

THE METAPHYSICS OF VICE: KANT AND THE PROBLEM OF MORAL FREEDOM

Jeppe von Platz

ABSTRACT In line with a tradition running from the Ancients through Christian thought, Kant affirms the idea of moral freedom: that true freedom consists in moral self-determination. The idea of moral freedom raises the problem of moral freedom: if freedom is moral self-determination, it seems that the wicked are not free and therefore not responsible for their wrongdoings. In this essay I discuss Kant's solution to this problem. I argue that Kant distinguishes between four modalities of freedom as moral self-determination and that the problem of moral freedom disappears when these distinctions are brought to light.

1. Introduction: the idea and problem of moral freedom

The idea of moral freedom is the idea that freedom is moral self-determination.¹ This idea is at the heart of Kant's moral philosophy: epistemically speaking, "freedom and unconditional practical law reciprocally imply each other," because, ontologically speaking, "a free will and a will under moral laws are one and the same."²

The problem of moral freedom arises from the idea of moral freedom. The problem is that since the wicked are not moral, the wicked do not achieve moral self-determination and so are not free. This implication is inconsistent with our theory and practices of responsibility.³ The problem of moral freedom can thus be presented as an inconsistent triad:

1. That freedom is moral self-determination and that the wicked, therefore, are not free.
2. That freedom is necessary for moral responsibility.
3. That the wicked are responsible for their wrongdoings.

This triad is inconsistent, since it entails that the wicked are free and not free.

It thus appears that Kant faces a dilemma:⁴ either he maintains the idea of moral freedom, in which case the wicked are not free. Or he maintains that the wicked are free, in which case he has to abandon the idea of moral freedom. The first horn is inconsistent with the common sense assumptions about morality that Kant's ethics is supposed to fit, explain, and justify.⁵ The second horn requires that Kant abandons the central claim of his moral philosophy.

In this essay I present an interpretation of Kant's theory of freedom that solves this apparent dilemma. I suggest that if we interpret Kant's theory of freedom in terms of the modalities of freedom as moral self-

¹ I use this neutral term absent from Kant's writings (except for *RGV*, 6:82) to avoid assuming my suggestions about how we should interpret the central concepts (autonomy, humanity, and so on) that I present in this essay.

² *KpV*, 5:29; *GMS*, 4:447. See also *KpV*, 5:33; *GMS*, 4:441, 4:446, 4:452-3; *V-Mo/Mrong.* 29:597-99; *V-MS/Vigil*, 27:499-501, 27:505-7; *Refl.* 6076, 18:443.

³ The problem is best known from Henry Sidgwick's "The Kantian Conception of Free Will," *Mind*, 13, 51 (1888), pp. 405-412, reprinted in *Methods of Ethics*, 7th edition, (London: Macmillan, 1907), pp. 511-6. But it had already been diagnosed by Carl C. E. Schmid in *Determinismus und Freiheit* (1790) and Karl L. Reinhold in the second edition of *Briefe ueber die kantische Philosophie* (1792). (For the relevant passages in Schmid and Reinhold, see Rüdiger Bittner and Konrad Cramer, eds., *Materialien zu Kants Kritik der praktischen Vernunft* [Frankfurt am Main: Suhrkamp Verlag, 1975], pp. 249-50; 252-74). For a presentation of the reception of Kant's theory of freedom and how the problem was raised by a variety of thinkers already in the 1780s and '90s see Bojanowski, *Kant's Theorie der Freiheit: Rekonstruktion und Rehabilitierung* (Berlin: Walter de Gruyter GmbH, 2006), chapter 7. For discussions of the problem and Kant's solution to it see Bojanowski, *Kant's Theorie der Freiheit* chapter 8; Guyer, "Nature, Morality, and Peace," in *Kant on Freedom, Law, and Happiness*, 423-4; "Kant on the Theory and Practice of Autonomy," in *Kant's System of Nature and Freedom* (Oxford: Oxford University Press, 2005), 123-5; Jens Timmermann *Kant's Groundwork of the Metaphysics of Morals* (Cambridge, UK: Cambridge University Press, 2007), 164-7; Henry Allison *Kant's Theory of Freedom* (Cambridge, UK: Cambridge University Press, 1990), chapter 7.

⁴ I follow Bojanowski in stating the dilemma this strongly, see *Kants Theorie der Freiheit: Rekonstruktion Und Rehabilitierung*. (Berlin: Walter De Gruyter, 2006), 230.

⁵ As Kant says in the *Religion*: "good or evil [...] must be an effect of his free power of choice [Wirkung seiner freien Willkür sein], for otherwise they could not be imputed to him, and [...] he could be neither **morally** good nor evil." (*RGV*, 6:44)

determination, then Kant can consistently maintain that the wicked are free and responsible, yet are not free in the sense that the virtuous are free.⁶

2. The Augustinian solution

Kant is usually thought to solve the problem of moral freedom by positing an original and inscrutable choice between virtue and vice – a solution analogous to the one suggested by Augustine’s treatment of the problem. In this section I examine this suggestion. In the following sections I present an alternative.

Augustine presents the idea of moral freedom through Jesus’s claim that the truth shall set us free.⁷ The Pharisees challenged Jesus to explain, “We [...] were never in bondage to any man, how sayest thou, Ye shall be made free?” Jesus replied by distinguishing between freedom from bondage to another man and freedom from bondage to sin: “[w]hosoever committeth sin is the servant of sin.”⁸ According to Augustine, this means that only the virtuous are truly free and that the sinner “will not be free to do right, until, being freed from sin, he shall begin to be the servant of righteousness. And this is true freedom.”⁹

According to Augustine, sin is either from ignorance or from weakness of will,¹⁰ so it is unclear if the sinner had it in her power to not sin and, if she did not, how she can be held responsible for her sins. Augustine replies that ignorance and weakness of will result from an original and inscrutable free choice of sin:

When someone acts wrongly out of ignorance, or cannot do what he rightly wills to do, his actions are called sins because they have their origin in that first sin, which was committed by free will. The later sins are the just results of that first sin.¹¹

As this passage indicates, Augustine’s solution to the problem of moral freedom is to distinguish between the freedom to choose between vice and virtue and the freedom that is realized only through the choice of virtue. The kind of freedom that is realized by the choice of virtue (what I call moral freedom) is “the only genuine freedom”.¹² Nev-

⁶ It is generally recognized that Kant works with several concepts of freedom, though I am not aware of any reconstruction that focuses on the modalities of moral self-determination. See for example Lewis White Beck’s “Five Concepts of Freedom in Kant,” in Predrag Cicovacki ed. *Essays by Lewis White Beck: Five Decades as a Philosopher* (Rochester: University of Rochester Press, 1998), pp. 181-98, chapter 3; *A Commentary on Kant’s Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), chapters 2, 7, 11; Guyer, “Kant on the Theory and Practice of Autonomy”; Uleman *An Introduction to Kant’s Moral Philosophy*, chapter 4; Thomas E. Hill Jr., “Kant’s Argument for the Rationality of Moral Conduct,” *Groundwork of the Metaphysics of Morals: Critical Essays*, pp. 257-8; Jochen Bojanowski *Kants Theorie der Freiheit*; Henry Allison *Kant’s Theory of Freedom* (Cambridge, UK: Cambridge University Press, 1990), chapters 1, 3; Bernard Carnois, *The Coherence of Kant’s Doctrine of Freedom*, D. Booth transl. (Chicago: University of Chicago Press, 1987); Stephen Engstrom “The Inner Freedom of Virtue,” in Mark Timmons ed. *Kant’s Metaphysics of Morals: Interpretative Essays* (Oxford: Oxford University Press, 2002), pp. 289-315.

⁷ John, 8:32. David Foreman has recently suggested that we can understand Kant’s idea of moral freedom in light of the Stoic idea of *eleutheria*, “Kant on Moral Freedom and Moral Slavery”, *Kantian Review*, 17, 1 (2012), pp. 1-32. Regrettably, a discussion of Foreman’s illuminating article would take me too far from the main argument of this essay. We find the idea of moral freedom throughout Christian thought and its predecessors. Thus the Jewish philosopher Philo, claimed that the wicked are not free, but “every good man is free,” because “he who has God alone for his leader, he alone is free [...] all whose life is regulated by the law [of reason] are free.” (Philo, “Every Good Man is Free”, in *Philo*, vol. IX, Loeb Classical Library [Cambridge, MA: Harvard University Press, 1954], sections 45-46). Plotinus wrote that “the soul becomes free when it presses without hindrance to the Good by means of Intellect.” (Plotinus, “Free Will and the Will of the One”, Sixth Ennead, Tractate 8, in *Plotinus*, vol. VII, A.H. Armstrong transl., Loeb Classical Library [Cambridge, MA: Harvard University Press, 1988], 247). Anselm argued that “nothing is more free than right will” and that “one who sins is a slave of sin” (*On Free Will*, chapters 9 and 10), and Thomas Aquinas wrote that: “when a man passes from sin to grace, he passes from servitude to freedom,” and “since man, by his natural reason, is inclined to justice, while sin is contrary to natural reason, it follows that freedom from sin is true freedom.” (*Summa Theologiae*, II-II, Q 184, 4 and 183, 4). The idea is famously stated by Rousseau, who wrote that “moral freedom [...] alone makes man truly the master of himself; for the impulsion of mere appetite is slavery, and obedience to the law one has prescribed to oneself is freedom” (*SC*, 3:364-5/54) and that “[t]he only slave is the man who does evil, for he always does it in spite of himself.” (*E*, 4:857/473)

⁸ John, 8:34.

⁹ From *Enchiridion*, in *The works of Aurelius Augustine: A new translation*, vol. 9, 197. Likewise in *On Free Choice of the Will* (T. Williams transl. [Hackett Publishing, 1993]): “This is our freedom, when we are subject to the truth; and the truth is God himself, who frees us from death, that is, from the state of sin.” (57)

¹⁰ See *On Free Choice of the Will*, Book III.

¹¹ *On Free Choice of the Will*, 107.

¹² *On Free Choice of the Will*, 25.

ertheless, the original and inscrutable free choice between virtue and vice is sufficient for moral responsibility: “Every rational nature that was created with free choice of the will undoubtedly deserves praise if it abides in the highest and unchangeable good [...] every nature that does not abide there, and does not will to act so that it might abide there, deserves to be condemned for precisely those reasons.”¹³

Kant is usually interpreted as adopting a variety of the Augustinian solution to the problem of moral freedom. Paul Guyer, for example, suggests that we find the solution to the problem of moral freedom in Kant’s idea that persons freely choose their basic maxim – the choice that Kant in the *Religion* presents as a choice between subjecting morality to self-love or self-love to morality: “the choice made by any individual, whether to comply with both the inclination of nature and the demand of morality, or to subvert both of these and be evil, is free, but the decision to go one way rather than the other is always inexplicable or inscrutable.”¹⁴ So, according to Guyer, Kant rightly “came to retract the thesis that freedom of the will entails autonomy”¹⁵ – rightly because “ought implies can it does not imply does and indeed always leaves open the possibility of does not.”¹⁶ Guyer thus suggests that Kant came to distinguish between two different kinds of freedom: the freedom to choose between morality and self-interest, and autonomy, which is realized only through the choice of morality.¹⁷ So, where Kant in the *Groundwork* and perhaps also in the *Critique of Practical Reason* affirms the identity of freedom and autonomy, he changed the relation between freedom and autonomy in the *Religion* and the *Metaphysics of Morals*: there the free will is no longer self-determination by pure reason (autonomy), but the capacity for such self-determination. Freedom is a necessary but not sufficient condition for autonomy, and autonomy is not a fact about human beings but an ideal that we ought to achieve.¹⁸ Though I agree that we need to distinguish between the kind of freedom that is sufficient for responsibility and the kind of freedom that is realized through virtue, I disagree with the Augustinian reading of Kant.

To begin with, I doubt that Kant changes rather than clarifies his theory in the *Religion* and the *Metaphysics of Morals*.¹⁹ That is, I think that Kant consistently worked with a distinction between the capacity for self-determination (being a subject of the moral law) and actual moral self-determination (obeying the moral law). True, Kant wrote in the *Groundwork* that “freedom and the will’s own lawgiving are both autonomy and hence reciprocal concepts,” (*GMS*, 4:450) and there defined autonomy as “the will’s property of being a law to itself.” (*GMS*, 4:446) But both of these definitions of autonomy are ambiguous between autonomy as a capacity for and as an ideal of moral self-determination; between the will as capable of obeying the moral law that it gives to itself (the capacity for moral self-determination), and the will’s determining itself through the moral law (the ideal of autonomy). So, even if the distinction is not clearly drawn, the distinction is readily available to Kant already in the *Groundwork*.

A more direct reason for being skeptical of the Augustinian reading is that Kant appears to reject just this proposal in his rejection of Reinhold’s friendly suggestion for how Kant could deal with the problem of moral freedom. Reinhold’s solution starts from a distinction between two different faculties of the will: a pure legislative faculty (*Wille*) that issues the moral law and a choosing faculty (*Willkür*) that decides whether the person acts on the moral law or on self-interested motives.²⁰ Freedom of the will is this capacity for free choice between morality and self-love. In a negative sense, this freedom of choice consists in independence from sensible determination. In a positive sense, “freedom of the will is its capacity for self-determination through a choice for or against the moral law.”²¹ Pure and impure willing are, accordingly, equal expressions of free choice and thus the problem of moral freedom is solved. Yet, Kant rejects Reinhold’s solution:

¹³ *On Free Choice of the Will*, 97.

¹⁴ “Nature, Morality, and Peace”, 424.

¹⁵ “Theory and Practice of Autonomy”, 125.

¹⁶ “Nature, Morality, and Peace”, 423.

¹⁷ “Theory and Practice of Autonomy”, 125.

¹⁸ “Theory and Practice of Autonomy”, 126. Allison offers a similar interpretation of Kant’s solution to the problem of moral freedom, though Allison prefers to state it in terms of the *Wille/Willkür* distinction and a choice of basic character, see *Kant’s Theory of Freedom*, 129–45. See also Allison’s *Kant’s Groundwork for the Metaphysics of Morals: A Commentary* (Oxford University Press, 2011), 296–300.

¹⁹ Allison might agree – he prefers to call the development of Kant’s theory a “thickening” rather than a change, see *Kant’s Theory of Freedom*, 140 and in a footnote to *Kant’s Groundwork for the Metaphysics of Morals: A Commentary* Allison states that Kant had the resources for the solution already in the *Groundwork* (p. 297, n.41).

²⁰ Reinhold, “Erörterung des Begriffs von der Freiheit des Willens” in *Materialen zu Kants Kritik der Praktischen Vernunft*, R. Bittner & K. Cramer eds. (Suhrkamp, 1875), 255–6.

²¹ Reinhold, “Erörterung des Begriffs von der Freiheit des Willens”, 256 (my translation).

[F]reedom of choice [Willkür] cannot be defined – as some have tried to define it – as the ability to make a choice for or against the law [...] freedom can never be located in a rational subject's being able to choose in opposition to his (lawgiving) reason²²

So, Kant denies that freedom could be located in a choice for or against morality - *Willkür* is not a capacity for a choice of acting against the moral law, *Willkür* is, rather, the faculty that makes moral self-determination possible.

More importantly, I think that the Augustinian solution to the problem of moral freedom is poor and that Kant offers a better solution.

At first glance, it appears that the Augustinian distinction between the freedom to choose between vice and virtue and the freedom that is realized by the choice of virtue solves the problem of moral freedom: free choice is sufficient for responsibility, but true freedom is realized only through virtue. But this solution and the way Augustine, and supposedly Kant, arrives at this solution are unsatisfying. The solution is unsatisfying since it pushes the basis of moral responsibility out of the reach of theory and experience – we are supposed to rely on the existence of an original and inscrutable free choice. Moreover, the solution asserts what should be shown. What is the argument that shows that virtue and vice result from an original and inscrutable free choice? It seems that the only argument for the claim that virtue and vice are freely chosen is that we need them to be freely chosen to solve the problem of moral freedom. In other words, the basis for positing an original free choice of character is that we need to posit this choice to solve the problem of moral freedom. But then we really have not solved the problem as much as we have acknowledged that there is a problem in need of a solution.

Of course, even if unsatisfying, Kant might embrace the Augustinian solution for lack of an alternative. Yet, interpretive charity dictates that we should not read Kant as affirming the Augustinian solution, if we find a more satisfying one in his writings. In the following sections I argue that we do.

3. *How the wicked are free: humanity as the capacity for moral self-determination*

The Augustinian solution points us in the direction of a solution to the problem of moral freedom: to solve the problem we need a distinction between kinds of freedom, so that the wicked can be free in one sense and unfree in another. But we need independent arguments for why we should accept this distinction between kinds of freedom and a justification for why they apply to human beings in a manner so that the wicked are free in the sense relevant for moral responsibility, yet are not free in the sense that the virtuous are free.

A place to start is the already indicated distinction between freedom as a capacity for moral self-determination and the ideal of moral self-determination that is the ideal of autonomy. If the capacity for moral self-determination is sufficient for moral responsibility, then the wicked and the virtuous alike are free and responsible. Does Kant offer a concept of freedom as the capacity for moral self-determination that fits this description? In this section I suggest that he does; it is the concept of humanity.

In the *Groundwork*, Kant defines humanity as rational nature.²³ Rational nature differs from everything else in nature in that it acts in the pursuit of ends: “[r]ational nature is distinguished from the rest of nature by this, that it sets itself an end.”²⁴ Kant echoes this definition in the *Metaphysics of Morals*: “[t]he capacity to set oneself an end – any end whatever – is what characterizes humanity.”²⁵

For rational but embodied beings like us, even the pursuit of the satisfaction of an inclination makes an end of the satisfaction of the inclination: “freedom of the power of choice [Willkür] [...] cannot be determined to action through any incentive except so far as the human being has incorporated it into his maxim.”²⁶ Indeed, *all and only* the actions of free beings are in the pursuit of ends: “[a]n **end** is an *object* of free choice, the representation of which determines it to action [...] since no one can have an end without *himself* making the object of his choice into an end,

²² *MdS*, 6:226. Bojanowski and Timmermann agree that Kant here rejects Reinhold's suggested solution. See Bojanowski *Kants Theorie der Freiheit*, 245-7; Timmermann *Kant's Groundwork of the Metaphysics of Morals*, 164-5.

²³ *GMS*, 4:428-9, 4:431, 4:437.

²⁴ *GMS*, 4:437. Another way to distinguish between animals and humans in terms ends is to note that ends are conceptual. An end is “the concept of an object insofar as it at the same time contains the ground of the reality of this object.” (*KU*, 5:180)

²⁵ *MdS*, 6:392. See also *MdS*, 6:387; *V-MS/Vigil.*, 27:671.

²⁶ *RGV*, 6:24, italics removed. Likewise in *KpV*: “it [the will] is never determined directly by the object and the representation of it, but is instead a faculty of making a rule of reason the motive of an action.” (*KpV*, 5:60) Allison, Reath, Engstrom, Wood, Guyer, and McCarty all draw attention to this aspect of Kant's theory of action (the so-called “incorporation thesis”), see Allison *Kant's Theory of Freedom*, 39-40; 51-2, 189.126-7; Reath “Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination,” in *Agency and Autonomy in Kant's Moral Theory* (Oxford: Oxford University Press, 2006), pp. 8-32, 17-8; Engstrom *The Form of Practical Knowledge*, 33; Wood's *Kant's Ethical Thought*, 51-3; Guyer *Kant on Freedom, Law, and Happiness* (Cambridge, UK: Cambridge University Press, 2000), 294-8; McCarty *Kant's Theory of Action*, 71-87.

to have any end of action whatsoever is an act of *freedom* on the part of the acting subject.” (*MdS*, 6:384-5) So, any action of a being with humanity is in the pursuit of an end and is, therefore, a determination of the will to act for reasons.

Humanity is the capacity to set and pursue ends and this capacity has two dimensions corresponding to Kant’s characterizations of the negative and positive aspects of freedom: the capacity to set and pursue ends involves both freedom from having one’s will determined by one’s natural inclinations (independence from sensible determination) and the capacity to freely set and pursue ends (the capacity for self-determination).²⁷ Since any setting of an end is an act of freedom, and since the central claim of Kant’s practical philosophy is that freedom and morality mutually imply each other, it seems fair to infer that humanity involves both independence from sensible determination and the capacity for *moral* self-determination; freedom, humanity, and moral subjectivity are in this sense reciprocal concepts. Thus, humanity is the *capacity* for moral self-determination. Moreover, since any being with humanity freely sets and pursues her ends subject to the requirements of morality, this capacity is sufficient for moral personality: humanity is sufficient both for being an object that must be regarded and treated with respect and for being held accountable for one’s doings.

While my interpretation of Kant’s concept of humanity is in agreement with most interpreters,²⁸ it conflicts with two other interpretations. To elaborate and defend my interpretation I shall briefly discuss these.

First, Richard Dean has proposed that humanity is the good will.²⁹ Dean bases this reading on Kant’s claim that the good will is the only thing good without qualification read together with Kant’s argument that humanity, and only humanity, must be treated always as an end in itself and never merely as a means. Dean argues that the only way to avoid inconsistency between these claims is if the good will is humanity.³⁰ I find it hard to accept Dean’s suggestion. For one thing, Kant says that all rational beings are ends in themselves: “every rational being [jedes vernünftige Wesen] exists as an end in itself, *not merely as a means*” (*GMS*, 4:428), but not all rational beings have a good will, so not all things that are ends in themselves have a good will. It is, moreover, clear that Kant equates humanity and rational nature, both from the passages quoted above where Kant identifies rational nature and humanity as the capacity to set and pursue ends, but also from how the just quoted sentence echoes the formula of humanity as well as from the context of the sentence, which leads to the inference that since “*rational nature exists as an end in itself*. [...] The practical imperative will therefore be the following: *So act that you use humanity, whether in your own person or in the person of any other, always at the same time as an end, never merely as a means.*” (4:429) So, humanity is rational nature, which is a property of all (or at least all minimally mature) human beings, which explains Kant’s remark in the *Metaphysics of Morals* that “Every human being has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other.” (*MdS*, 6:462). Even a wicked person is a rational being and must, as such, be treated with respect: “humanity [...] is worthy of respect, and even though somebody may be a bad man, the humanity in his person is entitled to respect.”³¹ In short, humanity is coextensive with rational being, not all rational beings have a good will, so humanity is not the good will.

²⁷ See also *KpV*, 5:33; *GMS*, 4:441, 4:446; *V-MS/Vigil.*, 27:499, 27:505; *Refl.*, 6076, 18:443.

²⁸ My understanding of Kant’s concept of humanity is a version of the standard reading – variants of it are presented by: Christing Korsgaard *Creating the Kingdom of Ends* (Cambridge, UK: Cambridge University Press, 1996), 17, 110-1, 346; Thomas E. Hill Jr. *Dignity and Practical Reason in Kant’s Moral Theory* (Ithaca: Cornell University Press, 1992), 40; Onora O’Neill *Constructions of Reason* (Cambridge, UK: Cambridge University Press, 1989), 137-8; Paul Guyer *Kant* (London: Routledge, 2006), 186-7; Allison *Kant’s Groundwork for the Metaphysics of Morals: A Commentary*, 215-8.

²⁹ Richard Dean *The Value of Humanity in Kant’s Moral Theory* (Oxford: Clarendon Press, 2006).

³⁰ Dean *The Value of Humanity in Kant’s Moral Theory*, 38-40.

³¹ *V-Mo/Collins*, 27:373. Yet, what about Kant’s claim that humanity has dignity and that the claim to dignity, a worth beyond all price, the claim of “a morally good disposition, or virtue” (*GMS*, 4:435) and that there is a certain “sublimity and dignity in the person who fulfills all his duties [...] there is indeed no sublimity in him insofar as he is *subject* to the moral law, but there certainly is insofar as he is at the same time *lawgiving* with respect to it and only for that reason subjected to it” (*GMS*, 4:439-40)? Do these passages not indicate that dignity presupposes a good will and thus that humanity, as the object of dignity, is the good will? In reply: Kant’s position on dignity is ambiguous between dignity being a property of all rational beings and of only the good willed. Thus, the paragraph that the previous quote is taken from concludes: “the dignity of humanity consists just in this capacity to give universal law, though with the condition of also being itself subject to this very lawgiving” (*GMS*, 4:440, my emphasis) and earlier he in the *Groundwork* Kant wrote that “morality, and humanity insofar as it is capable of morality, is that which alone has dignity.” (*GMS*, 4:435, my emphasis) A similar reply goes to Timmermann’s suggestion that Kant’s discussion of servility in the *Metaphysics of Morals* suggest that immoral actions divest a person of her dignity (*Kant’s Groundwork of the Metaphysics of Morals: A Commentary*, 108). Kant there wrote that “a subject of a morally practical reason [...] possesses a dignity (an absolute inner worth) by which he exacts *respect* for himself from all other rational beings in the world.” (*MdS*, 6:434-5) Yet, being a subject of morally practical reason is not the same as being moral: there is, again, a difference between being a subject of pure practical reason and obeying its commands. The ambiguity of Kant’s position is illustrated by Timmermann’s ambi-

Second, following a passage in the *Religion* (RGV, 6:26-8), Wood has suggested that humanity is a “capacity for setting ends through reason” that has “no specific reference to morality.”³² In the relevant passage Kant contrasts predispositions of animality, humanity, and personality. Where personality implies humanity, the converse implication does not hold. In a footnote Kant even emphasizes that the concept of humanity does not imply responsibility for “from the fact that a being has reason does not follow that [...] this reason contains a faculty of determining the power of choice unconditionally, and hence to be “practical” on its own.” (RGV, 6:26n) This passage is doubly troublesome for my suggestion that humanity is the capacity for moral self-determination sufficient for moral personality: first, it appears to clearly state that humanity is a capacity for rationality and not for moral self-determination, and, second, it explicitly rejects that humanity is sufficient for moral responsibility, which is what I claim.

It is hard to make sense of this passage. Kant is, I believe, playing with Rousseau’s distinction between different sources of incentives: *amour de soi-meme*, *amour-propre*, and conscience. We thus find Kant elaborating on the predispositions to humanity as a physical and comparative self-love that is the origin of “the inclination to gain worth in the opinion of others” and which is the source of the “vices of culture” – envy, ingratitude, and *Schadenfreude* (RGV, 6:27). If Kant is playing with Rousseau, it may be that we should interpret the passage as drawing attention to the different layers of human nature of relevance to understanding the social sources of human evil and not as defining the concepts of humanity and personality.

It is, in any case, impossible to find a consistent reading of the contrast between humanity and personality drawn in the *Religion* and the way that Kant relates humanity and personality in his other ethical writings. Kant defines a person as “a subject whose actions can be imputed to him. Moral personality is therefore nothing other than the freedom of a rational being under moral laws.”³³ So, personality implies freedom and moral subjectivity, the capacity for moral self-determination. So far, so compatible with Wood’s reading. Yet, in the *Groundwork* Kant defines personality thus: “Rational beings are called *persons* because their nature already marks them out as an end in itself, that is, as something that may not be used merely as a means [...] (and is an object of respect).” (GMS, 428) So, rational beings are called persons, because they are ends in themselves and as such must be respected and treated never merely as means, but always as ends in themselves. And then we return to Kant’s claim that there is only one thing in nature that is an end in itself in this sense: humanity. So, rational beings are called persons, because they have humanity and this would indicate, not only that personality implies humanity, but also that humanity implies personality.³⁴

In addition we should ask: Why is humanity the only end in itself?³⁵ According to Kant, rationality as such is not the source of dignity or a proper object of respect. As Kant says in the second *Critique*: “the human being (and with him every rational being) is an *end in itself* [...] humanity in our person must, accordingly, be *holy* to ourselves: for he is the *subject of the moral law* and so of that which is holy in itself. (KpV, 5:131-2) Here Kant indicates that it is humanity as moral subjectivity and not mere rationality that makes human beings ends in themselves and thus proper objects of respect. This reading is confirmed by what Kant is reported to have said in lectures given at the same time as he was working on the *Groundwork*: “If only rational beings are capable of being ends in themselves, it cannot be because they have reason but because they have freedom. [...] Reason does not give us dignity. [...] freedom and freedom alone warrants that we are ends in ourselves.”³⁶ Since freedom implies moral subjectivity it follows that it is humanity understood as the capacity for moral self-determination that is the basis of dignity, a conclusion that Kant affirms in the *Groundwork*: “the dignity of humanity consists just in this capacity to give universal law, though with the condition of also being itself subject to this very lawgiving.” (GMS, 4:440)

So, though my interpretation fails to make sense of a passage in the religion, it is supported by several other central passages in Kant’s ethical writings that contradict Wood’s reading.

4. Virtue and virtuousness

alent claim that “a person possesses dignity in so far as he or she is able to lead – i.e. succeeds in leading – a moral life.” (*Kant’s Groundwork of the Metaphysics of Morals*, 109, n119)

³² Allen W. Wood *Kant’s Ethical Thought* (Cambridge University Press, 1999), 119; likewise *Kantian Ethics* (Cambridge University Press, 2008), 88.

³³ *MdS*, 6:223. Likewise in the second *Critique*: “freedom and independence from the mechanism of the whole of nature, regarded nevertheless as also a capacity of a being subject to special laws – namely pure practical laws given by his own reason.” (KpV, 5:87). See also *GMS*, 4:428; *RGV*, 6:26; *Anth.*, 7:324; *Refl.*, 6713, 19:139

³⁴ Wood agrees: “Kant writes as if humanity and personality are necessarily coextensive. I think they *are* necessarily coextensive.” (*Kantian Ethics*, 94)

³⁵ I owe this line of argument to Kate Moran.

³⁶ *V-Na/Feyer*, 27:1321-2, my translation. The idea and the reference of this paragraph are Paul Guyer’s.

In the previous section I argued that the wicked and the virtuous alike are free and responsible, because they have humanity: the capacity for moral self-determination sufficient for moral personality. Yet, we still need to identify the sense in which the virtuous are free and the wicked are not free. For that we need to take a look at Kant's theory of virtue.³⁷

In line with a tradition running from the ancients to Rousseau, Kant defines virtue as a kind of "moral strength of the will," (*MdS*, 405) a "*fortitudo moralis*," (*MdS*, 6:380) or "a firmly grounded disposition to fulfill one's duty strictly." (*RGV*, 6:23n) But virtue is not merely the disposition to do what duty requires; habitual or instinctual conformity with duty is not yet virtue.³⁸ Virtue is, rather, the disposition to do what duty requires, *because* it is what duty requires.³⁹ In addition, virtue must be distinguished from holiness: a holy being is disposed to always do the right thing, because it is the right thing; "their will is adequate to the law." (*V-Mo/Collins*, 27:463) Such is not the condition of humanity. Embodied rational beings like us are inevitably subject to counter-moral interests, and it is in the struggle with this "rabble element in man" (*V-Mo/Collins*, 27:360) that virtue is required and shows its worth. In this sense virtue is not simply moral willing or moral disposition, but "moral disposition *in conflict*."⁴⁰ The moral strength that is virtue, then, is the strength of moral self-mastery: "virtue [...] means strength in mastering and overcoming oneself, in regard to the moral disposition",⁴¹ that results in "a moral preparedness to withstand all temptations to evil, so far as they arise from inclinations."⁴² Bringing all these features together, we can initially define virtue as a disposition or kind of character that makes rational but embodied agents do what they ought to do, because it is what they ought to do, even when it conflicts with their self-interests.

We get a more precise definition of virtue by looking at its ingredients: a good will's self-mastery in the principled performance of duty. Virtue should not be confused with either of these alone; virtue is not mere good will, conformity with duty on principle, or self-mastery.

To start with, virtue is not the good will. The good will is a principled commitment to give priority to morality over self-interest and, therefore, do what is right even when it conflicts with the imperatives of self-love.⁴³ One indication that the good will and virtue are different things is that the good will does not presuppose the presence of obstacles to overcome – the good will is not necessarily embattled; a holy will is good, but a holy being cannot be virtuous in the strict sense.⁴⁴ Another difference is that it is possible, so Kant maintains, to want to do what is right, because it is right, and nevertheless fail to do what is right – and, therefore, fail to be virtuous (e.g. *V-MS/Vigil*, 27:570-1). Two such sorts of failures come to mind:⁴⁵ first, the Socratic failure, where a person is mistaken about what she ought to do, in which case she fails to do her duty, in spite of her good intentions and in spite of having enough strength of will to carry out these intentions. Second, the Aristotelian failure, where a person knows what she ought to do and would really prefer to do it, but nevertheless lacks the moral fiber to actually do it. In each of these cases of moral failure – from ignorance or weakness of will – the person may have a good will, but nevertheless lack virtue, because she lacks one of its other two ingredients. So, I suggest the following distinction between the good will and virtue: the good will is the will that has the proper ordering of basic maxims so that the agent is genuinely committed to give her moral interests priority over interests of self-love and is therefore willing to do the

³⁷ The following draws on Engstrom "The Inner Freedom of Virtue;" Lara Denis "Kant's Conception of Virtue," in P. Guyer ed. *The Cambridge Companion for Kant and Modern Philosophy* (Cambridge, UK: Cambridge University Press, 2006), pp. 505-537 and *Moral Self-Regard: Duties to Oneself in Kant's Moral Theory* (Garland Publishing Inc., 2001), 163-71; Paul Guyer "Moral Worth, Virtue, and Merit," in *Kant on Freedom Law and Happiness*, chapter 9; Allen Wood *Kant's Ethical Thought*, 329-33; Anne Margaret Baxley's study *Kant's Theory of Virtue: The Value of Autocracy* (Cambridge, UK: Cambridge University Press, 2010), chapters 1-2.

³⁸ *MdS*, 6:383-4; *V-Mo/Collins*, 27:463; *MS-Vigil*, 27:492; *Anth*, 7:147. Kant, accordingly, distinguishes virtue from "kindness of heart," which is an instinctual disposition to well-doing: "One may have kindness of heart without virtue, for the latter is good conduct from principles, not instinct." (*V-Mo/Collins*, 27:463).

³⁹ *MdS*, 6:387, 6:394. To clarify: it is not enough that a person acts in conformity *as a rule*, nor that she does it *because it is a rule* (cf. Hart's distinction in *The Concept of Law* [Oxford: Oxford University Press, 1994], 55-7). She must do it *out of respect for the rule*.

⁴⁰ *KpV*, 5:84. Likewise in the Vigilantius lecture, *V-MS/Vigil*, 27:489-92, and in the lectures on logic, *V-Lo/Dohna*, 24:728.

⁴¹ *V-Mo/Collins*, 27:300. Likewise *V-Mo/Collins*, 27:465; *V-MS/Vigil*, 27:570-1; *Anth*, 7:147. See also Baxley's discussion of the relation between virtue and autocracy in *Kant's Theory of Virtue*, chapter 2.

⁴² *V-Mo/Mron II*, 29:604. Likewise *V-MS/Vigil*, 27:570; *MdS*, 6:394.

⁴³ This understanding of the good will should not be too controversial, it is a variant of the one suggested by Allison in *Kant's Groundwork of the Metaphysics of Morals: a Commentary*, chapter 3 and Thomas E. Hill Jr. in "Is a Good Will Overrated?" *Human Welfare and Moral Worth* (Oxford: Oxford University Press, 2002), chapter 2.

⁴⁴ For a helpful discussion of the difference between virtue and holiness, see Baxley *Kant's Theory of Virtue*, 54-7.

⁴⁵ Note: to my knowledge Kant does not distinguish between these two, and it is likely that he would consider the Socratic failure a true moral failure, since wrongdoing here follows from a failure of the understanding and not of the will.

right thing, come what may. Virtue is the good will plus correctness of judgment about what morality requires plus strength of will to carry the good will into practice.⁴⁶

Second, Kant draws a distinction between *virtus phaenomenon* and *virtus noumenon*, a distinction that clarifies that a commitment to doing one's duty is not yet virtue. A person has phenomenal virtue or "virtue in appearance" (*V-MS/Vigil*, 27:583) if she always does what duty requires, because she is committed to conform with the requirements of duty, which secures outward conformity with duty and even makes it appear to the agent herself that she has a principled allegiance to virtue.⁴⁷ Yet, even when this commitment to duty is coupled with strength of will, it is not yet virtue, for the agent may lack a good will. Indirect egoism is still egoism; the shop-keeper that maintains a commitment to honesty because honesty is the prudential strategy does not have a good will and so is not virtuous – the same shop-keeper would be committed to dishonesty, should that be what prudence dictates. A good will requires that the agent is committed to duty and acts as she ought to, not only as a matter of principle, but out of a principled commitment to morality – from a recognition of the authority of the moral law (cf. *GMS*, 4:397; *KpV*, 5:81). Only an agent with such a commitment to morality for the right reasons has a good will and so can really be virtuous.

Third, virtue is not mere self-mastery. Self-mastery, generically speaking, is strength of will: to be in control of one's sensible inclinations and therefore act from reason.⁴⁸ The result of self-mastery is a sort of integrity of the agent: she can lay plans and carry these out without being determined or deterred by affects or passions. But self-mastery can serve self-love as well as morality.⁴⁹ Of course, acting according to a principle of self-love entails a loss of self-control: the principle of self-love is a principle of heteronomy, since the interest in happiness unrestricted by the moral principle is an interest in the satisfaction of sensible desires. Thus, *real* self-control is moral and prudential self-control is only an "analogue of self-mastery."⁵⁰ Yet, the point remains that self-mastery alone is not virtue, for virtue is self-mastery in the service of a good will.

We can summarize what virtue is not, by restating what virtue is. The ideal of virtue combines three elements: understanding of what morality requires (moral understanding), the principled commitment to do it because it is what morality requires (a good will), and strength of will to do it (self-mastery).

Next, we need to clarify how Kant uses the concept of virtue in two distinct ways.⁵¹ First, considered as an ideal, virtue is the ideal of moral perfection relative to our condition: "[v]irtue is the moral perfection of man." (*V-Mo/Collins*, 27:465) The ideal of virtue is an ideal of complete possession of the ingredients of virtue: a well-informed good will and strength of will to carry it out. Second, virtue as virtuousness: the degree to which a person has actually achieved the ideal of virtue.⁵² The degree of virtuousness of a person will be a function of the three variables: her moral understanding, the goodness of her will, and the strength of her will.

The ideal of virtue is an ideal of complete moral self-control: the virtuous person is generally governed by the moral law that she gives to herself and never determined by her counter-moral inclinations. It follows that the ideal of virtue is as close to the ideal of complete moral self-determination that embodied rational beings like us could get. It also follows that the degree to which a person has approximated the ideal of virtue, i.e. her virtuousness, is her degree of *actual* moral self-determination. And it thus follows that virtuousness is the measure of freedom understood as the actual moral self-determination of an embodied agent.

Kant does not always state these connections between virtue and freedom as clearly as one could wish. But in notes from Kant's lectures on ethics we find them expressed clearly enough: "[t]he more a man is virtuous, the more free he is." (*V-Mo/Collins*, 27:464) And in the lectures on logic, where Kant is reported to have used virtue as an example of a pure concept of reason, we read that "[v]irtue is a readiness, a freedom in action." (*V-Lo/Wiener*, 24:923) And, if we recall that virtue is moral self-mastery, the claim is stated clearly enough also in the *Metaphysics of Morals*: "the more [a human being] can be constrained morally [...], so much the more free he is," so that "only

⁴⁶ For a conflicting interpretation, see Richard McCarty *Kant's Theory of Action* 6-7.

⁴⁷ *MdS*, 6:219, 6:225, 6:394; *KpV*, 5:81; *RGV*, 6:14, 6:46-7; *V-MS/Vigil*, 27:583, 27:706.

⁴⁸ Kant's discusses self-mastery at length in the Collins lectures on ethics, see *V-Mo/Collins*, 27:360-9.

⁴⁹ Such "coolness of the scoundrel" makes him especially abominable (*GMS*, 4:394).

⁵⁰ *V-Mo/Collins*, 27:362. Relevant here is also this passage from the Collins lectures: "where there is no self-mastery, there is anarchy. Yet even if there is moral anarchy in man, prudence still steps in to replace morality, and reigns in its stead." (*V-Mo/Collins*, 27:362) Kant sometimes calls moral self-mastery "autocracy", perhaps to indicate that in moral self-mastery pure reason has acquired executive in addition to legislative power of the will (see e.g. *MdS*, 6:383-4; *V-Mo/Collins*, 27:362, 27:364, 27:368; *V-Mo/Mron II*, 29:626; *V-MS/Vigil*, 27:653).

⁵¹ In *Kant's Theory of Action*, 227-30, Richard McCarty similarly suggests that Kant has a distinction between virtue as ideal and the degree of virtue of a person, but the interpretation he offers in defense of this claim conflicts with mine.

⁵² Kant at times talks of this degree of virtue as a degree of moral health or robustness of persons (e.g. *MdS*, 6:385, 6:397, 6:405).

in its [virtue's] possession is he [the human being] "free".⁵³ A sentiment echoed in the lectures: "the more a person practices self-compulsion, the freer he becomes." (*V-Mo/Collins*, 27:269)

Virtue and virtuousness as modalities of moral self-determination differ from the modalities of humanity and the ideal of autonomy. *Humanity* is the *capacity* (or possibility) for moral self-determination that does not come in degrees, but is equally the capacity of the wicked and the virtuous. *Virtuousness*, by contrast, is the *actual* degree of moral self-determination that a person has acquired: "while the capacity (*facultas*) to overcome all opposing sensible impulses can and must be simply *presupposed* in man on account of his freedom, yet this capacity as *strength* (*robur*) is something he must acquire."⁵⁴ And the virtuousness of a person is a matter of degree: the degree to which she has acquired this strength or moral character. As the fullest or complete acquisition of moral strength of will possible for beings like us, the *ideal of virtue* is the highest degree of moral self-determination we can realistically aspire to achieve; what we ought to make of our humanity.

The ideal of virtue is akin to the ideal of autonomy in that it is an ideal of complete moral self-determination. Yet, for human agents, even if we could reach the ideal of virtue (which we cannot⁵⁵), moral self-determination would still be precarious and contingent and so fall short of the ideal of autonomy. I believe that this contingency of virtue is the underlying idea when Kant says that "[v]irtue is always *in progress* and yet always starts *from the beginning*." (*MdS*, 6:409) Virtue is always in progress, because no person could ever fully reach the ideal of virtue. And it always starts from the beginning, because no person is ever free from the temptations of self-interest, and so "virtue can never settle down in peace and quiet." (*MdS*, 6:409)

The distinction between the ideal of virtue and the ideal of autonomy is analogous to the distinction between the ideal of virtue and holiness of will. A holy will is autonomous and the ideal of virtue is as close to holiness that human beings can hope to get: "[a]ngels in heaven may be holy, but man can only get so far as to be virtuous."⁵⁶ Kant is, however, quite pessimistic about the possibility of reaching the ideal of virtue: "virtue is an idea and nobody can possess true virtue."⁵⁷ So, holiness and the ideal of virtue are both ideals that no human being can reach, but their impossibilities differ in kind. A holy will is not sensibly affected and it is, as such, metaphysically impossible for a rational but embodied being to achieve holiness: "[h]oliness is the absolute or unlimited perfection of the will [...] the human being can *never* be *holy*."⁵⁸ It is not, by contrast, metaphysically impossible for a human being to reach the ideal of virtue. Yet, given facts about the world and the frailty of human nature, no person could achieve full virtue (at least not in this life⁵⁹). The difference in impossibility is like the difference between my travelling a mile without moving in time (metaphysically impossible) and my running a mile in five seconds (cannot happen, though it is not metaphysically impossible). Moreover, since holiness is metaphysically impossible, we cannot approximate holiness of will.⁶⁰ But we can and should approximate the ideal of virtue: virtue "is an ideal and unattainable, while yet constant approximation to it is a duty."⁶¹ And, even though we shall never achieve full virtue, continuous approximation to virtue is good in itself, for it increases the moral self-determination of the approximating agent. And so, to repeat, "[t]he more a man is virtuous, the more free he is." (*V-Mo/Collins*, 27:464)

5. Varieties of moral failure and how the wicked are both free and not free

I have indicated that every agent, no matter if she is wicked or virtuous, is free in the sense that she has the capacity for moral self-determination that is humanity. So, in this sense the wicked are free. But we still need to identify and

⁵³ *MdS*, 6:382, *MdS*, 6:405.

⁵⁴ *MdS*, 6:397. Likewise: "What follows from the constant clashing of the sensory and moral natures of man is simply that morality is not inborn, but only the capacity for arriving at it, and creating principles for ourselves; self-mastery, however must be acquired, in that we constantly resist our temperament or disposition [Sinnesart] by the laws of reason, and form principles of actions; and this acquisition gives a man character." (*V-MS/Vigil.*, 27:570-1)

⁵⁵ "Human virtue is always imperfect" *VRL*, 28:994.

⁵⁶ *V-Mo/Collins*, 4:465; likewise *KpV*, 5:122-3, 5:128. The qualitative difference between angelic virtue and human virtue and devilish vice and human vice is elaborated in *V-MS/Vigil.*, 27:691.

⁵⁷ *V-Mo/Collins*, 27:463. In the lectures on logic Kant uses virtue as example of an idea: "[w]e cannot encounter virtue among men. But my reason must nonetheless have a concept of virtue, as it must be in its complete perfection. We can perceive virtue in experience. But much must still be added[.] thus it is an idea." (*V-Lo/Wiener*, 24:906)

⁵⁸ *V-Phil-Th/Pöhlitz*, 28:1075, likewise in the *Vigilantius* lectures, *V-MS/Vigil.*, 27:490-1.

⁵⁹ In the second *Critique* and later writings Kant thought that he could warrant rational faith in immortality by showing that immortality was necessary for open-ended progress towards virtue (e.g. *KpV*, 5:122-3).

⁶⁰ Kant says in the *Religion* that "holiness is above all the goal for which the human being should strive," (*RGV*, 6:159), but a couple of pages later he adds that "'yet [...] the striving itself is called *virtue*.'" (*RGV*, 6:161).

⁶¹ *MdS*, 6:409. Likewise *MdS*, 6:446. The same sentiment that virtue is beyond reach, yet an ideal that we should approximate, is expressed in Kant's famous remark that "[o]ut of such crooked wood as the human being is made, nothing entirely straight can be fabricated. Only the approximation to this idea is laid upon us by nature." (*IaG*, 8:23)

articulate the sense in which the virtuous are free and the wicked are not free. We saw that the solution is not that the virtuous realize the ideal of autonomy and that the wicked do not, for virtuous and wicked alike fail to reach this ideal. I have, moreover, argued that virtuousness is the degree to which an agent has actually achieved moral self-determination. It follows that any agent, wicked or virtuous, has a degree of virtuousness. And it looks like this conclusion presents a further problem, for if the difference between the virtuous and the wicked is a matter of their degrees of virtuousness, then it seems that there is no conceptual space for a distinction in kind that would allow Kant to maintain that the virtuous have a kind of freedom that the wicked do not. In this section I try to identify the distinction in kind between the virtuous and the wicked and show why the virtuous are free in a sense that the wicked are not.

Wickedness is, of course, the opposite of virtuousness. But Kant carefully distinguishes between different varieties of moral failure and the sorts of opposition to virtue they present. So, we need to take a closer look at Kant's analysis of the varieties of moral failure to understand the distinctive kind of moral failing that is wickedness.

To begin, Kant identifies an important structural difference between varieties of moral failure that result from weakness of will and varieties of moral failure that result from badness of will. Kant states this distinction as follows: "Virtue = +a is opposed to *negative lack of virtue* (moral weakness = 0) as its *logical opposite* [...] but it is opposed to vice = -a as its *real opposite*."⁶² This is classic Kant: at first blush the distinction appears obscure and pointless, but at a closer look it is both clear and insightful.

The distinction is between two kinds of opposition: the opposition between virtue and weakness of will and the opposition between virtue and vice. Weakness of will (*defectus moralis*, mediocrity) is the condition where a person with a good will lacks the strength of will to carry through her commitment to morality and can be measured by the strength of temptations of self-love that a person can overcome. There are different varieties of moral failure from weakness of will: *impurity* of the will, where the agent respects the moral law, but nevertheless fails to give morality priority unless doing so is also attached with other incentives, and *frailty* of the will, where the agent has a sincere and pure commitment to morality, but lacks the self-mastery needed to overcome counter-moral inclinations (*RGV*, 6:29-30).

The varieties of weakness of will are on a continuum with the virtuous disposition. Assuming a good will, the scale of virtuousness goes from complete weakness of moral willing to complete strength of moral willing. At the extreme weakness end of the scale we find absence of virtue: "[n]o virtue can be found in human beings where a degree of temptation cannot be found which they can overcome." (*V-Na/Feyer*, 27:1323) At the other end we find the ideal of virtue that is complete moral self-determination. Though extreme weakness of will thus presents an opposite of virtue, it is only quantitatively different from the ideal of virtue (extreme weakness of will is the "logical opposite" of virtue).

Vice or wickedness is a different variety of moral failing, a different kind of opposite to virtue. Where failures from weakness of will are moral failures where a good will is coupled with insufficient strength of will, wickedness is the variety of moral failure where the person has a bad will.⁶³ The goodness or badness of a person's will boils down to whether or not the person has respect for morality or self-love as her basic commitment. The good will, again, is a basic commitment to do what is right, even when it conflicts with one's self-interest. The opposite of the good will is a bad will, which is committed to give priority to self-love if or when there is a conflict between what morality and self-love dictate (*RGV*, 6:36). According to Kant, any agent has either a good will or a bad will (exclusive disjunct and excluded middle), for there is no alternative to giving priority to either self-love or moral interests as a matter basic maxim. This is Kant's declared rigorism: "between the evil and a good disposition [...] there is no intermediate position." (*RGV*, 6:23n) A bad will gives priority to self-love over morality if and when these interests conflict. Having a bad will, wickedness, is thus qualitatively different from virtue (wickedness is the real opposite of virtue).

Thus, Kant's distinction between wickedness and weakness of will in terms of how they oppose virtue is quite clear. Since there is no middle ground, no continuum, between a good and a bad will, there is no continuum between wickedness and virtue (for virtue presupposes a good will). Weakness of will, by contrast, is the lack of the strength of will to carry out an otherwise principled commitment to morality – and weakness of will is thus a matter of degree. In other words, the difference between wickedness and virtue is qualitative, whereas the difference between weakness of will and virtue is quantitative: "the man who is not virtuous is not yet vicious on that account; he is merely lacking in virtue [...] vice is something positive. The want of virtue is mediocrity, but the contempt for

⁶² *MdS*, 6:384, see also *RGV*, 6:23n. Translation amended to remove a type-o.

⁶³ "See also *RGV*, 6:29-31. I here pursue the connection between vice as such and the various sorts of vices (bestial, human, devilish) that Kant is reported to have considered at great length in his lectures on ethics; cf. *V-Mo/Collins*, 27:439-40; *V-MS/Vigil*, 27:691-9.

moral laws is vice.”⁶⁴ Extreme weakness of will resembles wickedness, for both vice and complete weakness of will are ‘opposites’ of the ideal of virtue. Yet, resemblances deceive, for vice and weakness of will result in a zero degree of virtuousness for different reasons (I illustrate the difference shortly).

We can summarize the preceding by distinguishing between four basic varieties of moral failure. First, failure of moral understanding can lead to ignorance of one’s duty. Since this sort of failure is a failure of the understanding rather than of the will, these are not moral failures in the strict sense. Next, there are two failures of weakness of the will, namely, either (ii) impurity of will or (iii) frailty of will. While these are failures of the will, they presuppose the presence of a good will (a commitment to morality) and so are not as bad as the fourth kind: (iv) failures of the will’s basic maxim that give priority to self-love over virtue (depravity, vice, wickedness).⁶⁵ We can even imagine a fifth sort of failure that is worse than wickedness, namely, the principled commitment to disobey the moral principle, even when the moral principle coheres with the principle of self-love, but Kant appears to think that such a diabolical disposition is foreign to human nature.⁶⁶

With the above analysis of the varieties of moral failing we are finally in a position to understand the sense in which the virtuous are free and the wicked are not free. Every agent has a degree of virtuousness and that this degree correlates strictly with the degree to which the person has actually achieved the kind of freedom that is moral self-determination. The solution to how the virtuous and not the wicked are free is that the wicked have a zero degree of virtuousness and, therefore, do not actually achieve any degree of freedom that is actual moral self-determination, that is, moral freedom. Hence, the wicked are not free in the sense that those who have a positive degree of virtuousness are, for the wicked fail to achieve a positive degree of moral self-determination.

With one slight change, we can use Kant’s mathematics to illustrate why the wicked are not free in the sense that the virtuous are free.⁶⁷ The virtuousness of a person, which is also her degree of moral self-determination (moral freedom), is the product of her goodness of will and her strength of will. The will is either good (=1) or bad (=0), whereas the strength of will is a matter of degree, so that the strength of will of a person lies somewhere on the continuum between full strength of will (=1) and complete weakness of will (=0). So, the virtuousness and moral freedom of a person is a product of goodness of will (good=1 or bad=0) and strength of will (between 1 and 0). The wicked have a zero degree of virtuousness and moral freedom, because they have a bad will (= 0), which means that their virtuousness is zero no matter what their strength of will is ($0 \times \text{strength of will} = 0$). A virtuous person, by contrast has a good will (=1) and some degree of strength of will, and thus has some positive measure of virtuousness and moral self-determination ($= 1 \times \text{strength of will}$).

Of course, in extreme weakness of will, the weak and the wicked are equally lacking in virtuousness and thus are equally unfree. A person suffering from extreme weakness of will has zero degree of virtuousness, because her strength of will is zero, which means that, in spite of her good will (=1), she fails to achieve any positive degree of moral self-determination ($1 \times 0 = 0$). Thus, in the extreme a person that suffers extreme weakness of will is as unfree as a wicked person, but she is so for different reasons (which could explain why the wicked are more blameworthy than the weak).

So, the wicked are not free in the sense that the virtuous are free, since the wicked have a zero degree of actual moral self-determination, where the virtuous have a positive degree.

6. Conclusion

I have argued that we can understand Kant’s moral philosophy as structured by four modalities of freedom as moral self-determination. The first modality of moral self-determination is the *capacity* for moral self-determination (humanity); the second is the ideal of the kind of moral self-determination that we ought to realize. This ideal, in turn, splits into two: the ideal of autonomy that is the ideal of complete and necessary moral self-determination, and the ideal of virtue that is the most thorough moral self-determination possible for rational but embodied beings like us. Finally, any agent will have achieved the ideal of virtue to some degree (even if this degree is zero) and so the final modality of freedom of moral self-determination is the *actual degree* of moral self-determination that is the virtuousness and moral freedom of an agent. There are many other concepts of freedom in Kant’s philosophy. I have not, for example, discussed freedom as spontaneity or external freedom. Yet, I believe that these four are the important

⁶⁴ *V-Mo/Collins*, 27:463. Similarly, Kant says that “the distinction between virtue and vice can never be sought in the *degree* to which one follows certain maxims; it must rather be sought in the specific *quality* of the maxims.” (*MdS*, 6:404)

⁶⁵ For another way to reach more or less the same distinctions, see Baxley *Kant’s Theory of Virtue*, 81-2

⁶⁶ Cf. *RGV*, 6:37. Kant would assent to Grotius’s judgment that “There is hardly any Man wicked for nothing, and if there be any one who loves Wickedness for its own sake, he is a Sort of Monster.” (*The Rights of War and Peace* [Liberty Fund, 2005], 1003)

⁶⁷ The change is to use 0 instead of -1 for badness of will. I make this change to simplify the illustration of the argument. Since the argument does not rely on the illustration, but on the distinction between the moral failures of weakness of will and badness of will (quantitative /qualitative), this change should not affect the argument.

modalities of freedom *as moral self-determination* at work in Kant's philosophy. I also believe that once we understand how these are at work in Kant's moral philosophy, we understand that Kant does not, after all, suffer the dilemma that is the problem of moral freedom.

The wicked are free in the sense that they have the capacity for moral self-determination (humanity) and are, therefore, responsible for what they do. The wicked are not, on the other hand, free in the sense that the virtuous are free. For the wicked are not *morally* self-determining: they have a zero degree of virtuousness. But it should hardly be a surprise, and it certainly is no problem for Kant, that the vicious fail to achieve any degree of the kind of freedom that is actual moral self-determination. Taking self-love as one's basic maxim is to subject one's will to a principle of heteronomy, and is, as such, to turn one's back to the freedom *as moral* self-determination that is the ideal of autonomy.

Works cited

Allison, Henry E. *Kant's Theory of Freedom*. Cambridge, England: Cambridge UP, 1990.

---. *Kant's Groundwork for the Metaphysics of Morals: A Commentary*. Oxford, England: Oxford University Press, 2011.

Anselm. "On Free Will." *The Major Works*. Ed. Brian Davies and G. R. Evans. Oxford: Oxford UP, 1998.

Augustine. *The Works of Aurelius Augustine.: A New Translation*. Trans. Marcus Dods, John Richard King, Peter Holmes, Richard Stothert, J. G. Cunningham, Arthur West Haddan, William Findlay, S. D. F. Salmond, John Gibb, James Innes, Joseph Green Pilkington, and Robert Ernest Wallis. Vol. 9. Edinburgh: T. & T. Clark, 1872.

---. *On Free Choice of the Will*. Trans. Thomas Williams. Indianapolis: Hackett Pub., 1993.

Baxley, Anne Margaret. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge, UK: Cambridge UP, 2010.

Bojanowski, Jochen. *Kants Theorie Der Freiheit: Rekonstruktion Und Rehabilitierung*. Berlin: Walter De Gruyter, 2006.

Dean, Richard. *The Value of Humanity in Kant's Moral Theory*. Oxford: Clarendon, 2006.

Denis, Lara. "Kant's Conception of Virtue." *The Cambridge Companion to Kant and Modern Philosophy*. Ed. Paul Guyer. Cambridge: Cambridge UP, 2006. 505-37.

Engstrom. "The Form of Practical Knowledge." *Kant's Ethical Thought*. By Allen W. Wood. Cambridge: Cambridge UP, 1999.

Foreman, David. "Kant on Moral Freedom and Moral Slavery." *Kantian Review* 17.1 (2012): 1-32.

Grotius, Hugo. *The Rights of War and Peace*. Ed. Richard Tuck. Indianapolis, IN: Liberty Fund, 2005.

Guyer, Paul. *Kant*. London: Routledge, 2006.

---. "Kant on the Theory and Practice of Autonomy." *Kant's System of Nature and Freedom: Selected Essays*. Oxford: Clarendon, 2005.

---. *Kant on Freedom, Law, and Happiness*. Cambridge, U.K.: Cambridge UP, 2000.

---. "Nature, Morality, and Peace." *Kant on Freedom, Law, and Happiness*. Cambridge, U.K.: Cambridge UP, 2000.

Hill, Thomas E. "Is a Good Will Overrated?" *Human Welfare and Moral Worth: Kantian Perspectives*. Oxford: Clarendon, 2002.

Hill, Thomas E. *Dignity and Practical Reason in Kant's Moral Theory*. Ithaca: Cornell UP, 1992.

Kant, Immanuel: *Gesammelte Schriften*, Hrsg.: Bd. 1–22 Preussische Akademie der Wissenschaften, Bd. 23 Deutsche Akademie der Wissenschaften zu Berlin, ab Bd. 24 Akademie der Wissenschaften zu Göttingen. Berlin 1900ff.

---. *Religion and Rational Theology*. Allen W. Wood, and George Giovanni. Religion and rational theology. Cambridge New York, NY: Cambridge University Press, 2001.

---. *Practical Philosophy*. Ed. Mary J. Gregor. Cambridge: Cambridge UP, 1996.

- . *Critique of the Power of Judgment*. Ed. Paul Guyer. Cambridge, UK: Cambridge UP, 2000.
- . *Critique of Pure Reason*. Ed. Paul Guyer and Allen W. Wood. Cambridge: Cambridge UP, 1998.
- . *Lectures on Ethics*. Trans. Peter Heath. Ed. J. B. Schneewind. New York: Cambridge UP, 1997.
- . *Lectures on Logic*. Ed. J. Michael. Young. Cambridge: Cambridge UP, 1992.
- . *Anthropology, History, and Education*. Ed. Günter Zöller and Robert B. Loudon. Cambridge, UK: Cambridge UP, 2007.
- . *Notes and Fragments*. Ed. Paul Guyer. Cambridge, UK: Cambridge UP, 2005.
- Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge UP, 1996.
- McCarty, Richard. *Kant's Theory of Action*. Oxford: Oxford UP, 2009.
- O'Neill, Onora. *Constructions of Reason: Explorations of Kant's Practical Philosophy*. Cambridge, England: Cambridge UP, 1989.
- Philo. *Philo*. Vol. IX. Cambridge, Mass.: Harvard Univ. Pr, 1954.
- Plotinus. *Plotinus*. Trans. A. H. Armstrong. Vol. VIII. Cambridge, MA: Harvard UP, 1988.
- Reath, Andrews. "Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination." *Agency and Autonomy in Kant's Moral Theory*. Oxford: Clarendon, 2006.
- Reinhold, Karl L. "Briefe ueber die kantische Philosophie bd.2", *Materialen zu Kants Kritik der praktischen Vernunft*. Ed. Rüdiger Bittner and Konrad Cramer. Frankfurt am Main: Suhrkamp Verlag, 1975.
- Schmid, Carl C. E. "Determinismus und Freiheit", *Materialen zu Kants Kritik der praktischen Vernunft*. Ed. Rüdiger Bittner and Konrad Cramer. Frankfurt am Main: Suhrkamp Verlag, 1975.
- Sidgwick, Henry. "The Kantian Conception of Free Will." *The Methods of Ethic*. 7th ed. London: Macmillan, 1907.
- Thomas, Aquinas. *Summa Theologiae*. Ed. Timothy McDermott. Eyre and Spottiswoode, 1989.
- Timmermann, Jens. *Kants' Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge, UK: Cambridge UP, 2007.
- Wood, Allen W. *Kant's Ethical Thought*. Cambridge: Cambridge UP, 1999.