

Kantian Moral Psychology and Human Weakness

Jessica Tizzard

University of Connecticut

© 2021 Jessica Tizzard

*This work is licensed under a Creative Commons
Attribution-NonCommercial-NoDerivatives 3.0 License.
<www.philosophersimprint.org/o21016/>*

FEW ISSUES REVEAL MORE about Immanuel Kant's moral psychology than his take on human frailty or weakness.¹ Interpreters differ widely on the topic, for it gets to the heart of a fundamental and controversial subject for Kant's philosophical system. To give a reading of Kant's view of weakness, one must delve deeply into the question of how our rational and sensible capacities relate to one another. For example, is all human action, even weak action, expressive of our capacity to reason and adopt practical principles for ourselves? Or do moments of weakness show that sense-based desire can interrupt our practical lives and direct our behavior? However Kant's reader answers such questions, they become immediately committed to a general understanding of how our most fundamental human activities, sensing and knowing, relate to and inform one another. Their stance will inform a basic conception of human moral psychology and the structure of motivation. It will even fix a conception of how we as humans relate to other sensing creatures and a non-sensing purely rational being like God.² In short, we cannot understand Kant's view of weakness without saying a lot about human nature and its place in the universe.

1. Kant himself uses the German equivalents of these terms interchangeably in the *Religion*, though frailty [*Gebrechlichkeit*] is used more frequently as his chosen technical term. Cf. R 6:29: "We can think of three different grades to this natural propensity to evil. First, it is the general weakness [*Schwäche*] of the human heart in complying with the adopted maxims, or the frailty [*Gebrechlichkeit*] of human nature". In the *Metaphysics of Morals*, he is more likely to use the term 'weakness' (cf. MS 6:408), but he does sometimes opt for 'frailty' as well (cf. MS 6:446). I will use 'frailty' and 'weakness' interchangeably, typically prioritizing 'weakness' to maintain a connection to the standard English terminology. Although, it should be noted that Kant means something broader in scope than what is traditionally understood by 'weakness of will', i.e., an agent knowingly acting against their judgment about what it would be good to do.
2. For example, if Kant's commentator concludes that weak action is to be explained by looking to sensibility and the realm of empirically determinable causes, our moral psychology will share much with animals who sense but do not reason. On the other hand, if they conclude that weak action is to be explained by looking to the self-determining activity of reason, our moral psychology will be very different from such creatures, sharing more in common with other rational beings such as God. Though of course, insofar as the latter is for Kant a purely rational being possessing a radically different "intuitive" intellect that is wholly spontaneous — unlike our "discursive" intellect which

How we approach the topic also determines Kant's relation to his early modern predecessors. It is widely taught that Kant's systematic vision of philosophy can be read as an attempt to unify the insights of early modern rationalism and empiricism.³ But beyond this general thought, the details are subject to debate. Does Kant's view represent an amalgamation of two different sets of principles? Does he favor the insights of one school over the other? How does he resolve the apparent conflict between them? Developing an interpretation of human weakness will speak directly to these details, for the degree to which one takes sensibility to influence human action independently of reason will establish the extent of Kant's debt to the empiricists. We can use Kant's infamous distinction between phenomena and noumena, the realm of sense and the realm of reason, to illustrate. If his interpreter adopts an understanding of this distinction on which human action in the phenomenal world, the realm of everyday experience, is fully determined by mechanistic laws of nature that describe how sensible desire moves us, then a large part of Kant's moral psychology will be effectively Humean in character.⁴ The noumenal world of freedom will be almost entirely severed from this picture, and the strict division between them will delineate itself in response to Kant's need to address two very different kinds of philosophical pressure, each on its own terms. The nature of human action, understood as phenomenon, is explained with an eye to empiricist-inspired questions about the nature and origin of particular desires and their ability to

relies on the deliverances or "matter" of sensibility (cf. B 135) — there will be significant differences.

3. For this widespread view, we need look no further than the opening sentences of the main article on Kant in the Stanford Encyclopedia of Philosophy: "Immanuel Kant (1724–1804) is the central figure in modern philosophy. He synthesized early modern rationalism and empiricism, set the terms for much of nineteenth and twentieth century philosophy, and continues to exercise a significant influence today ..." (Rohlf, "Immanuel Kant", n.p., emphasis mine).
4. For interpretations of this stripe, see McCarty, *Kant's Theory of Action*, pp. xxi–ii, and Guyer, *Knowledge, Reason, and Taste: Kant's Response to Hume*, p. 165.

physically move us.⁵ The same action understood as the product of a noumenal will is also explained in deference to more rationalistically minded questions about the possibility of unconditioned, spontaneous freedom, or self-determination according to laws of reason.⁶ Such a view would hold that Kant was inspired by both schools more or less equally, and that this influence expresses itself through an amalgamation of ostensibly conflicting principles made consistent through the adoption of transcendental idealism and the pride of place it gives to the phenomena/noumena distinction. This will, no doubt, sound like the standard reading of Kant handed down to many of us.

The aim of this paper is to offer a radically different understanding of Kant's moral psychology and its relation to the early modern tradition by focusing on the particular topic of moral weakness. Exploring some of the basic tenets of Kant's position, I will argue that we should think of Kant's philosophy as constituting a much more profound break with the empiricist tradition that woke him from his "dogmatic slumber" (P 4:460).⁷ In particular, with respect to the topic of moral psychology and the possibility of weakness, I will claim that

5. See, for example, Guyer, *Knowledge, Reason, and Taste*, p. 165: "[Kant] also assumes that some sort of what Hume would call an 'affection' must be the proximate phenomenal or empirical cause of any action".
6. Again, Guyer, *Knowledge, Reason, and Taste*, p. 182: "Kant holds that we must suppose we have freedom of the will at the noumenal level, where the causal determinism that holds throughout the phenomenal world and that might there seem sometimes to stand in the way of our being free to do as morality requires does not obtain".
7. Kant citations are given in standard notation: *Critique of Pure Reason* is cited using A and B edition page numbers, and all other works are cited using the volume and page numbers from the Academy Editions. Works are abbreviated as follows:
 A: *Anthropology from a Pragmatic Point of View*
 G: *Groundwork of the Metaphysics of Morals*
 KpV: *Critique of Practical Reason*
 KU: *Critique of the Power of Judgment*
 LE: *Lectures on Ethics*
 MS: *The Metaphysics of Morals*
 P: *Prolegomena to any future Metaphysics*
 R: *Religion within the boundaries of mere reason*
 TP: *On the common saying: That may be correct in theory, but it is of no use in practice*

we must rethink the role that empirically discoverable, sense-based desire plays in human life. The picture described above demonstrates an untenable *dualism* that understands reason and sensibility to be radically distinct capacities best understood without reference to one another. As I would have it, this dualism is nowhere present in Kant's account. Most importantly, he does not develop his theory of the sensible world and our sensible capacities to meet empiricist-driven concerns about the physical-mechanical source of human action. As I will argue, Kant's position should not be understood as an amalgamation of rationalist and empiricist principles. Though he institutes more strict limits on metaphysical speculation and attributes a bigger role to sense experience than his rationalist predecessors, the important contributions that sensibility makes should not be understood as floating free of reason's influence. The constitutive role that our feelings and sensations play in motivating action can only be understood with reference to reason, as determined by and transformed through that reason's activity.⁸ As we will see, reckoning with the full significance of this claim requires moving beyond the "Incorporation Thesis" that has dominated Kant interpretation for decades, fixing the interpreter's gaze on maxim-formation as the locus for interaction between reason and sense-based desire.⁹

Reason's complex interdependence with and transformative influence upon sensibility is seen most fundamentally in Kant's identification of the will with practical reason in the *Groundwork of the Metaphysics of Morals*. For Kant, the essential form of our volition is captured through the idea of acting according to the representation of principles or laws, which leads him to conclude that "the will is nothing

8. For further discussion of the interdependence of reason and sensibility, see Tizzard, "Kant on Space, Time, and Respect for the Moral Law as Analogous Formal Elements of Sensibility" and Boyle, "Additive Theories of Rationality".
9. The "Incorporation Thesis" is the keystone of Henry Allison's seminal work *Kant's Theory of Freedom* (cf. p. 5 for its first introduction). More detailed discussion will follow, but its basic claim is that sensible desires or incentives must be incorporated into a maxim produced by reason in order to determine action.

other than practical reason" (G 4:412). On the reading I will suggest, this implies that sensible desire can have only a limited kind of efficacy in our practical lives because it is ultimately practical reason that moves us.¹⁰ We must reject the idea that sensibility can account for forces measured independently from and potentially conflicting with reason, forces that sufficiently explain human action in the empirical realm.¹¹ This will lead to a non-traditional account of weakness, rejecting the possibility that a weak agent could conduct a complete rational assessment of the wrongness of their proposed action, and then just do it anyway because some independently construed sensible desire is strong enough to outweigh this reflection.

It is, however, equally true that the reading to follow will not retreat into an overly rationalistic conception of Kant's position, explaining away the possibility that sensible desire could disrupt our practical lives in any meaningful way. In this respect, my view will differ from those that emphasize — or, as it may be, overemphasize — the so-called "Incorporation Thesis", and I will direct our attention away from the agent's adopted maxims to focus on the way that sensible desires can make good moral action difficult *after* these maxims have already been formed. As Kant continues his explanation of the will in

10. I thus also reject interpretations of the Wille/Willkür distinction according to which Willkür (usually translated as 'choice' or 'power of choice') depends on a sufficiently powerful incentive to be actualized. Cf. McCarty, "Motivation and Moral Choice in Kant's Theory of Rational Agency," p. 19. A full defense of an alternative interpretation of Willkür is beyond the scope of this paper, but I would point out that Kant's main remarks about the concept from MS 6:213–4 are consistent with my claim that we cannot isolate sensibility from reason's influence in explaining human action. Cf. MS 6:213: "Human choice, however, is a choice that can indeed be *affected* but not *determined* by impulses, and is therefore of itself (apart from an acquired proficiency of reason) not pure but can still be determined to actions by pure will". That said, I will not rely on the Wille/Willkür distinction here, though I cite and will note those passages where Kant employs the latter term.
11. As I will argue, this is not to deny that we sometimes feel strong desires to act against the moral law. My point is rather that we cannot fully understand the role these desires have in influencing human action with a dualistic moral psychology that takes Kant's conception of sensibility and sensible desire to answer to empiricist concerns.

the *Groundwork*, he gives central importance to the fact of our sense-dependence and attendant fallibility (G 4:412–3). This fallibility is, for Kant, inseparably connected with the structure of our sense-dependent thinking, and it is incumbent upon his interpreter to explain how the interdependence of reason and sensibility constitutive of human willing can produce such flawed activity. As we will see, a non-traditional account of weakness that overlooks the influence of empirically based sensible desire — sensibility’s contribution to practical reasoning — will likewise fail to capture Kant’s view. To avoid the dualism at work in the standard reading of Kant and his inheritance of the early modern tradition, his reader must grasp that reason and sensibility are not capacities with separate and contradictory realms as their objects. Brought together in one being, our capacities to reason and sense-give rise to a complex, unified moral psychology that can both invoke the role that reason plays in determining human action, and appeal to strong sensible desires and their ability to disrupt our practical lives.

After introducing the textual evidence typically used to interpret Kant’s view of weakness (§1.1), I will analyze two prominent readings offered by Henry Allison and Richard McCarty (§1.2–1.3). Each represents a different pole on the spectrum of interpretation, ranging from Allison’s extreme rationalism, which focuses on maxim formation and holds that all action is determined by reason, to McCarty’s extreme empiricism, which holds that all action is determined by the strongest available affective force. To avoid the dualism inherent to these positions, I will argue that Kant’s interpreter must instead locate the source of weakness in agential activity that involves both general principles of reason *and* particular sensible desires (§2.1). This, I will argue, is the activity of practical judgment “by which what is said in the rule universally (*in abstracto*) is applied to an action *in concreto*” (KpV 5:67).¹² Locating the source of weakness in practical judgment accordingly shifts the locus of explanation. While McCarty invokes strong sensible desires and Allison turns to self-deception regarding one’s maxims

to explain the possibility of weakness, I will instead look to the difficulties inherent to the application of universal practical principles to complex sense-dependent situations characterized by intersecting desires, affects, and ends. This form of explanation assumes that the agent’s principles are well-formed in the abstract, shifting focus to the issue of putting them into practice. As I will show, that we as moral agents are susceptible to weakness in practical judgment is indicative of our nature as sense-dependent reasoners. It is neither the affective strength of sensibility nor reason’s voluntary adoption of evil maxims that best explains our frailty, but rather the challenges that attend the practical activity of beings who must organize sensibly-given particulars according to universal principles of reason. The weak agent is, accordingly, one who fails to do this well.

As with most interpreters, my account of weakness takes its initial cue from Kant’s remarks about the human propensity to evil in *Religion within the Boundaries of Mere Reason* (1793). I will begin by introducing the particular descriptions of weakness offered there, and provide crucial context by framing these descriptions in light of Kant’s broader account of the ultimate ground of moral evil as a noumenal deed (§1.1).¹³ Unpacking the latter involves discussing the agent’s intelligible or noumenal character as that activity of reason which ultimately explains the moral bearings of our particular actions. These key passages will serve as interpretive anchors that any satisfactory account of Kantian weakness must heed, but my interpretive aims will move beyond them, incorporating resources from the *Groundwork* (1785), the *Anthropology* (1798), and especially the *Metaphysics of Morals* (1797) (§2.1–2.2). Accordingly, my aim is not to provide a focused reading of the remarks on weakness in the *Religion*, but to offer a systematic Kantian explanation of moral weakness that draws from a number of texts. While the *Religion* raises and quickly answers the important orienting questions required to give a philosophical account of weakness, I will argue that

12. Cf. KU 5:179, A 7:199 for further descriptions of the power of judgment, and MS 6:411 for another description of practical judgment in particular.

13. I am grateful to an anonymous reviewer for emphasizing the importance of including Kant’s discussion of the noumenal deed.

we must use these other texts to fill out this general framework and capture the rich moral psychology it implicates.¹⁴

The most important resource invoked for this purpose will be Kant's remarks on practical judgment in the Doctrine of Virtue from the *Metaphysics of Morals*. I will introduce two key obstacles to good practical judgment discussed there in connection with moral weakness: i) the Problem of Indeterminacy (§2.2.1), and ii) the Problem of Affect (§2.2.2). Each of these obstacles invokes both reason and sensibility, describing some difficulty that arises for judgment in virtue of their interaction. While Indeterminacy highlights the difficulties inherent to applying universal practical principles to particular sensibly-given contexts, Affect highlights the difficulties inherent to rational beings whose sensations influence them to think and act in various ways. Moral weakness is displayed insofar as the agent fails in respect to these obstacles (§2.2.3). I will then show how the detailed account of weakness developed using the Doctrine of Virtue is consistent with Kant's view of evil in the *Religion* (§2.3). Insofar as his aim in the latter is to show the ultimate noumenal ground of evil in general, i.e., the deeper agential structure that explains the particular moral failings we exhibit, focusing on practical judgment to explain weakness proves a promising avenue of interpretation. In the *Religion*, Kant identifies this deeper ground of evil insofar as the agent follows the law of duty "also with an eye to other aims" (R 6:42). As I will show, exhibiting weakness in practical judgment is the least culpable form of this evil, demonstrated simply insofar as the agent with good maxims *in abstracto* struggles to apply these principles *in concreto* because of the difficulties of indeterminacy and affect. A deep truth about the nature of our agency lurks here: because our sense-dependence entails that we cannot apply our moral principles except in response to sensibly-given content that naturally gives rise to other aims — namely, the aims of

14. A complete account of Kant's moral psychology will ultimately require a return to the *Religion* and especially Kant's discussion of the ethical community and historical faith in Part III, which is key to developing Kant's view of moral education. Though I do not have the space to deal with this issue in the present work, I am grateful to an anonymous reviewer for emphasizing this point.

happiness — we cannot avoid these difficulties. Kant's highly complex account of sense-dependent practical reasoning thus reveals weakness to be an inescapable part of human nature against which we must perennially struggle.

1. Empiricist vs. Rationalist Interpretations of Kantian Weakness

1.1 Textual Foundations

Kant's reader can find discussions of human weakness in a number of his published works, including *Religion within the Boundaries of mere Reason*, the *Metaphysics of Morals*, and *Anthropology from a Pragmatic Point of View*, as well as in student notes on his *Lectures on Ethics*.¹⁵ Most of these remarks are, however, engaged with particular empirical examples, and so are insufficient to determine the details of the broader account of weakness I aim to develop here.¹⁶ To approach the latter, commentators typically look to the first part of the *Religion*, where Kant introduces three grades of the human propensity to evil.¹⁷ The differences between these grades correspond roughly to the role that sense-based desire plays in each of them. That is, we distinguish between them based on the extent to which sense-based desire has been incorporated into the will as its determining ground. The first

15. Cf. LE 27:293–5, 27:349–51; 27:570–2, 27:605; 27:624; 27:701.

16. For a representative example, consider A 7:260: "It is true that hot temper can be diminished through inner discipline of the mind; but the **weakness** of an extremely delicate feeling of honor that manifests itself in shame does not allow itself to be removed so easily" (bold emphasis mine).

17. Kant defines a propensity (*Hang*) as "the subjective ground of the possibility of an inclination (habitual desire, *concupiscentia*), insofar as this possibility is contingent for humanity in general" (R 6:29). This is contrasted with a "predisposition" (*Anlage*), which is more deeply rooted: "by the predispositions of a being we understand the constituent parts required for it as well as the forms of their combination that make for such a being" (R 6:28). Humans have an ineradicable predisposition towards the good that makes our very nature as living moral beings possible. Our "propensity to evil" is, in contrast, the part of our subjectivity that accounts for the contingent (but equally ineradicable) possibility that we might fail to act from the moral law. Because propensities enable activity that is merely contingent, Kant thinks we can be held responsible for them. It is our agency that makes the difference as to whether they are actualized or not (R 6:29).

and most innocent of the three grades of evil is “the general weakness [*Schwäche*] of the human heart in complying with the adopted maxims, or the frailty [*Gebrechlichkeit*] of human nature” (R 6:29). Kant further unpacks this propensity in the passage directly following:

the frailty (*fragilitas*) of human nature is expressed even in the complaint of an Apostle: ‘What I would, that I do not!’ i.e., I incorporate the good (the law) into the maxim of my power of choice [*Willkür*]; but this good, which is an irresistible incentive objectively or ideally (*in thesi*), is subjectively (*in hypothesi*) the weaker (in comparison with inclination) whenever the maxim is to be followed. (R 6:29)

With frailty, the conflict or perversity characteristic of evil in general shows itself at the surface level: the agent who is interested in moral goodness acts on inclination — habitual sensible desire — instead. Despite this failure, however, the agent is still represented as being resolutely moral in their commitments. With impurity and depravity — the second and third more entrenched forms of evil — commitment to morality is supplemented with or entirely subordinated to sensible inclination (R 6:30). In the primary stage of evil, however, the weak agent understands themselves to recognize the moral law as the only legitimate determining ground of action, though they ultimately fail to act on this ground. As Kant’s first description suggests, at issue is not the character of the maxims the agent has adopted, but the matter of complying with them. He will sometimes even speak of “the frailty of human nature” in general, as the condition of “not being strong enough to comply with its adopted principles” (R 6:37).

This account is further complicated by Kant’s general discussion of the human propensity to evil, which draws heavily on the noumena/phenomena distinction as applied to moral agents. To explain how moral character can be responsible for weak, impure, or depraved actions, Kant identifies two different meanings of the term ‘deed’. While the phenomenal deed refers to the particular actions we can

empirically observe, the noumenal deed refers to the activity of reason that grounds or explains the former insofar as it is an act of willing. As that which explains the possibility of particular evil actions, the propensity to evil is a deed in this noumenal sense. Kant describes the propensity to evil in a number of ways:

- i) as “an intelligible deed, cognizable through reason alone apart from any temporal condition” (R 6:31);
- ii) as “the formal ground of every deed contrary to law according to the second [phenomenal] meaning” (R 6:31);
- iii) as “the use of freedom through which the supreme maxim ... is adopted in the power of choice [*Willkür*], as to the use by which the actions themselves (materially considered, i.e. as regards the objects of the power of choice) are performed in accordance with that maxim” (R 6:31).

With these remarks, Kant is attempting to show how explanations of human action can be traced to the morally relevant grounds that show such action to be imputable to the agent. To account for any particular phenomenal action, we must look beyond the empirical realm of contingent sensible inclination to consider the more foundational activity that shows this action to be freely adopted. As Kant says above, this activity is intelligible and atemporal, an act of reason that transcends the vicissitudes of the empirical realm and accounts for the “form” or basic character of the phenomenal action as an action with moral relevance that is either good or evil. As the third passage indicates, this rational deed can be expressed in terms of the adoption of a *supreme maxim* to prioritize either the moral law or self-love, which maxim provides the ultimate explanation for how the agent’s power of choice derives and performs particular actions. As the root explanation for our moral fallibility, Kant takes the propensity to evil to be innate and ineradicable (R 6:31). We are not necessitated to evil — this would, of course, destroy our freedom — but all human beings are susceptible to acting

on sensible inclinations, and we cannot escape this basic feature of our agency.

In keeping with Kant's identification of the propensity to evil as a noumenal deed, a satisfactory account of Kantian weakness should therefore explain how a particular case of weak action is to be grounded in the agent's more basic noumenal character. As Kant himself insists after canvassing the two types of deed:

in this section, from the very start, we sought the three sources of moral evil solely in that which affects the ultimate ground for the acceptance or observance of our maxims according to laws of freedom, not in what affects sensibility (as receptivity). (R 6:32)

This presents a particular difficulty in the case of weakness, since in the very same text, Kant depicts the weak agent as committed to their recognition that the moral law is the only appropriate determining ground of the will. With the impure or depraved agent, the propensity to evil is invoked to explain how the agent comes to endorse maxims that incorporate self-love into the determining ground of the will. But Kant's remarks on weakness in the *Religion* do not provide *prima facie* grounds for thinking that this is how to explain the first grade of evil. In this case, the agent incorporates the moral law into their maxim as the supreme determining ground, but produces particular actions that do not accord with this law. This raises the crucial question: how should we invoke the propensity to evil to explain individual weak actions, given that the weak agent is generally committed to morality?¹⁸

18. In his recent commentary on the *Religion*, Lawrence Pasternack shows that one way to deal with this question is to deny that the merely weak agent is less evil than the depraved one. On his reading, weakness also involves self-deception and the subordination of morality to self-interest, so it is fundamentally just as bad. (Pasternack, *Kant on Religion within the Boundaries of Mere Reason*, p. 119.) This view is admirable in its consistency, and thereby avoids the pitfall of unwittingly collapsing frailty into depravity that is present in other accounts. I will, however, advocate for a different reading of weakness, upon which we can say that it *is* the least severe stage of evil because the agent's basic moral orientation does not subordinate morality to self-interest.

Unfortunately for his interpreter, Kant does not delve any deeper into his account of weakness in the *Religion*.¹⁹ It is not surprising, then, that interpretations vary wildly. Given how little Kant actually says about weakness, commentators typically rely on the aforementioned big-picture commitments concerning the relation between reason and sensibility.

1.2 Two Well-Known Examples

For example, Richard McCarty's view of Kantian moral psychology invokes an empiricist version of dualism to explain human action. As I am using these terms, this means he holds reason and sensibility to be radically distinct capacities calling for independent accounts, with explanatory privilege given to the sensible part of human agency. These commitments enable him to comfortably handle the schism between reason and sensibility that cases of weakness seem to present. Because he already endorses a quasi-Humean, mechanical understanding of sensibility and its role in motivating action, moral weakness is for him just a prime illustration of the fact that the possibility of action depends on a sufficiently powerful affective force. McCarty thus fully embraces the model of the agent who acts against their knowledge: for him, "the pure in heart but frail ... are those who intellectually acknowledge the supreme authority of the moral law but may nevertheless choose,

19. Though there are numerous references to human weakness in the subsequent parts of the *Religion*, they are not particularly helpful for my purposes here. Most of these passages simply reference the fact of human weakness, taking it for granted in order to discuss how it manifests itself in religious practice (cf. R 6:103, 141, 169, 191). Some passages arguably stretch the term to a more general use that goes beyond the particular feature of moral psychology under examination here, e.g., when Kant speaks of that aspect of prayer which "weakens the effect of the moral idea" (R 6:197; cf. also R 6:43, 103). Though I think all such passages are ultimately related to Kant's general view of human weakness as I am seeking to develop it, they are not the most helpful places to look in order to arrive at this view. That is to say, how one reads these passages will depend on how one answers the more abstract questions about the nature of reason and sensibility posed at the outset of this paper, which are my concern here. I am grateful to an anonymous reviewer for pointing out that these passages should be addressed for the sake of accuracy and completeness.

through weakness, to act on opposing inclinations".²⁰ In a series of papers dealing with the phenomenon of weakness, McCarty takes Kant's commitment to an opposing-forces view of motivation to be obvious. What ultimately explains human action is the "oomph" attached to the stronger incentive that overpowers its competitors: "if that moral incentive is stronger than competing, non-moral incentives, it will subsequently determine the outcome of the moral-choice event".²¹ On this view, it is perfectly conceivable for the weak moral agent to rationally endorse the law and be moved by this endorsement, but ultimately be overpowered by an independently acting non-moral incentive.

As McCarty clarifies elsewhere, this view of moral psychology is firmly grounded in a "two-worlds" interpretation of the noumena/phenomena distinction that invokes the actual existence of two separate worlds, each of which accounts for part of our character as human agents.²² On the one hand, human action exists in the empirical world and is subject to psychological determining forces beyond our control. These forces are completely sufficient to explain what we do in the context of our sensible existence. On the other hand, to account for the moral responsibility we bear for such actions, we must also posit their adoption in a separately existing noumenal world that enjoys a unilateral grounding relation to the phenomenal world. This is, on McCarty's view, the only way Kant can adequately account for free but empirically explainable action: "we simply must presume this evil character has been freely chosen by every human moral agent in, as it must appear to us, a prior, or *noumenal* choice".²³ The noumenal deed or supreme maxim thus plays two important functions for McCarty's Kant: i) it accounts for the imputability or moral significance of the

20. McCarty, "Kantian Moral Motivation and the Feeling of Respect", p. 431. See also, McCarty, *Kant's Theory of Action*; McCarty, "Moral Weakness as Self-Deception"; and McCarty, "Motivation and Moral Choice".

21. McCarty, "Motivation and Moral Choice", p. 26.

22. McCarty, *Kant's Theory of Action*, chap. 4, p. 105.

23. McCarty, "Motivation and Moral Choice", p. 28.

action, and ii) it does so by expressing the agent's voluntary endorsement of evil actions due to weakness. As he concludes,

it is entirely appropriate to impute to moral agents any transgression due to moral weakness, even though there remains a sense in which we may say their transgressions depend on a factor beyond their immediate control: namely, the contingent strength of their capacity for moral feeling.²⁴

So although an appeal to deterministically construed affective psychological forces does most of the explanatory work in McCarty's picture, his "two-worlds" view of human agency leaves room for freedom.

While McCarty's emphasis on the contingent strength of affective forces constitutes the empiricist pole of the dualist interpretive spectrum, Henry Allison's emphasis on maxim-formation marks the rationalist one. Again, with Allison's view, we find that reason and sensibility are viewed as independent capacities calling for separate explanations, this time with the conviction that the contributions of sensibility cannot, as such, explain human action. This is expressed through Allison's commitment to the well-known "Incorporation Thesis",²⁵ which

24. Ibid.

25. Allison, *Kant's Theory of Freedom*, p. 5. The central passage he draws upon is R 6:24:

freedom of the power of choice [*Willkür*] has the characteristic, entirely peculiar to it, that it cannot be determined to action through any incentive *except so far as the human being has incorporated it into his maxim* (has made it into a universal rule for himself, according to which he wills to conduct himself); only in this way can an incentive, whatever it may be, coexist with the absolute spontaneity of the power of choice [*Willkür*] (of freedom).

This passage may seem to settle the matter definitively for Allison's camp, but Kant makes remarks elsewhere that suggest there is more to the Kantian picture. For example, when discussing the difference between a pure and impure commitment to morality, Kant insists that "a human being, who incorporates this purity into his maxims, though on this account still not holy as such (*for between maxim and deed there still is a wide gap*), is nonetheless upon

holds that sensible incentives must be incorporated into maxims for action in order to move the agent, thus giving priority to the rational activity of maxim formation in the explanation of human action. This rationalist emphasis fixes much of Allison's view on weakness, ruling out the traditional model on which the agent is fully committed to and aware of the moral constraints in play but lacks the motivational strength to carry out the required action. Because, according to Allison's view, sensible incentives must ultimately be rationally endorsed by the agent if they are to be acted upon, there must be some sense in which the weak agent's maxims do not demonstrate a total commitment to morality. The weak agent's claim to have adopted the moral law as their supreme maxim is, therefore, the result of self-deception:

the so-called lack of sufficient strength to follow moral principles when they conflict with the claims of inclination reflects the lack of a full commitment to these principles in the first place. Thus, self-deception enters the picture at the very beginning, depicting what is in reality a free evaluation on one's part as a 'weakness' for which one is not responsible.²⁶

So on this strongly rationalist interpretation, all action must reflect one's considered judgment. Though the subject might consciously disavow or unconsciously distort their reasons for acting, we must conclude that their actions are indicative of a commitment to morality that, while perhaps earnest in its striving, admits of exceptions.

Returning to the distinction between noumenal and phenomenal deeds, it is natural to read Allison's explanation as drawing heavily on the former. He equates weakness with "the bare propensity to evil itself", so that "the susceptibility to temptation, now characterized as

the road of endless progress towards holiness" (R 6:46–7; emphasis mine). As he suggests with the parenthetical remark, one can have the correct maxims in place and still fail to do what the law requires. This is just what it is to act weakly. So as I will argue, to explain weakness, we should not look to the agent's maxim, but elsewhere.

26. Allison, *Kant's Theory of Freedom*, p. 159.

weakness, is the direct result of the original primacy granted to the claims of one's sensuous nature".²⁷ It is this intelligible deed which explains the character of weak action and, Allison thinks, necessitates an appeal to self-deception. Insofar as the propensity to evil is construed as the voluntaristic determination to give the inclinations primacy, the agent cannot really be fully committed to morality. This is summed up in a particularly helpful passage that hits upon many of the key points required to understand Allison's view:

what was initially regarded as straightforward weakness in the face of temptation turned out, on further analysis, to be an expression of freedom, more specifically, an evaluation placed on certain ends of inclination. Consequently, if as free agents we are tempted, it is only because we, as it were, allow ourselves to be. In Kant's model of rational agency (the Incorporation Thesis) 'yielding' to temptation is not to be conceived as being overcome by a superior psychic force but rather as a kind of inner voting in favor of certain ends, a taking of these ends as one's own.²⁸

McCarty's prized empirical model of explanation is, on Allison's view, thus only helpful as a welcome illusion that allows us to forgo some responsibility for our actions. Moreover, what makes this "inner voting" merely weak as opposed to depraved is also explained by appeal to the noumenal deed. Though Allison's commitment to the Incorporation Thesis leads him to conclude that the weak agent does act on immoral maxims, he tempers this claim by recognizing Kant's distinction between a good will and an evil heart in R 6:37. On Allison's interpretation, the propensity to evil or "evil heart" grants primacy to the inclinations, but the will is still "in general good" — as opposed to wholly good — insofar as it is genuinely committed to the struggle towards

27. Allison, *Kant's Theory of Freedom*, p. 159.

28. *Ibid.*, p. 164.

virtue.²⁹ This latter commitment is thought to be an inextirpable aspect of the personality of the moral agent that must be assigned to their intelligible character, and invoked to accurately account for any human action.³⁰

1.3 *Difficulties with Dualism*

Despite the significant differences between these two accounts, both suffer interpretive difficulties that stem from a common root insofar as each isolates reason from sensibility in their respective explanations of human weakness. These difficulties map onto the distinction between noumenal and phenomenal deeds: a dualistic understanding of reason as the source of the former lays root to complications in explaining how this noumenal deed grounds the phenomenal actions of a sensibly-affected empirical agent. Bracketing the relevant differences between these two views, we find that both McCarty and Allison view the noumenal deed constitutive of weak action as some voluntary choice or commitment to prioritize the inclinations. They both agree that without this foundation, the resultant action could not be seen as free and imputable, i.e., as action in the robust moral sense of the term. I call this a “dualistic” understanding of the noumenal deed because reason is here isolated from sensibility and prioritized in the explanation of its function within Kant’s system. Though of course only a sensibly-affected being could pursue happiness instead of morality, McCarty and Allison cast the ground of this possibility in terms of pure practical reason — the will — making the clear-eyed choice to prioritize one’s happiness. So while the capacity to be sensibly affected must be presupposed if there is to be a choice between good and evil, the choice

29. *Ibid.*, p. 160. The passage Allison is drawing on says that “*Dieses [böse Herz] kann mit einem allgemeinen guten Willen zusammen bestehen*” (R 6:37). There is some disagreement about the best way to translate *allgemeinen* here. While Allison settles on “general”, di Giovanni prefers “in the abstract”, rendering the clause as follows: “an evil heart can coexist with a will which in the abstract is good”. As I will go on to argue, di Giovanni’s translation more accurately captures Kant’s view on the matter.

30. Allison, *Kant’s Theory of Freedom*, pp. 160–1.

itself is construed as one made by reason functioning in isolation from sensibility. One might think this is simply an interpretive necessity: as an atemporal activity attributed to our rational faculty, the noumenal deed should be explained by reference to reason alone. While this may seem an uncontroversial, textually well-supported interpretation, I would contend it is a questionable one that leads to insurmountable difficulties. Insofar as the noumenal activity of reason constitutive of weakness is construed as the clear-eyed choice to pursue happiness, there will inevitably be problems in specifying a grounding relation between this choice and the phenomenal actions it is meant to explain. These difficulties will differ accordingly as one chooses the rationalist or empiricist species of dualism to explain particular phenomenal actions and their relation to this foundational noumenal choice.

Consider, first, McCarty’s empiricist dualism. His emphasis on the affective forces constitutive of the sensible world leads him to completely isolate our noumenal actions from our phenomenal ones and conclude that all incentives — even the incentive to be moral — operate in a closed economy of psychological forces whose general thrust is indicative of our empirical character. Because the intellectual recognition of morality attributed to the noumenal deed, and the empirical motivational strength that explains the phenomenal deed have been so completely separated, the occurrence of weakness is effectively chalked up to dispositional luck: “no matter how strong the moral incentive is in us, dispositionally, circumstances can arise where competing incentives of inclination just turn out to be stronger than our feeling of respect for the moral law”.³¹ When it comes to explaining weakness, this leads to the strange interpretive stance that the weak agent is both i) morally responsible for the fact that the moral incentive happens to be too weak, because they have endorsed this weakness through a noumenal deed, and ii) unable to directly improve their moral situation, since the most important explanatory factor for

31. McCarty, “Motivation and Moral Choice”, p. 26.

phenomenal action is the affective strength of their incentives, a contingent empirical matter that lies beyond their control.

This makes for a shockingly precarious view of virtue, even by Kant's lights.³² Since McCarty is insistent that rational recognition is only contingently related to motive force, moral education and the cultivation of virtue cannot be a predominantly cognitive process. It must instead involve training psychological responses so as to strengthen one's ability to withstand temptation from inclination. Leaving aside the question of whether one wants to make room for the latter phenomenon within the Kantian picture, I would argue that it cannot stand as his model for virtue. Such a view cannot fit with Kant's broadly cognitivist commitments. Consider the following passage from the *Metaphysics of Morals*:

but virtue is not to be defined and valued merely as an *aptitude* and ... a long-standing habit of morally good actions acquired by practice. For unless this aptitude results from considered, firm, and continually purified principles, then, like any other mechanism of technically practical reason, it is neither armed for all situations nor adequately secured against the changes that new temptations could bring about. (MS 6:383–4).

I take passages like this one to require a much tighter relationship between our rational and sensible capacities than McCarty makes room for. Kant explicitly denies that virtue could result from the habitual work of instilling behavior without reflecting upon the rational considerations that should be non-accidentally motivating it. Placing such

32. Kant's own conception of virtue is itself precarious in a different sense. Virtue, for the human being, is still described as "moral disposition in conflict, and not holiness in the supposed possession of a complete purity of dispositions of the will" (KpV 5:84). Some commentators take this to be the most important point of difference between Kant and Aristotle's respective accounts, marking the philosophical impasse for those who would like to narrow the divide between their views and place less emphasis on the deontology/virtue ethics distinction. For an extensive account of this comparison, see Baxley, *Kant's Theory of Virtue: The Value of Autocracy*.

a large gap between reason and sensibility, the empiricist is thereby unable to articulate a properly Kantian account of moral life.

The rationalist dualism, on the other hand, allows for a similarly fatal gap, developing the account too much on the side of reason. Allison's view is the paradigmatic example: he prioritizes the Incorporation Thesis so much that there appears no viable alternative but to deny the traditional conception of weakness and conclude that all action reflects our considered principles. Sensibility can, in effect, only influence reason by transforming into something rational—namely, a practical principle endorsed by the subject. So with Allison, too, we see reason and sensibility held at arm's length from one another. While this account fares better when accounting for the way our rational commitments non-accidentally explain particular phenomenal actions, it encounters critical difficulties in accounting for weakness. For the rationalist version of dualism allows reason to completely overshadow sensibility and the way it shapes practical life.

This, too, leads to an untenable interpretation of Kantian moral psychology, for Kant himself is not one to ignore or reductively explain away the contributions that sensibility makes. Consider, for example, his reflections on affect from the *Metaphysics of Morals* and *Anthropology from a Pragmatic Point of View*: "affect is surprise through sensation, by means of which the mind's composure is suspended. Affect is therefore rash, that is, it quickly grows to a degree of feeling that makes reflection impossible (it is thoughtless)" (A 7:252). It is difficult to see how Allison could make sense of affect without somehow invoking self-deception and the Incorporation Thesis. If, for example, an agent is momentarily gripped by the affect of anger and responds disrespectfully to a colleague, Allison would claim that the affected agent rationally endorses acting in this way through a process of inner voting, and is mistaken in thinking that their anger overwhelmed them to influence what they did. Kant, however, explicitly denies that this is how to explain action influenced by affect in the *Anthropology*, contrasting affects with passions and claiming that only the latter can be explained by appealing to one's maxims:

since passions can be paired with the calmest reflection, it is easy to see that they are not thoughtless, like affects, or stormy and transitory; rather, they take root and can even co-exist with rationalizing. ... Passion always presupposes a maxim on the part of the subject, to act according to an end prescribed to him by his inclination. (A 7:266)

Kant therefore distinguishes affect from passion insofar as the former cannot, like the latter, express itself through desires that get incorporated into maxims. This suggests that at least some agential failures cannot be accounted for by appealing to maxim-formation. Now, we can ask: is this kind of failure an instance of weakness? I will argue in the following section that it is. The rationalist dualism I have attributed to Allison thus cannot effectively show how weak actions are grounded in the noumenal deed as he conceives it, because it reduces all explanations of agential failure to explanations about maxim formation, denying that sensibility can influence reason in other ways. Distinguishing the weak agent from the depraved one insofar as the former is “generally good”, as Allison does, will not help here, for this distinction still bottoms out in an understanding of the propensity to evil that equates it to the formation of immoral maxims. Such an account might capture some forms of agential failure, but it is not able to explain the kind of failures described in Kant’s passages on weakness and affect. To view Kant’s moral psychology in all of its nuance and complexity, we must look beyond this overly rationalist picture and the dualism that underlies it.³³

33. Though I have not discussed them here, other well-cited accounts of weakness also fit into this rationalist vs. empiricist framework. Consider those offered by Iain Morriison and Robert N. Johnson. Both attempt to square the possibility of weakness with Allison’s Incorporation Thesis, though with differing results. Morriison takes a more resolutely rationalist route and argues that the weak-willed subject acts “somewhat rationally”, choosing to pursue principled ends that have been deemed simply or immediately good, but do not figure in the totality of ends that constitute their vision of happiness. Weak action thus constitutes a failing to the extent that it is short-sighted but is nevertheless guided by principles that take the weak action as their object.

2. Rethinking Weakness: Invoking Practical Judgment

2.1 Initial Support for a Judgment-Based Account of Weakness

I have now diagnosed the interpretive shortcomings of prominent interpretations of Kant’s account of weakness. These have all been characterized as stemming from the assumption of dualism, i.e., the assumption that reason and sensibility are understood as isolated capacities best explained without reference to one another. As I have claimed, this will result in either something like McCarty’s view, which posits only a contingent and highly precarious connection between practical reasoning and what we desire to do; or something like Allison’s view, which insists that desire can only influence practical reasoning by being transformed into something *other* than a mere desire, namely, a general maxim for action. The assumption of a dualistic separation between reason and sensibility thus leads, in fact, to an impoverished kind of monism about moral psychology. Barring any conception of how these capacities might interact or depend on one another, one is compelled to choose between them, and explain action by appeal to either reason or sensibility. Neither view has the room for any kind of substantive and meaningful interaction between general principles (of reason) and particular desires (of sensibility). Tying this general analysis of Kant’s moral psychology to the phenomenon of weakness is now especially pressing, for it seems that we must appeal to exactly this kind of interaction in order to explain weakness. The fact that existing frameworks are unable to do so is, I would argue,

Cf. Morriison, “On Kantian Maxims: A reconciliation of the Incorporation Thesis and Weakness of the Will”. In contrast, Johnson offers an empiricist-friendly interpretation that holds an atypically expansive understanding of maxims: on Johnson’s view, some maxims motivate without justifying, making room for the possibility of irrational action that intentionally and voluntarily pursues the bad on the basis of strong desire. Cf. Johnson, “Weakness Incorporated”. It is also worth mentioning a seminal account of Kantian weakness published by Alexander Broadie and Elizabeth Pybus in 1982 (“Kant and Weakness of Will”). Broadie and Pybus offer a rationalist take on the matter, predating and sharing much with Allison’s account by arguing that we should explain weakness largely through appeals to self-deception and rationalization that cover up the immoral maxims the agent is actually acting upon.

strong evidence that we need to adopt a different conception of the relation between reason and sensibility to explain weakness, or, more specifically, to understand the relationship between the maxims we formulate and the sense-based desires that press upon us.

I suggested above that Kant's account of weakness is better understood as focused around the activity of practical judgment — “by which what is said in the rule universally (*in abstracto*) is applied to an action *in concreto*” (KpV 5:67) — rather than unruly sensible desires or maxim-formation, and it is to this thesis that I now turn. We can already see some *prima facie* evidence that this is a promising avenue of interpretation. Through analysis of the failings of other possible views, a few interpretive goals can be specified. First, drawing on Kant's account of the noumenal deed, the ultimate explanation of human action must be rooted in rational, as opposed to merely sensible, activity. But second, this rational activity cannot be dualistically construed. That is, our account of this rational activity cannot be completely isolated from the contributions of sensibility and framed as the clear-eyed rational choice to will the pursuit of happiness over morality. This implies, third, that a proper account of the noumenal deed that accounts for weak action must invoke the way sensibility influences or shapes our rational activity, without suggesting that it can determine or overpower it. I aim to show that understanding moral weakness as a failure of practical judgment — that is, a failure to apply an abstract universal principle *in concreto* — can satisfy all three conditions.

Before looking to other texts to develop Kant's account of practical judgment, it is worth looking to some textual support from the *Religion* itself. Though Kant does not invoke the power of judgment (*Urteilskraft*) explicitly in the explanation of weakness, he does describe the cognitive activity that is elsewhere attributed to it. As I have shown, Kant describes weakness as “the general weakness of the human heart in *complying with the adopted maxims*” (R 6:29, emphasis mine), suggesting that the issue is not bad maxims, but some failure in their application. When discussing the noumenal deed, he also marks a distinction between “the ultimate ground for the acceptance *or observance* of

our maxims” (R 6:32, emphasis mine), which can be read as implying that when accounting for various forms of evil, we must distinguish the ultimate ground for failing to *observe* adopted maxims from other grounds of moral failure involving the adoption of maxims themselves. Kant also makes room for the further activity of judgment when he acknowledges that properly oriented maxims are not sufficient for holiness in the human being, “for between maxim and deed there still is a wide gap” (R 6:47). Finally, in his longest remark on weakness, he says the following about the moral law: “this good, which is an irresistible incentive objectively or ideally (*in thesi*), is subjectively (*in hypothesi*) the weaker (in comparison with inclination) whenever the maxim is to be followed” (R 6:29). With some further analysis, the parenthetical Latin remarks are telling. Kant also uses them in his essay “On the common saying: That may be correct in theory, but it is of no use in practice”, where he clarifies that the saying “what sounds good in theory has no validity for practice ... is often expressed as, this or that proposition does indeed hold *in thesi*, but not *in hypothesi*” (TP 8:276). Kant uses these terms again in the third *Critique*, when discussing how enlightenment requires one to think for oneself: “one readily sees that while enlightenment is easy *in thesi*, *in hypothesi* it is a difficult matter that can only be accomplished slowly” (KU 5:294). Given the way Kant uses these Latin terms to indicate a distinction between abstract theory and concrete practice, we can read the above passage from the *Religion* as further indication that weakness arises from a failure in the *practice* of judgment, i.e., the failure to correctly apply maxims prioritizing the moral law to particular situations.³⁴

An initial worry can arise here insofar as my likening weakness to a failure of good principle application seems overly theoretical and

34. Not all commentators read these terms as I have. Stephen Palmquist, for example, reads ‘*in hypothesi*’ more literally, as part of the evidence that “Kant treats the evil propensity as a hypothesis throughout his exposition” (Palmquist, *Comprehensive Commentary on Kant's Religion within the Bounds of Bare Reason*, p. 77). I am grateful to Lionel Shapiro for pointing out that these Latin terms have a distinctive usage that supports my interpretation of Kant's view of weakness.

therefore not able to capture the unique features of moral weakness. The latter should involve an appeal to desire and motivation, or perhaps a lack thereof. A proper understanding of how Kant uses the term ‘practical judgment’ should, however, make room for the desiderative or motivational element required to explain moral weakness. Consider a helpful remark from the preface to the *Groundwork*. Here, Kant affirms that *a priori* laws of pure practical reason

no doubt still require a judgment sharpened by experience, partly to distinguish in what cases they are applicable and partly to provide them with access to the will of the human being and efficacy for his fulfillment of them; for the human being is affected by so many inclinations that, though capable of the idea of a practical pure reason, he is not so easily able to make it effective *in concreto* in the conduct of his life. (G 4:389)

As the passage suggests, Kant ascribes a twofold purpose to the cultivation of moral judgment: first, it prepares us for dealing with concrete, in-the-moment situations where the constraints of morality are felt. To grasp the applicability of moral conditions in the hurly-burly of practical life requires not just an abstract grasp of the law’s command to universal maxims, but the ability to read the situation, to differentiate particular incentives of self-love from those of morality, and to deal with unforeseen features that complicate the circumstances. Second, a cultivated sense of judgment grants more efficacy to the moral law, making it such as to actually move the subject to act on its basis. These may seem like separate aims, but Kant’s account of practical reason gives us the resources to see that they are in fact two sides of the same coin, representative of the objective and subjective aspects of the will’s determination. As Kant emphasizes when he equates the will with practical reason (G 4:412), practical rational representations are the source of their own efficacy. If reason is to be practical, it must be the act of representing itself that brings about its object. Practical representations are thus more or less efficacious for the subject on the

basis of their cognitive content, not some other measure of volitional strength that holds independently of this content.³⁵ So one cannot obtain a more determinate grasp of the objective constraints of morality without also increasing the efficacy that the representation of these constraints has for the subject. In short, objective cognitive content and subjective efficacy cannot be separated; to develop one is to develop the other. When one gains experience and learns how to better judge the situation, working out a more determinate picture of one’s moral commitments, their implications, and how they can be tested, this cognitive achievement strengthens the motivating power of the commitments themselves.³⁶ Practical judgment, for Kant, is thus better grasped as the species of judgment that moves the agent to act, rather than an inert judgment about matters of action.

We now have a preliminary account of how the cultivation of judgment could build virtue and thereby overcome moral weakness, as well as some textual evidence that the activity of judgment is implicated in Kant’s more direct account of the latter in the *Religion*. I have two remaining aims. First, Kant’s account of practical judgment should be filled out in more detail to show how failing to exhibit good judgment can result in weak action, taken in the phenomenal sense of the term. To do this, I will isolate two interrelated obstacles to practical judgment that Kant discusses in the *Metaphysics of Morals*. Second, this more detailed account should be examined in light of Kant’s remarks about the noumenal deed to show how understanding weakness as a

35. This thought can be expressed in well-known Kantian terminology by saying that it is the form and not the matter of a principle which should move us. That is, we should act based on the recognition of certain characteristics of our practical principles and the actions they describe — e.g., their universality, or the fact that they do not treat someone else as a mere means to an end — rather than the effects these actions bring about in the world (cf. KpV 5:27).

36. I take this to be an important upshot of the non-dualistic moral psychology I am attributing to Kant. In her recent book, Melissa Merritt also draws attention to this aspect of Kant’s view (*Kant on Reflection and Virtue*, p. 187). Her account of Kantian moral virtue is the only other published view I am aware of that develops a non-dualistic reading of Kant’s moral psychology.

failure of practical judgment gives the best account of the noumenal ground for these particular empirical examples of weakness.

2.2 *The Obstacles to Good Practical Judgment Constitutive of Weakness*

The interrelated obstacles to practical judgment that will be my focus in this section arise most prominently in the Introduction to the Doctrine of Virtue in the *Metaphysics of Morals*. As we will see, they are the Problem of Indeterminacy and the Problem of Affect. These are the two most basic obstacles to sound practical judgment, because they capture the difficulties inherent to the two-stemmed nature of sense-dependent practical reason. On the one hand, our *rational* nature allows us to act from universal principles that must be applied to particular empirically determined contexts. This aspect of our nature leads to all of the difficulties inherent to the task of successfully applying rules, making the abstract concrete, or producing a particular instance that is faithful to a universal norm. On the other hand, our *sensible* nature fixes it so that we stand in subjective practical relations to, and thus have feelings about, every practical situation we encounter. These feelings can often be very strong, influencing our practical attention in ways that make rational reflection and judgment more difficult. These two obstacles are thus inextricably bound up with our basic constitution as finite, embodied, sensing rational beings. They are not so much psychological quirks as *a priori* marks of our fallibility.

Turning now to Kant's discussion of these obstacles in the *Metaphysics of Morals*, it is helpful to begin with his remarks on the distinguishing features of the Doctrine of Virtue. While the first part of the *Metaphysics of Morals*, the Doctrine of Right, concerns duties that can be externally coerced by the moral community considered as a state, the Doctrine of Virtue concerns duties that the moral agent themselves must take on and fulfil. Kant construes this distinction as rooted in a necessary division in the concept of freedom. While considerations of right give principles to regulate external freedom, considerations of virtue yield principles for the realization of inner freedom. As Kant clarifies, inner freedom is thus "the condition of all *duties of virtue*" (MS

6:407). Central to this aspect of freedom is the activity of setting an end for oneself, where an end is construed as "an *object* of free choice [*Willkür*], the representation of which determines it to an action" (MS 6:384–5). The Doctrine of Virtue's central subject matter is, therefore, the ends it is a duty to have, or, more precisely, the ends one freely determines oneself to have in accordance with the moral law. It is in exploring this basic description of the Doctrine of Virtue that the first obstacle to good practical judgment comes to light, revealing deep difficulties for the exercise of freedom in sense-dependent rational beings.

2.2.1 *The Problem of Indeterminacy*

As soon as Kant specifies that the Doctrine of Virtue concerns those ends adopted through self-determination in accordance with universal laws, he recognizes the emergence of a crucial issue arising for the first time in the *Metaphysics of Morals*. In the Doctrine of Right, the concern was the principle in accordance with which particular actions could be externally constrained. That the agent has a particular end set by their private inclinations was taken for granted, and the question was only whether the action is permissible. When concern turns to the development of virtue, however, the question is not "do I have the right to do *x*?" where "*x*" is already specified, but "*what x* should I do to be virtuous?" A deep-seated difficulty arises insofar as Kant recognizes that the Doctrine of Virtue does not have the resources to specify particular actions on the basis of reason alone. Simply put, one does not cultivate virtue by adhering to a playbook of universal rules. The moral law only has the resources to specify *formal* constraints on the setting of ends, without giving particular actions that one must rigidly adhere to. As Kant says, "ethics does not give laws for actions ... but only for maxims of actions" (MS 6:388). This leads him to conclude that ethics can specify only *wide* duties:

if the law can prescribe only the maxim of actions, not actions themselves, this is a sign that it leaves a playroom

(*latitudo*) for free choice [*Willkür*] in following (complying with) the law, that is, the law cannot specify precisely in what way one is to act and how much one is to do by the action for an end that is also a duty. (MS 6:390)

I have termed this obstacle to virtue ‘the problem of indeterminacy’, which calls directly upon the cultivation of good practical judgment.³⁷ Kant is explicit about this towards the end of the Introduction to the Doctrine of Virtue when he writes: “ethics, because of the latitude it allows in its imperfect duties, unavoidably leads to questions that call

37. It is helpful to note that formal principles of morality and material principles of happiness must each deal with contrasting types of indeterminacy that reveal the basic nature of the principles themselves. As an *a priori* principle of reason, the moral law is general by nature, and concerned only with the universal character of the end in question. What matters is that the principle producing this end has the right form. This leads to the kind of indeterminacy identified above, insofar as careful judgment is needed to see how formal constraints can be applied to specify particular actions. In contrast, principles of happiness are based in sensibility and are conditioned directly by particular empirical circumstances and the contingent feelings we happen to have in response to them. This leads to a different kind of indeterminacy, insofar as the agent has difficulty unifying all the particular ends of happiness into a stable and coherent whole. With this type of indeterminacy, the agent can know that a particular action brings them pleasure in the present moment, but cannot know whether it will continue to do so or how this action will cohere with other things they want to do (cf. G 4:418–9). This difference can be summed up as follows: because morality is an ideal of reason, it represents an absolute totality whose necessary principles are easy to grasp with respect to general aims, but sometimes difficult to grasp in application to particular, concrete actions; and conversely, because happiness is an ideal of the imagination (G 4:418), it is easy to grasp in application to particular, concrete actions in the here and now, but difficult to grasp with respect to its representation as an absolute totality grounded in necessary principles that give a general picture of how to live. To invoke corresponding examples: it is easy to know I should care about the happiness of others, but sometimes difficult to know exactly how and when to act for this end; just as, conversely, it is easy to know that writing a philosophy paper this afternoon is bringing me satisfaction, but difficult to know how and whether the larger end of practicing academic philosophy will fit into a holistic vision of my lasting happiness. Given these contrasting types of indeterminacy, the weak agent in a particular empirical context could have a hazy sense of how to fulfill their moral duties, but a vivid sense of what would bring immediate pleasure or pain, the latter of which exercises influence on what they ultimately do. I am very grateful to an anonymous reviewer for pressing me to be clearer on this point.

upon judgment to decide how a maxim is to be applied in particular cases” (MS 6:411). Sharpening such judgment requires not just thinking about *a priori* principles of reason but engaging with the sensible particulars with which one is confronted. This calls for attention to the context of events as well as the practical responses — involving ends, feelings, inclinations, and affects — they evoke in the agent. It is thus with the problem of indeterminacy that we first find room for a non-dualistic account of moral failure, one which invokes the interdependence of reason and sensibility through the practice of judgment.³⁸

The basic tenor of the indeterminate questions to be decided by judgment is captured through Kant’s sub-section of “Casuistical questions” accompanying the treatment of each particular ethical duty. For example, in discussing the duty of beneficence, Kant asks:

how far should one expend one’s resources in practicing beneficence? Surely not to the extent that he himself would finally come to need the beneficence of others. How much worth has beneficence extended with a cold hand (by a will to be put into effect at one’s death)? (MS 6:454)

This is just a small subset of the questions Kant articulates to bring out the difficulties inherent to living out the duty of beneficence. While reason provides the principle in accordance with the formal constraints of the moral law, these constraints do not of themselves yield a determinate action. Instead, judgment must discern how to apply this universal law *in concreto*, navigating such issues as when to help others, how much to help them, and whether and to what extent such helping needs to heed the individual’s private conception of happiness (MS 6:454). Put differently, while the general end of beneficence is commanded by reason, whether a particular action in a given empirical

38. Merritt also calls attention to the importance of cultivating the capacity to recognize the relevant moral aspects of particular situations: e.g., “the moral law is grasped *in concreto* in the recognition of how particular facts about one’s situation give one reasons to do and not do certain things” (*Kant on Reflection and Virtue*, p. 192).

context counts as a well-executed case of beneficence is a further matter depending on sensitivity to a number of different contingent factors disclosed through the senses. The agent demonstrates virtue to the extent that their judgment is responsive to the relevant details that fill out a particular case.

Lest one think that the problem of indeterminacy is an obstacle only to be encountered in a few marginal cases or with respect to a particular subset of duties, it should be noted that any ethical duty covered by the Doctrine of Virtue is identified as a wide duty of virtue subject to this problem. Kant does distinguish between negative perfect duties and positive imperfect duties within the Doctrine of Virtue itself, but this further gradation should not obscure his more general claim that “ethical duties are of wide obligation, whereas duties of right are of narrow obligation” (MS 6:390). Perhaps the best evidence that Kant means all duties of virtue to be subject to the problem of indeterminacy is his decision to append his discussion of some perfect duties with a similar set of casuistical questions designed to point out the open-ended nature of practical judgment to his reader. For example, in the section on lying, Kant asks:

can an untruth from mere politeness ... be considered a lie? No one is deceived by it. — An author asks of one of his readers ‘How do you like my work?’ One could merely seem to give an answer, by joking about the impropriety of such a question. But who has his wit always ready? The author will take the slightest hesitation in answering as an insult. May one, then, say what is expected of one? (MS 6:431)

Through the example about the author and his reader, Kant highlights a crucial task for judgment: that of determining how different duties of virtue should limit one another when they converge in a particular sensibly given context. The reader is at once sensitive to their duty not to lie, and their duty to consider the happiness of the author. Though one of these duties is “negative” in character — lying

is always prohibited, and there is less latitude in determining whether a particular action constitutes a lie or not — there is still room for one to earnestly struggle in judging the situation at hand. Assessing how to limit one duty by another (must one cause unhappiness to avoid lying?)³⁹ and whether some duty applies in a particular situation (do well-meaning untruths count as lies?) are both crucial issues for judgment that arise because of the problem of indeterminacy, and they can confront the agent in connection with any duty of virtue.

2.2.2 *The Problem of Affect*

While the problem of indeterminacy highlights reason’s need to further specify *a priori* principles through concrete judgments about particular contexts, the problem of affect draws our attention to the way these latter contexts can be laden with affective meaning that influences judgment in various ways. Kant introduces the concept of affect towards the end of the Introduction to the Doctrine of Virtue in connection with his claim that virtue concerns the cultivation of inner freedom. He notes that two things are required for this inner freedom:

being one’s own *master* in a given case (*animus sui compos*), and *ruling oneself* (*imperium in semetipsum*), that is, subduing one’s affects and governing one’s passions. — In these two states one’s *character* (*indoles*) is noble (*erecta*); in the opposite case it is mean (*indoles abiecta, serva*). (MS 6:407)

Passion designates a deeper form of failure that involves actually taking up what is contrary to the moral law “into its maxim”, constituting what is “*properly* evil, that is, a true *vice*” (MS 6:408). Affect, however, can coexist with a will that is good in the abstract, and forms a key component of the account of weakness under development here, showing how sensible feelings can disrupt our practical lives and make it difficult to apply abstract principles.⁴⁰

39. For another passage where Kant discusses the issue of limiting one duty by another, see MS 6:390.

40. In aligning frailty with affect, the present reading differs sharply from another

Though Kant typically uses the term ‘affect’ in a more general sense to describe what sensory contact with an object produces in the subject, in this context, the term takes on a more specific meaning. Kant describes an affect as “belonging to feeling insofar as, preceding reflection, it makes this impossible or more difficult” (MS 6:407). Kant’s remarks on affect in the Doctrine of Virtue assign the concept a key role in overcoming weakness and developing virtue (cf. MS 6:407, cited previously), but his development of the concept itself is relatively sparse. Accordingly, it helps to supplement this account with his more in-depth treatment in *Anthropology from a Pragmatic Point of View*. In Book III, on the Faculty of Desire, he describes affect in similar terms: “affect is surprise through sensation, by means of which the mind’s composure is suspended. Affect is therefore rash, that is, it quickly grows to a degree of feeling that makes reflection impossible (it is thoughtless)” (A 7:252). Unfortunately, the talk of “degree of feeling” can lead us to conclude that Kant is working with the empiricist’s conception of feeling that focuses on its quasi-mechanical strength of force. This would pose a problem for the present reading, for, as we have seen, the quasi-mechanical view saddles us with a dualistic conception of reason and sensibility, on which feeling is a sensible response to one’s environment that arises independently of cognitive mental life. But Kant helpfully clarifies that this is not how he means to use the language of strength: “in general, it is not the intensity of a certain feeling that constitutes the affected state, but the lack of reflection in comparing this feeling with the sum of all feelings (of pleasure

or displeasure)” (A 7:254). This clarification helps get us to the desired view. Though we can understand the agent’s general capacity to sense and feel—including their more idiosyncratic dispositions towards particular feelings—in terms of psychological laws that track physical causes and degrees of force, this is not Kant’s focus in his moral philosophy or even his anthropological writings. Instead, we need to consider affect insofar as it relates to our higher capacity to reflect and judge. So Kant’s account of affect is not articulated within the economy of natural causes. Affect does not lead to thoughtlessness because it describes feelings that overpower on the basis of their psychological strength. Rather, it picks out the set of feelings we find it difficult to reflect upon. Understanding the scope of this set requires looking not into the mechanics of the mind, but into the subject’s personal history and their natural as well as cultivated abilities: the kinds of situations one has found oneself in so far; how self-aware one is about which affects one is particularly susceptible to; how one has reacted to similar moments of affect in the past; whether one has resolved to make progress; how attentive one is to the possibility of being caught off guard. Consider the example Kant gives of the rich man who flies into a rage when his server breaks a rare goblet, suffering from the affect of anger:

[he] would think nothing of this accident if, at the same moment, he were to compare this loss of *one* pleasure with the multitude of *all* the pleasures that his fortunate position as a rich man offers him. However, if he now gives himself over completely to this one feeling of pain (without quickly making that calculation in thought), then it is no wonder that, as a result, he feels as if his entire happiness were lost. (A 7:254)⁴¹

account offered by Morrisson, “On Kantian Maxims”, pp. 85–7. In defense of Allison’s position, Morrisson insists that even weak action must express a chosen maxim, suggesting that Kant’s view of weakness is better illustrated through examples of passion, since Kant is very clear that “passion always presupposes a maxim of the subject, namely, to act according to a purpose prescribed for him by his inclination” (A 7:266). I would argue that Morrisson’s commitment to the Incorporation Thesis here has resulted in an interpretation that does not fit with Kant’s general view of weakness: as I would have it, weakness should be explained by appealing to the activity of judgment and its ability to cope with indeterminacy and affect, rather than immoral maxim-formation, which better captures the later stages of evil.

41. Kant’s example can be variously interpreted as primarily moral *or* prudential in its focus. Insofar as the rich man cannot properly reflect on his good fortune, he makes himself more prone to unhappiness. One might also think of the poor servant subject to his rage, which makes the scenario morally relevant. The problem of affect is key to either reading, though matters of prudence are not my focus here.

What Kant describes here is an inability to appropriately relate one feeling to all the others that fill out one's practical life. This inability is, to be sure, dependent on one's natural propensity to feel in particular ways — something over which the subject does not have complete control. But it is fundamentally a failure of reflection, which one can work to control once it is brought to consciousness.

This language of control — or more precisely, self-control — reconnects us with Kant's insistence that to realize inner freedom and demonstrate virtue, one must be one's own master through subduing affect (MS 6:407). This general decree is borne out in his treatment of courage in the face of fear in the *Anthropology*. Consider the following passage: "anxiety, anguish, horror, and terror are degrees of fear, that is, degrees of aversion to danger. The composure of the mind to take on fear with reflection is *courage*" (A 7:256). How much fear one feels in response to a given situation is described as largely pertaining to the contingencies of sensibility, and therefore beyond the agent's control. One may be naturally "stout-hearted" and not quick to frighten, but the morally relevant fact is not how much fear one feels, but how reason is able to deal with it: "*stout-heartedness* ... is merely a quality of temperament. *Courage*, on the other hand, rests on principles and is a virtue. Reason then gives the resolute man strength that nature sometimes denies him" (A 7:256–7). Courage as a virtue thus indicates that one's reason is developed enough in its capacity for judgment so as to expertly apply the moral law to situations involving the affect of fear. As Kant indicates with his example of the rich man flying into a rage, this requires seeing how the particular feeling of fear and the desires it conditions are just one small part of the many considerations weighing on the agent in that moment. If the agent has failed to develop courage, they will be ill-prepared to situate the danger their fear is responding to in proper relation to other ends that are equally or more relevant to the circumstances. The feeling of fear stands out so singularly in the agent's limited capacity for reflection, that they are

unable to give due weight to other ends, including those duties of virtue commanded by the moral law itself. Insofar as the agent's reasoning is influenced by this strong feeling of fear, and they find it difficult to judge what is morally required in the present circumstances, they are likely to act out of a concern for self-interest that appears sufficient to warrant flight or some other form of inaction. Were this agent more practiced in their judgment, this feeling would be less influential, and they would opt for a different course of action, one that considers additional ends, such as, perhaps, that of helping others.

2.2.3 *Susceptibility to Indeterminacy and Affect as Weakness*

I have identified two obstacles to good practical judgment that, I am claiming, help fill out a non-dualistic account of moral weakness. As obstacles to practical *judgment*, they are to be encountered at the level of concrete action, of applying universal principles to particular situations in order to figure out what one should do. The *problem of indeterminacy* highlights a cognitive difficulty or a difficulty of understanding, insofar as the agent struggles to determine which duty or duties apply to a particular situation, and what concrete action should be undertaken to respond to them. For example, does my duty to refrain from lying require that I limit the happiness of another person by telling them a painful truth? This is essentially the difficulty of grappling with a complex manifold of sensible particulars in order to grasp the universal principles that determine or ought to determine them. The *problem of affect* highlights a difficulty attending to the practical efficacy of sensibility: this manifold of sensibly given particulars is such as to evoke strong feeling in us, and such feeling exercises influence on reason's activity. The affected agent struggles insofar as these strong feelings constrain their ability to navigate the problem of indeterminacy — to work out which duties apply and how to respond to them through concrete action — by effectively narrowing their practical vision, limiting it to the practical factors relevant to the affect(s) in question. These obstacles are, accordingly, closely interrelated: the better one is at navigating the problem of indeterminacy, the better one will

be at responding to affect; and the more susceptible one is to affect, the harder it will be to navigate indeterminacy. In the former case, the agent will be adept at seeing which duties are relevant to a particular context and how they should act in response to them, which means they will be better able to keep the whole practical context in mind, not allowing their moral vision to be narrowed by strong feelings of anger, fear, etc. In the latter case, the agent has not sufficiently developed the reflective capacities to sort out what is required and keep the whole practical context in view, and their judgment will be influenced by strong feeling, as demonstrated in the case of timidity discussed above.

Now that the obstacles to good practical judgment have been explored in more detail, we can examine the textual evidence linking them to moral weakness. The present account is not merely a reconstructive suggestion that we should unpack the latter issue using these obstacles. Kant himself references moral weakness in the Doctrine of Virtue in the course of discussing both indeterminacy and affect. Directly after introducing the claim that virtue is governed by wide or imperfect obligations that leave a playroom for judgment, we find the following passage:

imperfect duties alone are, accordingly, *duties of virtue*. Fulfillment of them is *merit (meritum)* =+a; but failure to fulfill them is not in itself culpability (*demeritum*)=-a, but rather mere deficiency in moral worth =o, unless the subject should make it his principle not to comply with such duties. It is only the strength of one's resolution, in the first case, that is properly called *virtue (virtuus)*; one's weakness, in the second case, is not so much *vice (vitium)* as rather mere *want of virtue*, lack of moral strength (*defectus moralis*). ... Every action contrary to duty is called a *transgression (peccatum)*. It is when an intentional transgression has become a principle that it is properly called a *vice (vitium)*. (MS 6:390)

Here, Kant distinguishes between two types of moral transgression: one that involves intentionally incorporating sensible incentives into one's maxim, and one that involves unintentionally lacking virtue due to weakness. Insofar as aiming for virtue is closely attended by the problem of indeterminacy, there opens up space for this difference. One can have a will that is good in the abstract, i.e., a will that has not incorporated selfish incentives, and yet still exhibit moral transgression insofar as one does not have a practiced judgment capable of navigating the indeterminacy problem. To lack this experience and the resolve that comes with it is to be weak and wanting in virtue.

Similarly, when Kant introduces the concept of affect later on in the Introduction to the Doctrine of Virtue, he claims:

an affect is called *precipitate or rash (animus praeceps)*, and reason says, through the concept of virtue, that one should *get hold of oneself*. Yet this weakness in the use of one's understanding coupled with the strength of one's emotions is only a *lack of virtue* and, as it were, something childish and weak, which can indeed coexist with the best will. (MS 6:407–8)

Once again, we find Kant framing moral weakness as a lack of virtue that is not to be explained by appealing to the incorporation of sensible incentives into the agent's general maxim for acting. The will is still "good" insofar as the moral law is recognized to be the only true incentive for action, but the agent's understanding is not practiced enough to grapple with the affects that influence and narrow their practical judgment. The result is a kind of transgression best understood as moral weakness.

2.3 Moral Weakness as a Description of Noumenal Character

The foregoing analysis of the main obstacles to practical judgment was largely conducted with an eye to explaining actions or deeds in the *phenomenal* sense. The focus was particular sensibly given contexts and the specific actions they call for, as well as the particular feelings

of fear, anger, etc. and the desires to which they give rise, which make these contexts especially difficult to navigate. My claim was that these obstacles are what open up the conceptual space for weakness or frailty as the least severe form of moral failure, a form of failure that occurs at the level of maxim application rather than formation. The question remains how to frame this account with respect to the *noumenal* deed that Kant emphasizes during his discussion of moral evil in Part One of the *Religion*. The key interpretive aim is to give an account of weakness that both i) counts it as a form of transgression occurring at the noumenal level – a level that implicates the agent’s freely determined disposition or character, and yet also ii) respects Kant’s insistence that the weak agent’s supreme maxim, which captures their most fundamental orientation to moral life, does *not* subordinate moral incentives to selfish ones. To fail in regard to ii) would effectively collapse weakness into the more severe stages of impurity or depravity. Weakness indeed counts as a grade of evil, and thus involves acting on sensible as opposed to moral incentives, but the weak agent’s character is not best described as acting on the basis of the unconditional or conditional prioritization of selfish concerns.

The question is thus: how can this less entrenched form of evil be expressed as a noumenal deed capturing the agent’s basic moral character or orientation? Given how I have been framing the issue, this general disposition should capture the agent’s susceptibility to falter in response to indeterminacy and affect. As we have seen, Kant never explicitly formulates how to express weakness as a supreme maxim in the *Religion*. He mainly gives it a negative characterization, depicting the weak will as “not being strong enough to comply with its adopted principles” (R 6:37). To find some basis for a positive characterization, it is helpful to look to Kant’s philosophical reading of the book of *Genesis*. In a passage worth quoting in full, he describes the first human sin as originating from a progression that starts in a state of innocence and proceeds through all three grades of evil:

instead of following this law absolutely as sufficient incentive (which alone is unconditionally good, and with which there cannot be further hesitation), the human being looked about for yet other incentives (III:6) which can be good only conditionally (i.e. so far as they do not infringe the law). **And he made it his maxim – if one thinks of an action as originating from freedom with consciousness – to follow the law of duty, not from duty but, if need be, also with an eye to other aims.** He thereby began to question the stringency of the command that excludes the influence of every other incentive, and thereupon to rationalize downgrading his obedience to the command to the status of merely conditional obedience as a means (under the concept of self-love), until, finally, the preponderance of the sensory inducements over the incentive of the law was incorporated into the maxim of action, and thus sin came to be (III:6). (R 6:42, emphasis mine)

Towards the end of the passage, we see descriptions of impurity and depravity that closely resemble earlier formulations.⁴² The agent depicted here demonstrates impurity when he questions the “stringency” of the law, coopting his obedience into a way of also pursuing ends of happiness. This turns to depravity when the agent fully incorporates the subordination of morality to happiness, turning happiness into his ultimate, unconditional end. Before these later stages take hold, the agent is described as looking about “for yet other incentives (III:6)” and as having the maxim “to follow the law of duty, not from duty but, if need be, also with an eye to other aims”. The biblical passage Kant makes direct reference to here is helpful in bringing out what this entails. It describes Eve’s responsiveness to the sensory aspects of her situation as she looks at the tree bearing fruit: “and the woman saw that the tree was good for eating and that it was lust to the eyes and the

42. Cf. R 6:30.

tree was lovely to look at" (III:6).⁴³ Two aspects of sensory contribution are noted here. First, any sensibly given context will reveal a number of converging practical factors, including various ends to consider. In this case, the tree's fruit is seen to be good for eating, raising the question, "should I eat it?" Does one's general principle to eat delicious, nutritious food when hungry need to be limited by another consideration, or should this end take precedence? Second, the strong affective element bound up with this practical context is also emphasized. On this point, the translator notes that the ancient Hebrew term rendered as "lust to the eyes" "means that which is intensely desired, appetite, sometimes specifically lust", while the term translated as "lovely" literally means "that which is desired".⁴⁴ On Kant's terminology, the passage thus describes desire born in response to an all-encompassing "affect" that arises in light of one's sensibly given context. As he implies with various progress-indicating terms that populate the rest of the passage, this practical responsiveness is the first step on the way to a full reversal of the agent's incentives, but it is not *yet* that. The agent's practical openness to ends other than morality itself (frailty), is what allows them to begin to view the moral law as a means to happiness (impurity), and finally to fail to treat morality as an end-in-itself at all, completely subordinating it to happiness (depravity). I thus take the passage to indicate that when Kant initially describes the first human agent as making it their maxim to follow the law "with an eye to other aims", he is not registering the eventual depravity that will grow out of weakness, but a far less extreme and logically prior moral orientation that is a condition on the possibility of later stages of evil. As I have argued using the passage at hand, this orientation is captured through the activity of earnestly trying to live out the moral law while also being practically responsive to one's sensory context, and thus liable to the obstacles of indeterminacy and affect. This is, on Kant's framework,

43. Alter (trans.), *Genesis*, p. 12.

44. *Genesis*, p. 12.

a noumenal description of weakness, understood as a basic practical orientation that grounds and explains particular phenomenal actions.

Herein lies the deep truth about the nature of our agency that was promised at the end of the introductory section. Kant's highly complex account of sense-dependent practical reasoning reveals weakness to be an inescapable part of human nature against which we must perennially struggle. As sense-dependent rational beings, we can *only* apply our moral principles in response to sensibly given contexts, and these contexts cannot help but involve sensible incentives pertaining to happiness. As soon as the agent is open to the world, sensory inducements assail them. As Kant notes towards the end of the second *Critique*, inclinations "always have the first word" (KpV 5:147). To stand in sensory relations to objects is to have feelings of pleasure and pain that evoke desires and the ends at which they aim. It is these subjective, sense-dependent ends, understood as forming a self-consistent whole, that constitute one's happiness, the subjective end that we must all will as embodied rational agents. While it is simply a matter of fact that happiness is our subjective end, the free use of our agency is implicated in our propensity to grant this end too much weight in choosing what to do, whether through the full-on depravity that subordinates morality to happiness, or the weakness in judgment that fails to grasp what is objectively called for in a given situation because of indeterminacy and affect. While the former manifestation of fallibility requires an entrenched and developed viciousness, the latter is so intertwined with our practical nature as to suggest that weakness is the general condition we persist in as finite moral agents. Indeed, when Kant first introduces the concept of weakness in the *Religion*, he speaks not of a propensity but of the outright "*frailty* of human nature" (R 6:29, and again at R 6:37).⁴⁵ This understanding of frailty as a universal

45. This is further corroborated by Kant's students' collected notes from his Lectures on Ethics. The Collins Lecture Notes mention "the frailty" or "fragility" of "human nature" (LE 27:295-6), and claim that "it is right and proper, indeed, for man to recognize his weakness, but not to be deprived of his good dispositions" (LE 27:351). In the Mrongrovius Lecture Notes, it is written that no human being could serve as the prototype of morality because of this perennial

condition we must struggle through also fits with Kant's characterization of virtue as "moral disposition in conflict" (KpV 5:84). Even the agent who demonstrates purity in their basic moral orientation must still develop the "capacity to master one's inclinations when they rebel against the law" (MS 6:383) and thus be able to subdue their affects through the reflective clarity that comes with an experienced faculty of judgment. This, as Kant emphasizes, is a lifelong project. Because we are affected by inclinations in the unavoidable manner I have been describing, "virtue can never settle down in peace and quiet with its maxims adopted once and for all but, if it is not rising, is unavoidably sinking" (MS 6:409).

2.4 *Tying Everything Together*

One final set of *Religion* passages discussing the relationship between noumenal and phenomenal deeds should help tie everything together by accounting for the various pieces of terminology that have been introduced. These are the passages where Kant speaks of the moral agent as undergoing a revolution with respect to their noumenal character and a slow reformation with respect to their phenomenal one (roughly, R 6:45–8). I take these passages to support the reading advanced above: that the weak agent is best described as one whose supreme maxim is pure but unstable because of its propensity to give undue weight to subjective considerations — those based on mere inclination — in judging what to do to fulfill the moral law. Given that this propensity is ineradicably bound up with our nature as sense-dependent beings, the road to moral goodness is not a matter of incorporating the moral law more completely into one's will or overcoming self-deception, but of gaining the practiced judgment that can deal with indeterminacy and affect. Consider Kant's short formulation of what needs to occur for good moral character to develop: "a revolution is necessary in the mode of thought but a gradual reformation in the mode of sense (which places obstacles in the way of the former)" (R

weakness: "a natural man can never be the ideal, for he is still always subject to weakness" (LE 27:605).

6:47). The revolution Kant speaks of here refers to the total reversal in the ordering of one's incentives necessary for the virtuous disposition. Insofar as morality is neither conditionally nor unconditionally subordinated to happiness, the supreme maxim is pure. Having a pure supreme maxim dedicated to morality, however, does not mean that one is perfectly holy and will always do the right thing, for, as Kant insists in a nearby passage, "between maxim and deed there is still a wide gap" (R 6:46). In accordance with the earlier passage at R 6:47, we should conclude that this gap occurs because our sense-dependence places obstacles in the way of even the purest will. While Kant sometimes speaks imprecisely and simply identifies the inclinations themselves as the obstacles that get in the way, he clarifies in the *Religion* that inclinations are an obstacle only because of how reason conducts itself in response to them (R 6:35). This suggests the obstacles in question are better understood as obstacles to the use of practical judgment, i.e., indeterminacy and affect, as I have identified them. Our propensity to be susceptible to such obstacles in the free use of our rational agency is the propensity to evil considered in its most basic and ineradicable form, i.e., as frailty or weakness.

This framework carefully maps onto the dual sense of 'deed' that has been used throughout. The revolution that comes from adopting a newly purified supreme maxim is a wholly complete noumenal deed. It has no middle ground, no sense of being in progress. One is either pure or impure in one's moral orientation. However, insofar as this noumenal deed is the ground of further phenomenal deeds, the notion of progress is highly relevant. At issue here is the agent's cultivated capacity to actualize this pure disposition in response to the potentially infinite sensibly given contexts that could confront them. As I have been arguing, this is a matter of applying one's principles *in concreto*, success in doing which requires the sharpened judgment required to navigate the obstacles of indeterminacy and affect. Such activity is never complete. As Kant says,

he is a good human being only in incessant laboring and becoming i.e. he can hope — in view of the purity of the principle which he has adopted as the supreme maxim of his power of choice [*Willkür*], and in view of the stability of this principle — to find himself upon the good (though narrow) path of constant *progress* from bad to better. (R 6:48)

Though all human beings are weak to the extent that this progress is never finished, I take this passage to suggest that one is more or less weak insofar as one's supreme maxim is more or less stable as a ground of phenomenal action. The reading of Kant I have offered here interprets this stability as cognitive and only thereby motivational. That is, the agent is more stable and less weak insofar as their judgment is better able to deal with the obstacles of indeterminacy and affect, and so arrive at a determinate representation of what to do — as a deed understood in the phenomenal sense. Such an agent demonstrates the kind of virtue Kant describes in the *Metaphysics of Morals*, one which is grounded in purified principles and whose firmness is “armed for all situations” and “adequately secured against the changes that new temptations could bring about” (MS 6:383–4).⁴⁶

3. Conclusion

I will now take stock of the implications for the view of moral weakness I have developed. First, we should return to the problem of dualism. At the outset of this paper, I claimed that developing a viable moral psychology from Kant's texts requires overcoming the prevalence of a certain kind of dualism — one which holds that the exercises of our

rational and sensible capacities are best understood without reference to one another. Two prominent and exemplary versions of this dualism helped to show the shortcomings natural to it. Common to both views is the assumption that for practical reason to demonstrate autonomy in the noumenal realm, we must think of the agent as making a voluntaristic meta-choice to at least sometimes prioritize one's own happiness over morality. Differences in moral character would then correspond to how frequent or all-encompassing this choice is. On this view, the depraved agent demonstrates the most minimal commitment to morality that is possible for a moral agent by completely subordinating it to the end of happiness. The frail or weak agent, in contrast, shows a firmer but still fallible commitment insofar as their will is usually or generally on the right track but still makes exceptions in favor of happiness, exceptions that show there is still room for this commitment to be strengthened.⁴⁷ On some accounts, this need for a stronger commitment is cashed out in terms of overcoming self-deception about one's willingness to make exceptions in following the moral law.⁴⁸ Sharing in this foundation, the two forms of dualism diverge in how they frame this account of noumenal agency in reference to the phenomenal realm. The empiricist dualist explains phenomenal action by appeal to quasi-Humean affective forces, while the rationalist does so by appeal to maxims formed by reason. As I argued, fatal interpretive difficulties arise in either case, for focusing only on sensibility or reason in isolation from the other to explain weak action — or indeed any action — leads to an impoverished moral psychology that does not capture the complexity of Kant's view.

My own reading of Kant on moral weakness frames it instead as a failure in practical judgment as opposed to a failure in commitment.

46. At this point, we can ask the further question of *how* the moral agent is to make progress against the constant obstacles of indeterminacy and affect, and it is here that Kant's discussion of the ethical community united in religious practice from part three of the *Religion* becomes absolutely critical. For some recent commentary on this historically neglected part of his view, see Palmquist (*Comprehensive Commentary*, chap. 7), DiCenso, *Kant's Religion within the Boundaries of Mere Reason: A Commentary*, chap. 5; and Pasterneck (*Kant on Religion within the Boundaries of Mere Reason*, chap. 5).

47. Cf. Allison: “the so-called lack of sufficient strength to follow moral principles when they conflict with the claims of inclination reflects the lack of a full commitment to these principles in the first place” (*Kant's Theory of Freedom*, p. 159).

48. Cf. Allison: “thus, self-deception enters the picture at the very beginning, depicting what is in reality a free evaluation on one's part as a ‘weakness’ for which one is not responsible” (*ibid.*).

Weakness results from the inability to successfully apply *in concreto* the moral principles to which one is committed *in abstracto*. On my view then, we need not appeal to self-deception or frame one's noumenal commitment as in need of shoring up at the noumenal level. Becoming less weak is not simply a matter of opting to act from the moral law on a more regular basis, even when tempted not to. It is rather a matter of sharpening one's judgment so that it is able to overcome the obstacles that attend to the application of universal principles in particular sensibly given contexts. As I have shown, it would be overly simplistic to identify these obstacles as dualistically opposed inclinations running contrary to the moral law. The obstacles of indeterminacy and affect involve sense-based feelings, desires, and ends, but the difficulties they occasion are fundamentally *cognitive* difficulties, highlighting just how demanding it can be to figure out what to do at ground level when unforeseen contexts arise and multiple duties and ends intersect. The problem of indeterminacy brings out that it can be no small task to assess how highly abstract laws and duties apply in particular situations to produce concrete actions. The problem of affect brings out that the practical efficacy belonging to our feelings and desires can cloud our judgment, making it difficult to see what to do. Together, these obstacles show the non-dualistic interdependence of reason and sensibility in human judgment and action. Acts of practical reasoning cannot be fully understood without reference to the sensible particulars to which they apply; and the significance of conative representations such as affect, feeling, and sensible desire cannot be fully understood without reference to their interaction with and influence upon practical reasoning. Kant's reader can give an adequate account of the relation between reason and sensibility only to the extent that this interdependence is recognized.

Moreover, this interdependence is implicated in the basic moral orientation of the weak agent understood as a noumenal deed. When Kant talks of the ineradicable human propensity to evil, he has in mind the fact that sense-dependent reason is always susceptible to falter in response to indeterminacy and affect. Our basic moral orientation

can thus be pure, but the work of realizing this disposition is never finished and becoming virtuous is always a matter of slow and continual progress in the cultivation of judgment at the phenomenal level. One's noumenal character as a human being who can never completely overcome weakness is thus best captured through the idea that in our use of reason we are always sensitive to non-moral aims. At the level of frailty, however, this sensitivity is not construed in terms of a weakened *commitment* to morality. Though the weak agent ultimately acts on a merely subjective ground of action when her practical judgment falters, it would be misleading to frame this in terms of the meta-level subordination of morality to happiness, which only occurs at the stages of impurity and depravity. We should say, rather, that the agent's practical judgment is uncultivated, meaning they do not fulfill their earnestly made commitment to morality in practice because they fail to arrive at a determinate action-guiding conception of how to fulfill this commitment at the level of concrete particulars. Though we are universally liable to such failure whatever we do, it is not a deterministic flaw for which we cannot be held responsible. Even on my reading of weakness, it makes sense to cast it as a basic moral orientation at the noumenal level, because it still describes the way the agent makes free, active use of their rational capacities — a use which can and should be properly cultivated — rather than a temporally conditioned natural occurrence.

I hope to have now articulated a textually founded, comprehensive, and attractive reading of Kantian moral weakness that is substantively different from those offered by interpreters with a more dualistic approach to Kant's moral psychology. Though I briefly discussed some concrete cases in §2.2, a much more detailed account of how this general view of moral weakness applies to a wide variety of examples would be desirable. I do not have the space for this here, but will close with some brief remarks about the implications this view has for the type of case that traditionally comes into focus when analytic philosophers discuss weakness of will. It can be difficult to describe such cases in theoretically neutral terms, but the distinctive feature that marks them

can, I think, be non-controversially described in terms of the agent knowing that they are violating a *moral*⁴⁹ commitment and so failing to act on some moral knowledge that they have. So for example, the agent who recognizes and regrets their lie even in the act of telling it. Because the obstacles of indeterminacy and affect can ground the failure of judgment *in concreto* in a multitude of ways, the account I have developed is extremely plastic, giving different explanations for different cases. Such detailed engagement is beyond what I can attempt now however, so I will instead turn to the general form of explanation for weakness of will in the traditional sense.

First, whatever the particular case, the agent's pure moral orientation should be invoked to explain that they have a basic understanding of and thereby are motivated towards abstract moral ends like respecting others, valuing their happiness, being honest, etc. But second, the instability of this moral orientation should also be invoked to show why they have failed to uphold the relevant moral end(s) in the case at issue. On my reading, this failure is explained by appeal to the uncultivated judgment that fails to come to a determinate, action-guiding representation of what to do. Importantly, just knowing, for example, not to lie is not enough here. What is needed in addition is a positive, determinate specification of what to do instead. How, for example, to tell a difficult truth in a considerate way that does not disrespect or unnecessarily hurt others. On the view articulated here, what is unique about the traditional case is that the agent has a developed enough judgment to immediately recognize that their response to their current situation is inadequate — a recognition that can be missing in other cases which could also fit under the broader sense of weakness Kant identifies — though this judgment is not yet developed enough

49. I emphasize the term 'moral', because prudential weakness that is not yet moral is a separate issue; for example, when one aims to exercise regularly, but sometimes finds oneself unmotivated to go to the gym at the end of a long workday. While maintaining one's physical health is broadly a duty of virtue, it is imperfect enough that this particular kind of practical failure (e.g., actively caring for one's health in practice, but skipping the gym sometimes) does not count as moral weakness.

to get all the way to a determinate, action-guiding representation of a particular action that fulfills the moral demands at issue. Third, given that the traditional case is described as one where the agent also has a strong momentary desire that runs counter to their moral knowledge, we should draw on the problem of affect as well. So, for example, anxiety about causing pain or fear of social repercussion influence the agent in judging what to do. Importantly, however, we cannot say that the agent's view of the practical situation has been so narrowed by affect that they are blind to its influence. In such cases, the agent will possess some level of awareness that their judgment is faltering in response to a strong desire that influences their ability to navigate the problem of indeterminacy. On my reading, then, traditional cases of moral weakness of will indicate that the agent is close to being able to do what the virtuous person would do, but not quite there. Their capacity to judge about concrete particulars is developed enough to recognize, to at least some degree, the various ends, desires, and duties involved, and it is also developed enough to recognize their own failure, a failure described here as judgment that lacks a fully determinate and thereby motivating conception of how to act. So although the weak action might be influenced by affect, its ultimate explanation is a form of cognitive failure that is striking in its self-conscious character. The agent fails to act not because the desire to do otherwise is very strong (though it may well be), but because their judgment is still lacking.⁵⁰

As this brief analysis of the traditional case shows, a non-dualistic Kantian moral psychology, once fully developed, should speak against a number of common assumptions about Kant's view. These include:

50. Though I would claim that both views are ultimately varieties of the rationalist dualism paradigmatically expressed by Allison, it is noteworthy that Thomas Hill Jr. (in *Virtue, Rules, and Justice: Kantian Aspirations*, pp. 120–1) and Laura Papish (in *Kant on Evil, Self-Deception, and Moral Reform*, pp. 195–6) also describe the weak agent as lacking in determinate knowledge of the concrete implications of their moral commitments. Merritt, who comes much closer to my non-dualistic view of Kantian moral psychology, also endorses a similar view of the traditional case of weakness of will, though she does not develop it (*Kant on Reflection and Virtue*, p. 192).

i) the assumption that practical reason is first and foremost a volitional capacity best described in terms of choice and commitment, as opposed to a cognitive capacity whose activity is best described in terms of judgment that aims at determinate knowledge; ii) the assumption that the strength of one's commitment or motivation is fundamentally separable from one's practical knowledge of what to do; iii) the assumption that Kant's egalitarian conception of practical reason means it is always easy to know exactly what to do to uphold the commands of morality; and iv) the assumption that cultivating virtue requires looking inward to scrutinize one's moral commitments and overcome self-deception, as opposed to looking outwards to cultivate one's capacity to judge about concrete sensibly given situations. While providing a more complete account of this non-dualistic moral psychology is beyond the present work, I hope to have given persuasive evidence that Kant's text is often suggestive of it, and that at least some of these common assumptions, often thought to be foundational elements of his view, should be open to debate.⁵¹

References

- Allison, Henry. *Kant's Theory of Freedom*. New York: Cambridge University Press, 1990.
- Alter, Robert (Trans. and Comm.). *Genesis*. New York: Norton, 1996.
- Baxley, Anne Margaret. *Kant's Theory of Virtue: The Value of Autocracy*. New York: Cambridge University Press, 2010.
- Boyle, Matthew. "Additive Theories of Rationality: A Critique." *European Journal of Philosophy* 24 (2016): 527–555.
51. This paper has been years in the making and there are many people to thank. I would especially like to mention Julia Peters, Melissa Merritt, Laura Papish, Mavis Biss, Alexandra Newton, Anastasia Berg, Matthias Haase, Jim Conant, Matt Boyle, Lionel Shapiro, and the graduate students at the University of Connecticut who took part in my Kant Seminar in Spring 2020. I am also grateful to audiences from the 2018 NAKS Biennial Meeting, the 2018 Eastern and 2019 Midwestern NAKS Study Groups, the 2019 International Kant Congress, and the German Philosophy workshop at the University of Chicago. Finally, I would like to thank two anonymous reviewers whose comments were extraordinarily helpful.
- Broadie, Alexander and Pybus, Elizabeth. "Kant and Weakness of Will." *Kant-Studien* 73 (1982): 406–412.
- DiCenso, James. *Kant's Religion within the Boundaries of Mere Reason: A Commentary*. New York: Cambridge University Press, 2012.
- Guyer, Paul. *Knowledge, Reason, and Taste: Kant's Response to Hume*. Princeton: Princeton University Press, 2008.
- Hill, Thomas Jr. *Virtue, Rules, and Justice: Kantian Aspirations*. New York: Oxford University Press, 2012.
- Johnson, Robert N. "Weakness Incorporated." *History of Philosophy Quarterly* 15 (1998): 349–367.
- Kant, Immanuel. "Anthropology from a pragmatic point of view." In G. Zöllner and R. B. Loudon (Eds. and Trans.), *Anthropology, History, and Education* (pp. 227–429). New York: Cambridge University Press, 2007.
- . "Critique of Practical Reason." In M. Gregor (Ed. and Trans.), *Practical Philosophy* (pp. 133–272). New York: Cambridge University Press, 1996.
- . *Critique of Pure Reason*. P. Guyer and A. Wood (Eds. and Trans.). New York: Cambridge University Press, 1998.
- . *Critique of the Power of Judgment*. P. Guyer and E. Matthews (Eds. and Trans.). New York: Cambridge University Press, 2000.
- . "Groundwork of the Metaphysics of Morals." In M. Gregor (Ed. and Trans.), *Practical Philosophy* (pp. 37–108). New York: Cambridge University Press, 1996.
- . "Lectures on Ethics." In P. Heath and J. Schneewind (Eds. and Trans.), *Lectures on Ethics*. New York: Cambridge University Press, 1997.
- . "Metaphysics of Morals." In M. Gregor (Ed. and Trans.), *Practical Philosophy* (pp. 353–604). New York: Cambridge University Press, 1996.
- . "On the common saying: That may be correct in theory, but it is of no use in practice." In M. Gregor (Ed. and Trans.), *Practical Philosophy* (pp. 273–310). New York: Cambridge University Press, 1996.
- . "Prolegomena to any future metaphysics that will be able to come forward as science." In G. Hatfield, M. Friedman, H. Allison, and

- P. Heath (Eds. And Trans.), *Theoretical Philosophy after 1781* (pp. 29–170). New York: Cambridge University Press, 2002.
- . “Religion within the boundaries of mere reason.” In A. Wood and G. di Giovanni (Eds. and Trans.), *Religion and Rational Theology* (pp. 39–216). New York: Cambridge University Press, 1996.
- Merritt, Melissa. *Kant on Reflection and Virtue*. New York: Cambridge University Press, 2018.
- McCarty, Richard. *Kant’s Theory of Action*. New York: Oxford University Press, 2009.
- . “Kantian Moral Motivation and the Feeling of Respect.” *Journal of the History of Philosophy* 31 (1993): 421–435.
- . “Moral Weakness as Self-Deception.” In H. Robinson (Ed.), *Proceeding of the Eighth International Kant Congress: Memphis 1995, Vol. II* (pp. 587–593). Milwaukee: Marquette University Press, 1995.
- . “Motivation and Moral Choice in Kant’s Theory of Rational Agency.” *Kant-Studien* 85 (1994): 15–31.
- Morrisson, Iain. “On Kantian Maxims: A Reconciliation of the Incorporation Thesis and Weakness of the Will.” *History of Philosophy Quarterly* 22 (2005): 73–89.
- Palmquist, Stephen R. *Comprehensive Commentary on Kant’s Religion within the Bounds of Bare Reason*. West Sussex: Wiley-Blackwell, 2016.
- Papish, Laura. *Kant on Evil, Self-Deception, and Moral Reform*. New York: Oxford University Press, 2018.
- Pasternack, Lawrence. *Kant on Religion within the Boundaries of Mere Reason*. New York: Routledge, 2014.
- Rohlf, Michael. “Immanuel Kant.” *Stanford Encyclopedia of Philosophy*, 2020. <https://plato.stanford.edu/entries/kant/>.
- Tizzard, Jessica. “Kant on Space, Time, and Respect for the Moral Law as Analogous Formal Elements of Sensibility.” *European Journal of Philosophy* 26 (2018): 630–646.