

# Kant on Cultivating a Good and Stable Will

Adam Cureton

## 1 Introduction

A person's character, according to Immanuel Kant, is her most basic commitment about how she will conduct her life along with the strength of her resolve to live up to the life-governing policy she has chosen for herself.<sup>1</sup> Someone with a good character not only has a *good will*, which is a fundamental commitment to morality above all else, but she is *virtuous* as well, she has a firm and considered resolve to satisfy this commitment despite internal obstacles she may face.<sup>2</sup> Someone with a bad character combines an *evil will*, which is a fundamental commitment to satisfy her non-rational desires, with *stubbornness*, which is a steadfast persistence to pursue her self-interest despite her moral compunctions.<sup>3</sup> People of good will can have a weak character when they, on occasion, deviate from the moral law for the sake of their own happiness or rely on external inducements to act as they should, while those of evil will can have a weak character when they lack the firm resolve to pursue their self-interest when doing so conflicts with morality.<sup>4</sup> On Kant's view, these are the most basic types of character that human beings can have—Kant denies the possibility of a diabolically evil person, one who is committed to evil simply for its own sake.

Kant's moral theory, with its emphasis on rules and principles, is often contrasted with those that highlight the role of virtue and character in a morally good life. As recent commentators have shown, however, virtue and character play an essential and distinctive role in Kant's ethical framework (Cureton and Hill 2015; Denis 2006; Engstrom 2002). Kant discusses many specific virtues, including self-respect, honesty, generosity, and self-improvement, but he defines virtue itself as a kind of fortitude or strength of will to do our duty from duty.<sup>5</sup> Reason prescribes duties of various kinds, including strict ones that can be coercively enforced by others as well as imperfect duties to adopt certain moral ends.<sup>6</sup> Kant's rationalist moral theory excludes the possibility of genuine moral

dilemmas, and it requires that our moral duties always override any other reasons we may have.

A person of good will, on Kant's view, is committed to fulfilling all of her duties from duty while a virtuous person also has the strength of will to do so. Her moral motivations may not be impersonal but instead arise from a sincere commitment to respect the dignity of all. And her actions have positive moral worth when she promotes morally valuable ends, such as the happiness of others, beyond what she is strictly required to do.<sup>7</sup> Kant thinks that a virtuous person is likely to be happy because of the contentment, moral pleasure, and esteem from others that it tends to bring; he claims that we have a duty to cultivate various natural sentiments that assist us in acting as we should; and he emphasizes the importance of moral education.<sup>8</sup> A virtuous person does not feel coerced by morality as an extorting tyrant but instead has a cheerful disposition and tranquil mind that comes from her recognition that she aims to do no wrong and that she is capable of resisting temptations to the contrary.<sup>9</sup> Yet in contrast to ethical theories that are inspired by Aristotle, Kant thinks that a virtuous person may be destitute and unhappy; she may also lack certain sentiments and emotions that she was unable to cultivate; and she may have been subject to an unfortunate upbringing.<sup>10</sup> Kant also claims that most everyone can become more or less virtuous regardless of their early family life or social station. On his view, a person's character is basically up to that person, so feelings and inclinations may play a role in helping us to become virtuous, but a necessary ingredient in virtue is that the person devotes herself to morality above all else and develops in herself the moral strength to follow through on this commitment.<sup>11</sup>

When we explore Kant's conceptions of virtue and character in more detail, we find a number of puzzles that can help us to understand the distinctive contributions that Kant's views should make to our understanding of virtue and character. For practical purposes, according to Kant, we must assume that our will is a spontaneous power of choice that is not determined by any "alien causes," including our desires, feelings, and inclinations.<sup>12</sup> Each of us has the power to choose his own character himself, he "must make or have made himself whatever he is or should become in a moral sense," by committing to a basic, life-governing principle and then deciding on particular occasions whether or not to conform to that fundamental personal standard.<sup>13</sup> It is therefore within each person's own power, including the most evil ones, to acquire a good character.

Yet Kant also claims that we must somehow work to cultivate a good character, in part by repeatedly practicing virtuous acts over time; but no matter what we do we can never reach moral perfection.<sup>14</sup> Given our nature as human beings, "striving after this end always remains only a progress from *one* perfection to another" so our duty is to cultivate a good character through labor and practice "but not to *reach* it (in this life)."<sup>15</sup>

If we must assume we are free in the robust sense Kant describes, however, how could repeated performance of good deeds in the past have any influence on

our spontaneously free wills? And what prevents us from achieving a perfectly good character merely by committing ourselves to morality and, on every occasion, choosing to act accordingly?<sup>16</sup> In addressing these questions, I will suggest an interpretation of Kant's conception of virtue according to which a person's good will is strong when her life-governing commitment to morality is *stable* and her will is weak when her most basic moral policy is *unstable*.

## 2 Character as Good Will plus Virtue

Let's first consider Kant's ideas about a good will and virtue. In the *Groundwork for the Metaphysics of Morals*, Kant characterizes a good will as a basic commitment to do one's duty from duty despite any inclinations or desires to the contrary.<sup>17</sup> A will of this sort is a species of a holy will, which lacks inclinations and necessarily wills in accord with the dictates of reason.<sup>18</sup> The basic principle of a good will, according to Kant, is to abide by the Categorical Imperative in its various forms.<sup>19</sup>

In his late work *Religion within the Boundaries of Mere Reason*, Kant distinguishes two aspects of our will that, as he sees it, we must presuppose about ourselves if we are to genuinely decide what to do and be responsible for our choices. *Wille* is our legislative practical reason, it is reason telling us what we must do and predisposing us to act accordingly, while *Willkür* is our power of choice, it allows us to decide whether to follow the dictates of practical reason, as expressed in the Categorical Imperative, or to act to satisfy our sensible nature. Our inclinations, on Kant's later view, are not evil in themselves, they merely tempt us to immorality, but it is ultimately our own choice whether to indulge or resist them when duty calls.<sup>20</sup>

In *The Metaphysics of Morals*, Kant distinguishes matters of law and justice from those of ethics, virtue, and personal morality. He argues in the second part, called the Doctrine of Virtue, that specific virtues are effective commitments to particular moral ends, such as the happiness of others and our own perfection, while duties of virtue are moral requirements to adopt those ends. Virtue itself, however, is described as the strength of will to fulfill all of our duties from a sense of duty. Thus, a person of good will, who is sincerely committed to morality, can also lack the strength and resolve to put her true and right moral conviction into practice while a virtuous person stands firmly ready to put her good will into effect even when she is inclined to do otherwise.

The character of a person, on Kant's view, does not include her temperament, which is her set of unchosen and innate dispositions of feeling and desire.<sup>21</sup> Being good-natured, malicious, cold-blooded, sanguine, or melancholic as well as having "courage, decisiveness, and perseverance . . . as qualities of temperament" cannot be imputed to us, although they can be aids to morality. Character also differs from habits, which are feelings or behaviors that have acquired "physical

inner necessitation” through frequent repetition.<sup>22</sup> We can be held responsible for choosing to acquire a habit, allowing one to take hold, or giving in to one on a particular occasion, but we are not accountable for our habits themselves because they are merely causal mechanisms that can be freely overcome.<sup>23</sup> Once we separate our character, understood as our basic life-governing commitment plus the strength of our will to put it into action, from our warm temperament, congenial habits, peaceful upbringing, luck avoiding strong temptations to immorality, and any other “gifts of fortune” that we are not responsible for, Kant doubts that we can be very sure about the quality of our own character or that of anyone else’s.<sup>24</sup>

### 3 Final Commitments

Our will, or power of choice, according to Kant, is not a psychological power that can conflict with our desires and inclinations. The maxims or personal policies that we will cannot be overwhelmed by our natural impulses, nor can they be fully explained by any causal descriptions of our behavior, for we must take ourselves always to have the power to act contrary to our sensuous nature even if our actions may not always have the effects that we intend. Yet Kant describes virtue as *strength* of will to do our duty and he describes lack of virtue as *weakness* of will in doing our duty. How should we understand these metaphors if Kant’s view is not, like Hobbes’s, that our will is simply the last and most causally efficacious desire? Weakness of will cannot be an incapacity, or lack of causal strength, on Kant’s view, so what is it? Let’s consider his discussion in the *Religion*:

(1) A maxim, according to Kant, is a personal policy or commitment to act that we freely adopt for ourselves. A maxim typically includes what we plan to do in what circumstances and for what ends. Maxims vary in their generality; some are specific intentions to act while others are life-governing commitments. We settle on some maxims consciously and explicitly while others, according to Kant, must be presupposed as genuine exercises of will in order to explain and rationalize what we do, even if we are not aware of having made those choices explicitly.

(2) Our human wills, on Kant’s view, are motivated by only two kinds of considerations, those of morality and those of self-love.<sup>25</sup> Despite its name, self-love in Kant’s sense is not necessarily selfish or egoistic. Self-love is the nonmoral interests *of* a self, which are not always interests *in* the self.<sup>26</sup> We have interests in pursuing particular chosen ends, plans, projects, and values that are suggested to us by our natural desires and inclinations as well as interests in effectively and efficiently pursuing a coherent conception of our good as a whole. Sometimes our more immediate self-interests conflict with our long-term rational self-interests.

Morality and self-love are the only live options available to us, but they also exert an active influence on our wills.<sup>27</sup> It is part of our human nature that we are disposed to take the fact that an action will further our personal plans, projects,

and values as a reason to do it, and to take the fact that an action would hamper our nonmoral interests as a reason not to do it. As human beings, we are also disposed to take the fact that an action is moral as a reason to do it, and that an action is immoral as a reason not to do it. But unlike Milton's Satan, we cannot regard the fact that an action is immoral as a reason in itself to do it.<sup>28</sup> Morality and self-interest each press their own claims on our wills, but with unequal authority because we cannot help but acknowledge that morality should be given precedence, yet we must ultimately decide for ourselves which motivations to endorse (see Reath 1989).

(3) When an agent makes it her policy to act in a specific way, we can ask: Why did she choose that particular maxim, what was her rationale for doing so?<sup>29</sup> Her reasons or "subjective grounds" for adopting the maxim, according to Kant, were not causally determined by her natural desires, for that would be inconsistent with her freedom.<sup>30</sup> Her freedom can be preserved if she chose her maxim for the sake of a more fundamental commitment of hers, perhaps as a means to satisfying it, or as a way of interpreting and applying that more general personal policy. But then we could ask whether that commitment is explained by an even more fundamental maxim she accepts.<sup>31</sup> There must be an endpoint to such progressions, according to Kant, for otherwise her original maxim would be inexplicable. A final commitment or "first subjective ground" is thus a freely adopted maxim with no underlying explanation for why it was chosen.<sup>32</sup> We may cite what we found appealing about the maxim itself, or mention circumstances that we took as relevant to our decision, but the choice of a final maxim is not explained by any other adopted maxim or by our natural desires and inclinations.<sup>33</sup>

(4) All maxims, according to Kant, are therefore either final maxims or ultimately explained by them. Because morality and self-love are the only sorts of considerations that can influence our wills, every specific maxim we choose must ultimately be explained by a freely chosen final commitment to morality or to self-love. We may act on a particular occasion in order to promote our overall conception of the good or we may act in order to satisfy some specific nonmoral goal or end, knowing full well that doing so is imprudent, but on Kant's view the choice of both kinds of maxims is ultimately explained by a commitment to self-love. Other specific maxims are chosen because of a commitment to morality, and some are chosen on both grounds.

(5) Perhaps then we have two freely chosen final commitments, one to morality and the other to self-love, and all of our other maxims are ultimately explained by one or both of these policies. Yet Kant argues that we cannot have more than one final commitment because that would make our wills internally inconsistent and so exclude the possibility of rational action altogether.<sup>34</sup> We could be presumptively committed to morality and to self-love, without thereby contradicting ourselves, but then Kant thinks we must also be committed to an even more basic priority rule in order to explain why we acted in one way or the other when self-love and morality conflict. An analogy may be

to Kant's argument that a political authority must be absolute because if any power placed limits on the state then that power would be the true authority; similarly, a commitment that ranks self-love and morality would itself be our final commitment rather than the presumptive commitments that it prioritizes.<sup>35</sup> Thus, according to Kant, we can have only one final commitment, which ultimately explains all of our other maxims: "the first subjective ground of the adoption of the maxims, can only be a single one, and it applies to the entire use of freedom universally."<sup>36</sup>

(6) We are disposed to pursue self-love and morality, yet "the two cannot stand on an equal footing," so Kant claims that our most basic maxim is either to subordinate morality to self-love or to subordinate self-love to morality.<sup>37</sup> A good will, according to Kant, is a final commitment to do duty from duty unconditionally, but otherwise to pursue self-interest, while an evil will is any other final commitment that places certain conditions in the name of self-love on satisfying the requirements of morality.

If these basic points are correct, then any maxim we choose must be explained either by a final commitment that makes morality lexically prior to self-love or by a final commitment that does not do so. In contrast to a prominent interpretation suggested by Thomas Hill, this means that we cannot adopt a specific maxim that conflicts with our life-governing commitment because a basic principle cannot explain or rationalize a maxim that conflicts with it.<sup>38</sup> We may know, for example, that some act is wrong, we may even know that the act would not be to our long-term advantage, but if we perform the act anyway to satisfy some nonmoral end of ours, then there must be an explanation for our doing so. We may not have acted for the sake of any other ends, but simply because we found the goal appealing, but on Kant's view choosing that maxim makes sense only because of our commitment to self-love. Yet because the action is immoral, and because we can have only one final commitment, there must have been some allowance in our life-governing maxim that permitted immorality, at least in this kind of case.

Any kind of deviation from the moral law, according to Kant, whether from special-pleading, making excuses, inattention, or distraction, makes us an "evil human being" in his technical sense because our basic commitment allows an immoral exception and any possible explanation of an immoral act is "eradicated" by the "maxim of the good."<sup>39</sup> Someone who fails to do her duty from duty "is conscious of the moral law and yet has incorporated into his maxim the (occasional) deviation from it."<sup>40</sup> Our basic commitment is thus partially revealed by our actions, but for all we know it may include additional escape clauses that have gone unnoticed because we were fortunate enough to avoid the circumstances that trigger them: "It must be possible to infer *a priori* from a number of consciously evil actions, or even from a single one, an underlying evil maxim, and, from this, the presence in the subject of a common ground, itself a maxim, of all

particular morally evil maxims.<sup>41</sup> Having a truly good will is thus a far-off ideal for many of us, who have shown by our deeds that we are not at times genuinely committed to morality in any and all circumstances, as much as we may wish otherwise.

#### 4 Virtue as a Stable Good Will

Our character, according to Kant, is determined both by our basic life-governing commitment and our strength of will in putting that policy into practice. If our final commitment cannot be overwhelmed by our desires and inclinations, how else might Kant explain strength and weakness of will? Kant describes a person with a “frail” will as someone who is committed to doing her duty from duty even though she sometimes chooses, in light of self-interest, not to adopt the subsidiary maxims that her moral commitment requires.<sup>42</sup> A natural way of understanding frailty in this sense, which is suggested by Thomas Hill, is that her will is conflicted, she is committed to morality unconditionally but chooses on a particular occasion not to do what it requires (see Hill 2012a). Yet if the interpretation of the last section is correct, then such conflicts of will are impossible on Kant’s view—her choice to act immorally must be explained by her final commitment, which means her final commitment must have included an exception for this case, so her will at the time could not have been fully good after all.<sup>43</sup> Another possibility is that frailty is an incapacity to put one’s basic commitment into effect because the non-natural inclinations produced by practical reason are not strong enough to overcome our natural inclinations (see Engstrom 1988, 2002). Kant’s transcendental conception of freedom, however, seems to rule out this possibility because, on that view, we are always free to act in spite of our natural desires and inclinations. Frailty may simply be inexplicable, although we must at least be able to explain how it is compatible with Kant’s transcendental account of freedom (see Sussman 2005). Finally, frailty may be a freely adopted commitment to give into temptation on certain occasions (see Allison 1990, 159–160). Yet Kant describes a frail person as having a good rather than evil life-governing commitment, she has incorporated “the good (the law) into the maxim of [her] power of choice,” which means that her will cannot include any moral escape clauses.<sup>44</sup>

The nature of frailty, or weakness of will as Kant understands it, has more to do with the following passage:

The original good is *holiness of maxims* in the compliance to one’s duty, hence merely out of duty, whereby 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*). (R 6:47; italics added)<sup>45</sup>

A further possible explanation of Kant's conception of strength and weakness of will is thus the following:

(1) A final commitment says how we resolve to prioritize claims of morality and self-interest over our entire lives. Assuming it is impossible to adopt a maxim that conflicts with our final commitment, forming an incompatible intention means that our final commitment must have changed to accommodate it.<sup>46</sup> We may commit ourselves to doing something but change our minds when it comes time to form a subsidiary intention needed to put our plan into action: "A person may have the intention [to kill someone in the room], and be entertaining such wickedness in his heart, but when he would proceed to the action, he is horrified by its atrocity, and so changes his intention."<sup>47</sup> We can thus distinguish between the *content* of our final commitment, whether we resolve to be good or evil, and its *stability*, whether we maintain that resolution in various circumstances.<sup>48</sup>

A person may have a good will, and so be committed to morality above all else, but when she recognizes on a particular occasion how difficult it would be for her to put her commitment into practice, she may consciously or unconsciously revise her life-governing commitment to include an exception for the case at hand. She would still be committed to morality for the most part, but her will would no longer be purely good if she decided to act immorally in this instance, for in doing so she would thereby incorporate an exception to her final commitment. Some people may vacillate radically between a good will and a thoroughly evil one while others may retain a general commitment to morality but incorporate and remove minor escape clauses.

(2) Our wills are subject to deliberative tendencies of various kinds that do not causally determine our choices but dispose us to choose in one way or the other. Our tendency to pursue self-love, according to Kant, disposes us to tinker with the moral law to make exceptions in the name of self-interest, to deceive ourselves about the morality of our actions and motives so that we may rest content with doing less than morality requires, to tell ourselves that our natural desires and inclinations prevent us from doing what morality seems to require, to substitute wishing for willing good deeds, and to passively wait for external help to achieve moral goodness rather than pursuing it by our own efforts.<sup>49</sup> Our tendency to morality, on the other hand, includes dispositions to scrutinize our actions and motives, take due care in our moral deliberations, and heed the verdicts of conscience.<sup>50</sup>

(3) When someone is deliberating about her life-governing commitment, these deliberative tendencies often dispose her to choose in conflicting ways, but when she finally makes her decision, her will is in a sort of equilibrium state because she has settled how conflicts among these "forces" are to be resolved. Particular contexts, however, may arouse some of her deliberative tendencies in a way that leads her to rethink and alter her basic commitment. She may not have been fully



aware of what her commitment would involve, she may have even planned for such eventualities, but she may nonetheless dither when faced with the prospect of fulfilling her commitment on a particular occasion (see Hill 1991, 132).

(4) We can say that someone's will is *stable* or *strong* if she tends not to alter her basic commitment too readily and she tends to revert back to it were it to change, while a person's will is *unstable* or *weak* if she tends to alter her basic commitment too readily and tends not to revert back to it were it to change.<sup>51</sup> A person's current resolution may be to do his duty from duty unconditionally, but when he is faced with putting his resolution into action on a particular occasion, the claims of self-interest may gain greater prominence in his deliberations and he may find himself furiously trying to convince himself that he is not morally required to act in this way after all. These self-interested tendencies may ultimately lead him to include an exception in his otherwise good will for the current case.<sup>52</sup> Perhaps later, after performing the immoral act, he feels a twinge of guilt for what he did but decides to dull his conscience as best he can, so instead of removing the illicit exception, he chooses to maintain it. This person seems to have a weak will. His will would have been stronger if he had managed to do the right thing despite the apparent plausibility of pursuing his self-interest or if his conscience and further reflection had led him to eliminate the illicit exception after having chosen to give in to temptation.

## 5 Moral Perfection and Practice

It is our duty, according to Kant, to "be holy" and "be perfect," to have a good will and perfect strength of will, and so "elevate ourselves to this ideal of moral perfection, i.e. to the prototype of moral disposition in its entire purity."<sup>53</sup> Kant also claims that it is impossible for us to reach this ideal, yet he maintains that we nonetheless have a duty to strive for it, so it must be that "compliance with this duty can, accordingly, consist only in continual progress."<sup>54</sup>

Why does Kant think that moral perfection is impossible for us? The duty to "be holy" does not require us to attain a holy will, which is a will that is not subject to incentives or imperatives but necessarily does what reason requires. Developing such a will is impossible for human beings, because we are disposed to pursue our self-interest, even at the expense of morality, which means morality is always a constraint for us.<sup>55</sup> Nor does the duty to "be holy" require that someone have always had a good will, for Kant argues that, in order for us to be responsible for our dispositions to immorality, we must represent ourselves as having chosen an evil will at least initially, which some people then manage to replace with a good one.<sup>56</sup>

Although we can, apparently, never reach moral perfection, we have an imperfect duty to strengthen our wills. Kant does not explain much about how, in particular, we are supposed to cultivate a good character, but he suggests that one of

the ways we can train our wills is simply “by *practicing* virtue (*exercitio*).”<sup>57</sup> Kant’s idea is apparently that through repeated attempts and successes at overcoming obstacle to doing our duty, perhaps even by seeking out contexts in which doing our duty is particularly difficult, we somehow fortify our wills against temptations to immorality as we approach but never achieve full virtue.

Kant’s account of cultivating character seems to conflict with the transcendental conception of freedom that he thinks agents must presuppose about themselves when deciding how to act. First, if someone is free in this sense, then it seems that her will can be good and firm after all, she can choose to commit herself to morality unconditionally and to maintain that commitment on every occasion. Kant himself even represents this ideal in human form as a rational “prototype,” so it is unclear why he thinks moral perfection is impossible for us.<sup>58</sup> Second, our will is not a muscle that can literally be strengthened through repeated use, so it is unclear how time, effort, and practice can increase the likelihood that we freely choose to conform to our basic moral commitment from the motive of duty. Kant even emphasizes that virtue cannot be merely a “*habit* of morally good action acquired by practice” because “this belongs to the natural constitution of the will’s determination.”<sup>59</sup>

We can begin to resolve these tensions by distinguishing what is in principle possible for every human agent from what we can know about ourselves. Our final commitment, according to Kant, is inscrutable; we cannot gain direct knowledge of whether we or anyone else has a will that is good or evil, stable or unstable.<sup>60</sup> We can only make some limited educated guesses about the quality of someone’s will. When we notice a person acting immorally, we can conclude that her fundamental commitment must make an allowance for such cases.<sup>61</sup> And when our conscience punishes us, we can infer that our final commitment includes an immoral exception as well.<sup>62</sup> But we cannot know whether these escape clauses were present all along or whether the person had a good will but simply changed his or her mind at the last minute and incorporated an exception into it.

In order to determine the overall quality of someone’s will with much accuracy, we would have to know how they plan to choose in any conceivable circumstances, not just the ones they have actually faced, and how they would in fact choose were those circumstances to come about. We only encounter a limited range of contexts in our lives, and which ones we do face depends to a significant extent on luck, so at any point we cannot determine whether the next moment would lead us to alter our good or evil will. Because we can never know with much certainty whether we or anyone else has a good and stable will, we can never know whether our duty of moral self-improvement is satisfied. Our best option, in light of this ignorance, is to continue striving to adopt and maintain a good will and hope that we will be successful in doing so.<sup>63</sup> Kant makes a similar point with regard to friendship, claiming that we have a duty to seek perfect friendships, which are in principle possible, although we should regard them as

ideals we should strive for but never fully attain because both parties can never be sure that the conditions of friendship are fully satisfied.<sup>64</sup>

Our freedom to choose is strictly independent of our past deeds, so why does Kant suggest that practicing virtue can help us to cultivate our character? Practice can help us develop various aids to virtue, such as moral judgment, conscience, reason, attention, presence of mind, and feelings of various kinds.<sup>65</sup> But more than that, moral practice, which can involve moral failure as well as success, helps us to get to know ourselves better, which can in turn help us to adopt a good and stable will.<sup>66</sup> First, how we act in a variety of contexts partially reveals our character to us. Immoral acts can alert us to possible exceptions or instability in our basic maxim while difficult situations can reveal possible trouble spots even if we managed to conform to our final commitment. Second, practicing virtue can tell us something about our moral nature as well—that our wills are capable of overcoming our sensuous nature and that practical reason by itself can motivate us to act as we should. The self-knowledge we acquire by practicing virtue, which sometimes involves failing to achieve it, can lead us to redouble our efforts at moral self-perfection and focus our attention on specific aspects of our moral commitment that need improvement. Past examples where we managed to overcome our sensuous nature even though we thought doing so was impossible can combat the deliberative tendency to suppose that our desires and inclinations prevent us from acting as we apparently should. Self-knowledge acquired through moral practice can even strengthen our moral dispositions by highlighting the presence of that disposition itself, which “we cannot cease viewing with the highest wonder” and which dispels “*fanatical* contempt for oneself as a human being.”<sup>67</sup>

The “First Command of All Duties to Oneself,” according to Kant, is to “*know* (scrutinize, fathom) *yourself*,” to “penetrate into the depths (the abyss) of one’s heart,” because the knowledge we gain about the character we have can be used to form an even better one.<sup>68</sup>

## Notes

1. I wish to thank Thomas Hill, Markus Kohl, Robin Dillon, Jordan MacKenzie, Melisa Seymour Fahmy, Iskra Fileva, and participants at a 2013 workshop on Kant’s conception of character for their valuable comments and suggestions. MM 6: 407; CPrR 5: 153; G 4: 393; A 7: 292; R 6: 199; LP 9: 488; V 27: 571.

2. Allison (1990, 140) and Engstrom (1988, 440) cite some passages in which Kant also seems to think of character as just our basic life-governing commitment.

3. LP 9: 488.

4. C 27: 287; V 29: 624; MM 6: 407; A 7: 292–293.

5. MM 6: 621, 435, 429–430, 452, 456, 473, 461.

6. A virtuous person, for Kant, does not “perceive” external moral reasons or wait for her non-rational desires to motivate her, but instead recognizes and accepts principles of rationality that pick out what facts are reasons for what. See Sinhababu, this volume.

7. For discussions of supererogation in Kant, see Baron 1995; Hill 1992; Hill and Cureton 2013.

8. MM 6: 457. CPrR 5: 97–100.

9. MM 6: 485.

10. For discussions of “moral luck,” see Moody Adams 1990; Nagel 1979.

11. G 4: 394, 398–399. Al Mele, this volume, drawing on Aristotle, describes a continent person as someone who acts as she should in all matters but has to overcome temptation in order to do so. On Kant’s view, a fully virtuous person not only has a considered resolve to do her duty despite obstacles but she also does so cheerfully, with a tranquil mind (MM 6: 409). This is an ideal of perfect virtue that Kant thinks human beings must strive for but can never reach, so the best we realistically achieve is a kind of continence in which we stand ready to overcome desires and inclinations that will inevitably tempt us away from doing our duty. For further discussion of Aristotle’s conception of virtue, see Sinhababu, this volume.

12. CPrR 5: 97; G 4: 446.

13. R 6: 43.

14. MM 6: 397, 387; C 27: 465, 366; LT 28: 1077; V 27: 571; R 6: 48, 66–68.

15. MM 6: 446.

16. V 27: 572.

17. G 4: 394–402. A good will in this sense is an individual’s choice, resolve, or will to do one’s duty, not simply the inevitable predisposition to morality that, in Kant’s view, is shared by all moral agents, even the worst. At G 4: 455, however, Kant apparently uses “good will” to refer to the latter.

18. G 4: 414; 439.

19. G 4: 413–421.

20. MM 6: 484–485; R 6: 34–35, but also see G 4: 428, R 6: 21, 34–35.

21. A 7: 285–286, 293.

22. A 7: 149, 286; MM 6: 479, 407, 409; LP 9: 463.

23. V 27: 569; R 6: 23–24, 38; CPrR 5: 98. See Allison 1990.

24. EAT 8: 330.

25. R 6: 23–24; CPrR 5: 72–75. See Korsgaard 1996, ch. 6.

26. This is how Rawls (1999, 111) puts the distinction between these two senses of “self-interest.”

27. R 6: 26–29; V 27: 571.

28. R 6: 35, 37.

29. CPrR 5: 95; R 6: 20–21. See Allison 1990, 136–137; Hill 2000; Korsgaard 1996, ch. 6.

30. R 6: 21.

31. R 6: 21.

32. R 6: 22.

33. R 6: 21.

34. R 6: 24–25.

35. MM 6: 319–320.

36. R 6: 20, 25, 22. See Allison 1990; Korsgaard 1996; Wood 1970, ch. 6.

37. R 6: 36.

38. See Hill 1991, 2012a. Commonsense seems to allow someone to maintain a commitment to, for example, protect the environment, while occasionally acting in ways that

deviate from it by, for instance, using plastic grocery bags when she forgot to bring her reusable ones. Kant's view, as I interpret it, does not allow deviations from a commitment, for in acting contrary to a commitment we thereby show that we either were not committed after all or that our commitment changed to allow for the apparent aberration.

39. R 6: 44, 31.

40. R 6: 32, 36–37.

41. R 6: 20.

42. R 6: 29–31.

43. Although Kant says that frailty and impurity can “coexist with a will (*Wille*) which in the abstract is good” (R 6: 37) he means that even those with evil wills are nonetheless disposed to recognize the authority of the moral law.

44. R 6:30.

45. For further discussions of this problem in Kant, see Baron 1993; Johnson 1996.

46. Although Kant sometimes describes our final commitment as a timeless choice, he regularly allows for a “change of heart,” which is a “revolution” that makes us “a new man,” when we switch from one life-governing maxim to another (R 6: 38, 33, 46). See Allison 1990, 144.

47. C 27: 292. Such a person, on my interpretation, has a will that is not completely good when he decides to kill an innocent person for personal gain because the presence of the immoral intention entails that his most basic commitment included at least one exception to doing his duty. If he abandons that intention for moral reasons before carrying out the act then he has removed the exception from his basic commitment, which means he is closer to having a good will. His life-governing policy may be relatively stable if he rarely forms such nefarious intentions and quickly abandons them when he does or if he rarely abandons his immoral plans and easily returns to them when his conscience and pure practical reason derail him from his immoral ways. Harry Frankfurt (1988) argues that in some cases a person cannot remain herself and perform certain kinds of actions—perhaps the person in Kant's example is so thoroughly committed to his immoral pursuits that he cannot abandon them while still remaining the same person. On my interpretation, Kant would agree that we cannot perform certain actions without also changing our basic life-governing policy, but Kant thinks we are nonetheless the same person who undergoes such changes because of our will or power of choice. Nomy Arpaly (2002) argues that Huck Finn, as she understands him, is virtuous for refusing to turn in Jim even though he believes he should do so. Kant would likely agree that Huck should not betray Jim, but he would question whether Huck has adequately satisfied his second-order duties of self-scrutiny and due deliberation in forming his belief about what he ought to do. Kant's hope is that engaging in such rational reflection about our moral beliefs will lead us to correct judgments about how we should act.

48. Kant distinguishes between a basic commitment to morality and its “stability,” “perseverance,” and “immutability” at R 6: 22, 48, 63, 71.

49. R 6: 27, 32–33, 38, 57, 161; G 4: 424; MM 6: 441, 477; CPrR 5: 73–75, 98; C 27: 293.

50. R 6: 184–187; MM 6: 437–443. See Hill 2012b.

51. V 29: 611. See Rawls 1999, 399–400. For purposes of this chapter, I leave the phrase “too readily” vague and rely on commonsense ideas about those who regularly flip-flop without much temptation or incentive.

52. R 6: 66, 76.

53. MM 6: 446; R 6: 61.  
 54. MM 6: 446; R 6: 48, 74–77. See Allison 1990, 171–179.  
 55. G 4: 413–415; R 6: 61, 64; MM 6: 222, 383, 397, 379, 405.  
 56. R 6: 38, 59, 61, 72.  
 57. MM 6: 398.  
 58. R 6: 60–61, 64; C 27: 591.  
 59. MM 6: 383–384.  
 60. R 6: 21–22, 51, 63, 71; MM 6: 447; G 4: 406.  
 61. R 6: 20–21, 59, 75, 77.  
 62. R 6: 20; MM 6: 392–393.  
 63. CPrR 5: 33.  
 64. MM 6: 469–470.  
 65. MM 6: 328, 401, 457, 495; G 4: 389; V 27: 362, 370; IUH 8: 18–19.  
 66. V 27: 365.  
 67. R 6: 49; MM 6: 441. For related discussions, see Grenberg 2010; Sussman 2005.  
 68. R 6: 441.

## References

- Kant's works will be referenced as follows: C—Collins' Lecture Notes; CJ—Critique of the Power of Judgment; CPrR—Critique of Practical Reason; EAT—The End of All Things; G—Groundwork for the Metaphysics of Morals; IUH—Idea for a Universal History; LP—Lectures on Pedagogy; LT—Lectures on Theology; MM—The Metaphysics of Morals; R—Religion within the Boundaries of Mere Reason Alone; V—Vigilantius' Lecture Notes.
- Allison, Henry E. 1990. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press.
- Arpaly, Nomy. 2002. "Moral Worth." *Journal of Philosophy* 99(5): 223–245.
- Baron, Marcia. 1993. "Freedom, Frailty, and Impurity." *Inquiry* 36(4): 431–441.
- Baron, Marcia. 1995. *Kantian Ethics Almost Without Apology*. Ithaca, NY: Cornell University Press.
- Cureton, Adam, and Thomas E. Hill. 2015. "Kant on Virtue and the Virtues." In *Cultivating Virtue: Perspectives from Philosophy, Theology, and Psychology*, ed. N. Snow, 87–110. Oxford: Oxford University Press.
- Denis, Lara. 2006. "Kant's Conception of Virtue." In *Cambridge Companion to Kant and Modern Philosophy*, ed. P. Guyer, 505–537. Cambridge: Cambridge University Press.
- Engstrom, Stephen. 1988. "Conditioned Autonomy." *Philosophy and Phenomenological Research* 48(3): 435–453.
- Engstrom, Stephen. 2002. "The Inner Freedom of Virtue." In *Kant's Metaphysics of Morals: Interpretative Essays*, ed. M. Timmons, 289–315. Oxford and New York: Oxford University Press.
- Frankfurt, Harry G. 1988. "Rationality and the Unthinkable." In *The Importance of What We Care About: Philosophical Essays*, ed. H. G. Frankfurt, 177–190. Cambridge: Cambridge University Press.
- Grenberg, Jeanine. 2010. "What Is the Enemy of Virtue?" In *Kant's Metaphysics of Morals: A Critical Guide*, ed. L. Denis, 152–169. Cambridge: Cambridge University Press.

- Hill, Thomas E. 1991. "Weakness of Will and Character." In *Autonomy and Self-Respect*, 118–137. Cambridge: Cambridge University Press.
- Hill, Thomas E. 1992. "Kant on Imperfect Duty and Supererogation." In *Dignity and Practical Reason in Kant's Moral Theory*, 147–175. Ithaca, NY: Cornell University Press.
- Hill, Thomas E. 2000. "Basic Respect and Cultural Diversity." In *Respect, Pluralism, and Justice: Kantian Perspectives*, ed. T. E. Hill, 59–86. Oxford: Oxford University Press.
- Hill, Thomas E. 2012a. "Kant on Weakness of Will." In *Virtue, Rules, and Justice: Kantian Aspirations*, 107–128. Oxford: Oxford University Press.
- Hill, Thomas E. 2012b. "Moral Responsibilities of Bystanders." *Virtue, Rules, and Justice: Kantian Aspirations*, 343–357. Oxford: Oxford University Press.
- Hill, Thomas E., and Cureton, Adam. 2013. "Supererogation." *International Encyclopedia of Ethics*.
- Johnson, Robert. 1996. "Kant's Conception of Virtue." *Jarhbuch für Recht und Ethik/ Annual Review of Law and Ethics* 5: 365–387.
- Kant, Immanuel. 1996a. "The End of All Things." In *Religion and Rational Theology*, ed. A. W. Wood and G. Di Giovanni, 217–232. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1996b. "Lectures on Theology." In *Religion and Rational Theology*, ed. A. W. Wood and G. Di Giovanni, 335–452. Cambridge: Cambridge University Press.
- Kant, Immanuel. 1998. *Religion within the Boundaries of Mere Reason*. Trans. Allen W. Wood and George Di Giovanni. Cambridge: Cambridge University Press.
- Kant, Immanuel. 2001. "Moral Philosophy: Collins's Lecture Notes." In *Lectures on Ethics*, ed. P. L. Heath and J. B. Schneewind, 37–222. Cambridge: Cambridge University Press.
- Kant, Immanuel. 2006. "Idea for a Universal History from a Cosmopolitan Perspective." In *Toward Perpetual Peace and Other Writings on Politics, Peace, and History*, ed. P. Kleingeld and D. L. Colclasure, 3–16. New Haven: Yale University Press.
- Kant, Immanuel, and Paul Guyer. 2000. *Critique of the Power of Judgment*. Trans. Eric Matthews. Cambridge: Cambridge University Press.
- Kant, Immanuel, and Thomas E. Hill. 2002. *Groundwork for the Metaphysics of Morals*. Trans. A. Zweig. Oxford: Oxford University Press.
- Kant, Immanuel, and Robert B. Loudon. 2007. "Lectures on Pedagogy." In *Anthropology, History, and Education*, ed. I. Kant, G. Zöllner, and R. B. Loudon, 486–527. Cambridge: Cambridge University Press.
- Korsgaard, Christine M. 1996. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- Moody Adams, Michele. 1991. "On the Old Saw that Character is Destiny." In *Identity, Character, and Morality: Essays in Moral Psychology*, ed. Owen Flanagan and Amelie Rorty, 111–131. Cambridge, MA: MIT Press.
- Nagel, Thomas. 1979. "Moral Luck." In *Mortal Questions*, ed. T. Nagel, 24–38. Cambridge: Cambridge University Press.
- Rawls, John. 1999. *A Theory of Justice*. Rev. ed. Cambridge, MA: Belknap Press of Harvard University Press.
- Reath, Andrews. 1989. "Kant's Theory of Moral Sensibility: Respect for the Moral Law and the Influence of Inclination." *Kant-Studien* 80(1–4): 284–302.
- Sussman, David. 2005. "What's Wrong with Torture?" *Philosophy and Public Affairs* 33(1): 1–33.
- Wood, Allen W. 1970. *Kant's Moral Religion*. Ithaca, NY: Cornell University Press.