful) ones" (R 6: 21); it "can only be one" and it "applies universally to the entire use of freedom" (R 6: 25). By adopting this supreme maxim an agent adopts a governing principle for the exercise of their will in its entirety. Such a principle thereby constitutes the agent's moral character at the highest level of generality. Since this principle establishes the agent's character as good or bad, and since every action is ultimately guided by this principle, any action's moral worth or lack thereof is derived fundamentally from the overarching principle of activity in which it is grounded, the *Gesinnung*.

How exactly does this principle determine moral worth? The supreme principle determines the subordination of the two "*Triebfedern*", moral or sensible, to one another. If the agent subordinates the sensible to the moral, the agent's *Gesinnung* is good, if the agent subordinates the moral to the sensible, the agent's *Gesinnung* is bad. What then is a *Triebfeder*? Kant defines *Triebfeder* as the "subjective ground of desire" (*G* 4: 427) or "subjective determining ground of the will" (KpV 5: 72*).* But what is it that "determines the will"? *Triebfeder* is most commonly translated as 'incentive.' This is somewhat misleading, however, because, as interpreters have noted, Kant does not have in mind an object or circumstance that incites the agent to act, like a tax incentive.[5] Interpreters have proposed instead that a *Triebfeder* is best thought of as something like "the dynamic or conative factor in willing" (Beck 1960, 216), or "the driving or propelling force from which action or effort springs" (Herrera 2000, 395), as opposed to the cognitive or representational factor.

But, as Stephen Engstrom has convincingly argued, the strict identification of the *Triebfeder* with the conative aspect of willing misses the implication of a key part of Kant's account in the second *Critique*: namely, the idea that when the agent acts from consciousness of the moral

[5] This is the sense in which Korsgaard, for example, interprets it: "An incentive is a motivationally loaded representation of an object." (Korsgaard, *Self-Constitution*, 104-105). Earlier however she claims more simply that "incentive" is just Kant's "own language" for "desire or impulse" (Ibid., 72). The translation of *Triebfeder* by incentive is persuasively objected to by Larry Herrera, "Kant on the Moral Triebfeder," 395fn1 and Stephen Engstrom, "The Triebfeder of pure practical reason," 91-93. I leave the term untranslated.

law the moral law *itself* is the *Triebfeder* of the human will.[6] Thus, Engstrom writes, in order for the moral law to move you, the objective principle (the moral law) must become the subjective principle (your maxim) and thereby become itself *Triebfeder*. In other words, the *Triebfeder* is not something that motivates the agent to adopt a certain principle, but it is a principle of action—in this case the moral law—*as motivating* (Engstrom 2010, 93).

Engstrom is only concerned with the moral *Triebfeder*, but I propose to extend the suggestion we ought to understand the *Triebfeder* as an actually motivating principle, to both of its instances: moral and sensible. Accordingly, I propose we interpret the distinction between the moral and non-moral *Triebfedern* against the background of Kant's distinction between moral and non-moral *principles* of action. That is, I propose we understand the contrast between non-moral *Triebfedern* and moral *Triebfedern* as the contrast between acting on *material* (instrumental) principles and acting on *formal* (moral) ones. Accordingly, to admit the non-moral *Triebfeder* into one's "supreme maxim" is therefore to admit instrumental principles of action into one's will. To admit both kinds of *Triebfeder* into one's supreme maxim, as Kant says we do, is therefore to admit both kinds of principles into one's organizing principle of action. And to subordinate principles to one another means subordinating the activity from one principle to activity from the other.

Good and bad characters reflect the way the two kinds of *Triebfedern*, moral and sensible, are subordinated to one another. That is, they reflect which kind of principle of action an agent prioritizes in their practical life. To subordinate one kind of *Triebfeder* to another means determining whether one acts for particular ends only as subordinated to moral ones, or prioritize the pursuit of particular ends as if they were good in themselves. We can therefore think of the moral *Gesinnung* as the subordination of principles of action to one another by which an agent maintains their practical identity. Moral goodness is activity grounded in a motivational structure whereby activity from instrumental principles in the pursuit of particular, sensible ends is subordinated to activity whose principle is the moral law. Evil actions are grounded instead in giving, within one's *Gesinnung*-constituting maxim, the sensible *Triebfedern*, i.e., purely instrumental pursuits, *absolute* priority in action.

---

[6] Indeed, while in the *Groundwork* Kant uses the term *Triebfeder* exclusively with reference to non-moral activity (Cf. G 4: 400, 404, 407, 411, 412, 419, 425, 427, 431, 439, 441, 444, 449, 461). from the second *Critique* onward, Kant employs the notion of the *Triebfeder* also to express what it is for reason to be practical in a finite rational being: *the moral law itself* must serve as the will's *Triebfeder* (KpV 5: 72).

Although Kant mentions the need for a single, unifying principle for the power of choice for the first time in the *Religion*, it is implicit in his earlier, central thesis that the moral worth of an action depends on the principle on which an agent acts. In the *Religion* Kant makes explicit an implication of this idea, namely, that because the moral law commands universally, properly recognizing its authority means recognizing its authority universally, or, over all of one's actions. In particular it means that an agent cannot be thought of as only occasionally acting from genuine recognition of the moral law. To allow oneself to occasionally deviate from the commands of the moral law is simply to refuse its authority as universally binding, which is to refuse its authority *tout court*. An agent is thus either committed to the supremacy of moral considerations in all matters, or, in effect, in none.[7]

This is why 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 22), and therefore that he is not "both simultaneously, namely in some points good, in others evil" (*ibid*.) His argument is as follows. Kant holds that in order for action to be free and imputable to the agent, the power of choice can be determined to an action by a certain kind of ground only insofar as the subject has admitted this ground into her character-constituting maxim as sufficient for the determination of one's will. Therefore, if someone acts such that the law does not determine her power of choice, then she acts on a non-moral ground, or an instrumental principle that was not in turn subordinated to a moral one. This means, in turn, that it must have been freely admitted as a ground into her character-constituting maxim. But Kant also holds, as we have seen, that morally evaluating any action requires attending to its maxim, its principle, and that all particular principles are ultimately grounded in the one character-constituting maxim.

Therefore, Kant concludes, 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)

[7] The claim that that maxims stand to one another in a hierarchical system have been defended by Beck, *A Commentary on Kant's Critique of Practical Reason*, 118; Korsgaard, *Creating the Kingdom of Ends*, 58; Onora O'Neill, *Constructions of Reason*, 83–85; and Allison, *Kant's Theory of Freedom*, 91–94.

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 ground into one's most general maxim of choice. However, this recognition of the necessity and universality of the law is not consistent with the occasional admission of exceptions. If one gives priority to the pursuit of a sensible/particular end, a purpose, in one's determination to action, one cannot be said to have adopted the good character-constituting principle, for the latter would preclude the prioritization of a sensible ground. Therefore, since one is good if and only if one has recognized the necessary and universal validity of the law, and this is incompatible with occasional deviation, one cannot be good at one time and evil at another. Kant rejects by this very argument the idea that incorporation of a *Triebfeder* can be a local affair, but aims to show why any seemingly local deviation from moral commitment is wholly determining of the subject, why any admission of a sensible *Triebfeder* is just thereby a constitution of the supereme, *Gesinnung*-constituting maxim as an evil one.

On the face of it, 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 complexity of moral imperfection, namely, his 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 grounds of the power of choice, every one of its species ought to be a species of this general perverse reversal: each of the levels of evil should instantiate this structure and offer a further determination of it. But Kant's account of the levels of evil, which appears just a few pages after his brief defense of rigorism, fails to meet this expectation.

Firstly, the definition of wickedness (*Bösartigkeit*), the third and highest level of evil, is identical to the general definition of evil: wickedness, Kant writes, "reverses the moral order in regard to the incentives [*Triebfedern*] of the *free* power of choice" (R 6: 30). Wickedness is not so much a sub-kind of a general type, but the type itself.

Secondly, the first two levels of evil are not characterized in terms of the general definition, at all, and in fact appear to be in conflict with it.

Frailty (*Gebrechlichkeit*) 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. The frail agent admits the good into the maxim of her power of choice and yet fails to act on it. In other words, they are, say, aware of the principle upon which they should act, but do not follow through with action. The putative explanation—that the *Triebfeder* which is objectively insurmountable is subjectively the weaker—does little to help us make sense of this. Which maxim an agent ends up prioritizing cannot be a matter of which is "weaker" and which "stronger"; rather, it is a matter of how the agent herself has subordinated the moral to the sensible, or vice versa. If an agent acts on a sensible ground, then they must have freely admitted it as sufficient for the determination of action in the character-constituting maxim. A sensible ground cannot, of itself, overcome the determination to act in recognition of the authority of the moral law or else the action would no longer be free. If it were no longer free it would not be imputable to the agent. It is therefore hard to see how frailty is a species of Kant's general account of evil.

Impurity (*Unlauterkeit*), a state where the performance of good action requires the cooperation of sensible grounds 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, but usually (perhaps always) still needs other incentives 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 non-moral interests she is pursuing. 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 grounds is not simply

perversely reversed, the *Triebfedern* are instead "mixed." Compliance with the law is "intended"; the maxims are maxims "of the good."[8] In his description of impurity as well as the brief discussion 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 of her non-moral interests alone. She looks for incentives to act only upon recognizing 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 looks for other advantages the action might afford her in order to motivationally prop herself up, so to speak.

Actions grounded both in frailty and impurity seem therefore 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 character-constituting maxim or that of the first two levels of evil—must be jettisoned.

However, dismissing the account of the character-constituting maxim would leave Kant with no resources to explain how evil actions determined by non-moral grounds are free and therefore imputable to the subject, and hence with no resources to explain how can they be morally evaluable. On the other hand, dismissing frailty and impurity would likewise come at a high philosophical cost. First, frailty and impurity 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 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. These distinctions capture the way most people take themselves to be. To deny the possibility of evil actions as occasional

[8] This characterization points to a familiar moral phenomenon, and recalls Kant's description in the Preface to the First Edition of the *Religion* 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). Kant 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," he describes a moral failure of the following sort: "When we bring any flattering thought of merit into our action, then the incentive is already somewhat mixed with self-love and thus has some assistance from the side of sensibility" (KpV 5: 159). The agent compensates herself for the sacrifice of some particular ends (purposes) by identifying other particular ends that her duty-conforming action may promote.

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, ultimately, 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—has always been sensibly conditioned, after all. The bigger problem however is that, 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—denying any gradation in states of imperfection renders it impossible to account for moral differences between moral subjects. In particular, denying moral grades makes it impossible to render intelligible 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. 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*.

There are two obvious obstacles to this proposal, and we would do well to meet both head on:

First, 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) (§3-4). Once we overcome the first challenge we will be able to attend to the second. This is the worry that 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 (§5). I will argue that though a morally immature agent cannot be evaluated strictly based on the moral quality of their *Gesinnung*, they can be morally evaluated according to whether they are becoming relatively morally better, or worse.

III. 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 same text, as well as from the *Critique of Practical Reason,* seem explicitly to presuppose just such an account.

First, Kant's claims that a person's *Gesinnung* is "laid … in earliest youth" (R 6: 21), and is conceived as present in the human being simultaneously with birth (R 6: 22), seem to leave no room for the idea of acquiring a moral *Gesinnung*, for they apparently preclude the possibility of any *morally* evaluable state that is not already a good or evil *Gesinnung* itself.

Kant likewise seems hostile to the idea of gradated morally relevant 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 […] must be brought about through a *revolution* in the attitude in the human being […]; 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 moral 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. There, Kant claims that "personality", that constituent of "human nature" by virtue of which reason can determine the power of choice unconditionally, i.e., the essential determination of a being in whom reason is practical, is "the receptivity to respect for the moral law, as an incentive, sufficient by itself, of the power of choice" (R 27). He explains,

> 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. (R 6:27, emphasis mine)

Here Kant claims that good moral character, the admission and prioritization of the moral *Treiebfeder* in one's highest maxim, is itself an acquired capacity.

Likewise, in both of Kant's accounts of moral education, in the *Religion* and the second *Critique*, moral education is characterized as requiring 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 the *Gesinnung*, the moral character. In both the *Critique* (KpV 5: 152-161) 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 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 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. This however must be gradually "cultivated" through the presentation of examples of actions that conform to the law. Hereby "moral apprentices" [*moralischen Lehrlinge*] are invited 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).[9]

[9] My aim in this section has been to establish that Kant recognized that the adequate exercise of our rational practical capacities requires habituation, that is to say, moral education. For recent discussions focused on the nature and significance of moral

IV. 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 with his descriptions of the gradual cultivation of a capacity for moral feeling, which is necessary for the prioritization of the moral *Triebfeder* in one's *Gesinnung*? To do so we must attend to a distinction Kant introduces between an agent's *Denkungsart* and her *Sinnesart*, roughly, the way of thinking and the way of sensing. Moral goodness, Kant claims, requires a revolution in the agent's way of thinking, but it must be supplemented with a gradual reform in the way of sensing:

> 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). (R 6: 47)

But what might this mean? On the face of it, it might seem that the distinction between the way of thinking and way of sensing could simply be mapped onto the distinction between intelligible character and sensible character—i.e., an agent's real moral character and the apparent conformity of her actions to the moral law. On such a reading the *Denkungsart* is simply identical to the intelligible *Gesinnung*, and the *Sinnesart* is morally superfluous.[10]

---

education in Kant see Munzel, "'Doctrine of Method' and 'Closing'" and *Kant's Conception of Pedagogy*; Koch, *Kants ethische Didaktik*; Moran, "Can Kant Have an Account of Moral Education?"; *Community and Progress in Kant's Moral Philosophy*, Ch. 3; Louden, *Morality and Moral Theory*, Ch. 11; and the essays in Roth and Surprenant's *Kant and Education*. For a discussion of how to reconcile the significance of moral education, which necessarily relies on the presence of individuals and institutions other than the agent herself, and the agent's *autonomy*, see Beck *A Commentary on Kant's Critique of Practical Reason*, 235; "Kant on Education", 22 and Formosa, "From Discipline to Autonomy: Kant's Theory of Moral Development", 163–77. For a discussion of the tension between moral education, and Kant's optimism that the *common agent* is already *morally competent* see Sticker, "Educating the Moral Agent: Kant on the Varieties of Moral Education".

[10] 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

The main reason that the gradual reform in the *Sinnesart* appears morally superfluous is that in surrounding passages, the revolution in the *Denkungsart* is indeed contrasted with a kind of "gradual reform" that is not morally laden. What is too often sorely overlooked, however, is that there are *two distinct* processes that Kant characterizes by speaking of "gradual reform." First, as we have seen earlier, 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," the one that occurs while the foundation of maxims is impure, is coextensive with Kant's notion of merely phenomenal virtue: permanent performance of "lawful actions." This kind of virtue is "acquired little by little, and means to some a long habituation (in observing the law)," it involves "gradual reform of his conduct and stabilization of his maxims" (R 6: 47). This 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 Kant claims must accompany the revolution in the way of thinking as identical to the gradual reform of conduct that goes along with habituation in performing merely lawful actions, and therefore dismiss the gradual reform in the *Sinnesart* as morally irrelevant. This conclusion would be however hasty. This is because the gradual reform in the Sinnesart is not identical to a gradual reform in mores.

Kant glosses the revolution in *Denkungsart* and gradual progress in *Sinnesart* in the following way. First, Kant indeed singles out the revolution in the *Denkungsart* as a precondition of the constitution of good character:

---

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 (see Munzel, 1999). Barbara Herman provides an example of the second type of reading when she 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 (1996, 37; cf. 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" (1997, 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 the constitution of moral 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. (R 6: 47-48)

A so-called revolution in the *Denkungsart*, in the way of thinking, is thus 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. 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)

A gradual reform in *Sinnesart* must follow upon a revolution in *Denkungsart*. Thus, not every "gradual reform" pertains merely to the legality of actions. The good *Denkungsart* renders the subject receptive to the good. The arousal of the moral feeling, i.e., the cultivation of the predisposition to the good, renders the subject receptive to the good in actuality, and is therefore necessary for complete good *Gesinnung*. The *Denkungsart*, in other words, is a necessary but not sufficient condition for the actualization of the predisposition to personality. Full actualization requires the further gradual reform of a *Sinnesart*, as well.

This might sound like a strange reversal of what is customarily taken to be the Aristotelian insight concerning moral education. On that account, 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.[11] I would like to suggest that the priority of the constitution of the *Denkungsart* over the gradual reform in the *Sinnesart* should not be understood as standing in any tension with this important insight. In particular, I suggest that the constitution of the *Denkungsart* is not a state of practical wisdom, i.e., full knowledge of the good (knowledge of what courses of action are good and why they are good), but is instead an attitude toward the good. The agent with a good *Denkungsart* is positively

[11] 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.

inclined toward the good. To be oriented toward the good is a precondition of the acquisition of desires to perform actions that are good, as well as a precondition of the acquisition of the knowledge of *why* particular 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 by another 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 (to do the good for its own sake). In a Kantian idiom, they would learn to act according to the good, but not from the good. 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 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, "The supernatural being will be happy not to have to punish you, as it 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 supernatural beings. 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. The child must not only be motivated by threats of punishments and promises of material rewards, but by appeal to the moral effects and moral quality of their

actions.[12] To the degree the instruction is explicitly moral, the child participates in the process of moral habituation as a moral apprentice. To this extent this 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). It is thereby that the moral teacher is able to help the child gradually develop her own *Sinnesart*, and thus become good *in concreto*.

V. Levels of Evil

We are now in a position to address the challenge with which I opened this essay, 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 gradually gaining aptitude in the exercise of practical reason, or the gradual development of a *Sinnesart*. Thus, we see the way in which a stable moral character requires acquisition, after all.

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 they recognize is good but only with the aid of motivational propping. In both cases, a necessary orientation towards the good is present, but sufficient cultivation of one's *Sinnesart* is lacking.[13]

---

[12] For Aristotle, such an appeal to the good must be mediated by the appeal to honor and shame. Only those who are taught to care about honor and shame can come to care about the good.

[13] 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*, (New York: Cambridge University Press, 1990), p. 157. As well as, Alexander Broadie and Elizabeth M Pybus, "Kant and Weakness of Will" and Stephen Engstrom, "Conditioned Autonomy".

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, to identify 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 the state of immaturity is one that is directed towards the acquisition of a moral character or not. This allows us to secure the applicability of moral praise and blame in the absence of a stable moral *Gesinnung*.

This solution is of course 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 *The Metaphysics of Morals* Kant claims just this. Concerning virtue, Kant claims, "if it is not rising, [it] is unavoidably sinking" (MS 6: 409). But why is it that if someone is not moving towards virtue one is necessarily moving away from it? Why, in other words, can't one be stably immature? I suggest that this is because a person who fails to do the right thing, and is subsequently unconcerned with her moral failure—who does not take a practical interest in preventing a similar failure from recurring in the future—is just *thereby* adding moral insult to moral injury. And, more importantly, by failing to take any measures that would prevent the recurrence of such a moral lapse in the future, is making it increasingly more likely that she will fail again. Her moral badness will harden into bad habit. In this context it might be worth recalling that for Aristotle one cannot be weak of will but on occasion, or else the agent turns vicious.

This idea gives us the means to answer yet another question that hovers over 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 and 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.[14] Only in the case of frailty although the good (the law) is admitted into the maxim of the power of choice, an inclination is gratified instead. Why then is the first grade the least bad and the last the worst?

[14] "Legally good (lawful) actions can always still consist with this [with wickedness]" (*R* 6:30).

The account I have proposed allows us to see that the increasing severity of the levels of evil is an expression of the state'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, and she will experience this with displeasure.[15] This pain has the potential to increase the agent's 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 in character to the adoption of a bad *Gesinnung*, whereby the sensible *Triebfedern* are taken as sufficient conditions for determination to action. The impure agent ends up acting in accordance with the law and is therefore less likely to be made aware of their moral failure.

VI. Conclusion

On the interpretation I offered, the *acquisition* of character in Kant emerges as a rational accomplishment: the development and determination of our uniquely rational capacities for feeling, and, in particular, for moral feeling. This reading not only allows us to understand better the attractions of rigorism, but also to gain insight into the nature of the complexity of moral failure. Thereby, I hope, it will allow us to better appreciate Kant's nuanced understanding of practical cognition, as an account that recognizes our forms of dependency (our capacity for feeling and the character grounded in feeling) not as threats to the exercise of reason and obstacles to ethical life but as the necessary conditions of actual freedom.

---

[15] Interpreters, following Marcia Baron's account in "Freedom, Frailty, and Impurity," have drawn 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. This is often done in an attempt to refute Allison's cognitivist denial of the possibility of a schism between objective principles and subjective motivation. Richard McCarty follows suit in "Moral Weakness as Self-Deception". The objection has been echoed by Robert N. Johnson, "Weakness Incorporated," and Iain Morrisson, "On Kantian Maxims: A reconciliation of the Incorporation Thesis and Weakness of the Will."

Works Cited

Kant's works

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–). Unless otherwise stated, 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–).

Other Sources

Allison, Henry E. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 1991.

Baron, Marcia. "Freedom, Frailty, and Impurity." *Inquiry* 36, no. 4 (December 1, 1993): 431–41.

Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. Chicago: University Of Chicago Press, 1960.

—. "Kant on Education." In *Education in the 18th Century*, edited by J. D. Browning, 10–24. New York: Garland, 1979.

Broadie, Alexander, and Elizabeth M. Pybus. "Kant and Weakness of Will." *Kant-Studien* 73, no. 1–4 (1982): 406–412.

Coble, Kelly. "Kant's Dynamic Theory of Character." *Kantian Review* 7 (March 2003): 38–71.

Engstrom, Stephen. "Conditioned Autonomy." *Philosophy and Phenomenological Research* 48, no. 3 (1988): 435–453.

—. "The Triebfeder of Pure Practical Reason." In *Kant's Critique of Practical Reason*. Cambridge Critical Guides. Cambridge University Press, 2010.

Formosa, Paul. "From Discipline and Autonomy: Kant's Theory of Moral Development." In *Kant and Education: Interpretations and Commentary*, edited by Klas Roth and Chris W. Surprenant, 163–176. New York: Routledge, 2011.

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 University Press, 1996.

Herrera, Larry. "Kant on the Moral Triebfeder." *Kant-Studien* 91, no. 4 (2000): 395–410.

Hill, Thomas E. “Kant on Weakness of Will.” In *Virtue, Rules, and Justice: Kantian Aspirations*. Oxford: Oxford University Press, 2012.

Johnson, Robert N. “Weakness Incorporated.” *History of Philosophy Quarterly* 15, no. 3 (1998): 349–367.

Koch, Lutz. *Kants Ethische Didaktik*. Würzburg: Ergon, 2003.

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

—. *Self-Constitution : Agency, Identity, and Integrity*. Oxford ; New York: Oxford University Press, 2009.

Louden, Robert B. *Morality and Moral Theory: A Reappraisal and Reaffirmation*. New York: Oxford University Press, 1992.

McCarty, Richard. “Moral Weakness as Self-Deception.” *Proceedings of the Eighth International Kant Congress* 2 (1995): 587–593.

Moran, Kate. *Community and Progress in Kant’s Moral Philosophy*. Washington, D.C: Catholic University America Press, 2012.

Moran, Kate A. “Can Kant Have an Account of Moral Education?” *Journal of Philosophy of Education* 43, no. 4 (2009): 471–84.

Morrisson, Iain. “On Kantian Maxims: A Reconciliation of the Incorporation Thesis and Weakness of the Will.” *History of Philosophy Quarterly* 22, no. 1 (2005): 73–89.

Munzel, G. Felicitas. “‘Doctrine of Method’ and ‘Closing.’” In *Immanuel Kant: Kritik der praktischen Vernunft*, edited by Otfried Höffe, 151–63. De Gruyter, 2002.

—. *Kant’s Conception of Moral Character: The “Critical” Link of Morality, Anthropology, and Reflective Judgment*. Chicago: University of Chicago Press, 1999.

—. *Kant’S Conception of Pedagogy: Toward Education for Freedom*. Evanston, Ill: Northwestern University Press, 2012.

O’Connor, Daniel. “Good and Evil Disposition.” *Kant-Studien* 76 (1985): 288–302.

O’Neill, Onora. *Constructions of Reason: Explorations of Kant’s Practical Philosophy*. Cambridge University Press, 1989.

Peters, Julia. “Kant’s Gesinnung.” *Journal of the History of Philosophy* 56, no. 3 (2018): 497–518.

Roth, Klas, and Chris W. Surprenant, eds. *Kant and Education: Interpretations and Commentary*. 1 edition. New York: Routledge, 2011.

Schapiro, Tamar. “Kantian Rigorism and Mitigating Circumstances.” *Ethics* 117, no. 1 (2006): 32–57.

Sherman, Nancy. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge University Press, 1997.