

Kant & Moral Character



Carl Hildebrand

University College and Faculty of Philosophy
University of Oxford

A thesis submitted for the degree of
Doctor of Philosophy
Trinity 2017

Kant and Moral Character

Carl Hildebrand, University College and Faculty of Philosophy

A thesis submitted for the degree of Doctor of Philosophy, Trinity 2017

This thesis argues that Kant has a viable, intellectualist account of moral character that is much richer and more interesting than has often been thought. This account is consistent with his broader practical philosophy, in particular, his account of moral worth. Chapter one establishes that Kant has a theory of education on which a child's inclinations are to be trained in preparation for her to grasp the moral law and acquire full moral agency. It argues that his account of habit is complex, recognizing a kind of moral value that is broader than his definition of moral worth. Chapter two argues that sympathy is, for Kant, a primarily cognitive disposition of special importance; this is because it provides knowledge of how the moral law applies in particular circumstances, therefore enabling an agent to fulfill her duties toward others. This chapter also resolves a puzzle concerning Kant's dual concept of character (as both intelligible and empirical) by drawing an analogy with one account of weakness of will. Chapter three develops an account of moral worth that incorporates these more palatable elements of Kant's account of moral character with the seemingly more austere elements familiar from the *Groundwork*. This theory allows for positive, participating inclinations alongside ascriptions of moral worth. Further, it introduces a distinction between full and mitigated moral worth, to account for agents who, for example, act rightly but for confused reasons as in the case of Twain's Huckleberry Finn. Chapter four responds to two objections to Kant on the basis of moral demandingness, one concerning psychological integration and personal relationships, the other concerning the value of non-moral goods more broadly. It then responds to some objections to his account of the highest good, or the idea of a world in which happiness is distributed in proportion to virtue.

Dear Pan, and all ye other gods that dwell in this place, grant that I may become beautiful within, and that such outward things as I have may not war against the spirit within me.

~ Plato

I regard it as the peculiar function of virtue to win over men's minds, and to harness them to its purposes.

~ Cicero

Contents

Introduction	1
1. Education & Habit	11
The Moral Education of Children	14
The Grounding of Character	22
The Moral Value of Habit	27
Kantian Autonomy and the Influence of Inclinations	33
2. Sympathy	37
Kantian Sympathy	40
Moral Anthropology's Contribution to Sympathy	65
The Change of Heart and the Puzzle of Dual Character	71
3. Moral Worth	87
Kant on <i>Charakter</i>	91
Moral Worth: the Locus Classicus	99
The Nature of Maxims	111
A Unified Account of Moral Worth	116
4. The Good	141
Moral Demandingness: Psychological Integration and Personal Relationships	143
Moral Demandingness: Happiness, Saints, and Moral Constraints	152
The Highest Good: Moralism and Practical Import	159
Conclusion	170
Bibliography	175

Abbreviations for Kant's Texts

<i>AN</i>	<i>Anthropology from a Pragmatic Point of View</i>
<i>CJ</i>	<i>Critique of Judgment</i>
<i>CPR</i>	<i>Critique of Pure Reason</i>
<i>CPrR</i>	<i>Critique of Practical Reason</i>
<i>EA</i>	<i>The End of All Things</i>
<i>GR</i>	<i>Groundwork of the Metaphysics of Morals</i>
<i>LP</i>	<i>Lectures on Pedagogy</i>
<i>MM</i>	<i>The Metaphysics of Morals</i>
<i>REL</i>	<i>Religion Within the Boundaries of Mere Reason</i>
<i>SRL</i>	<i>On a Supposed Right to Lie from Philanthropy</i>
<i>TPP</i>	<i>Toward Perpetual Peace</i>

Full bibliographic details may be found at the end of the thesis. All references to Kant's work cite the volume and page numbers of the German Academy of Sciences edition of his collected work, *Kant's gesammelte Schriften*.

Introduction

There is a tradition of venerable thinkers who have misunderstood Kant's ethics. Kant was greatly admired by many of his contemporaries and those who followed after him for what he achieved in the *Critique of Pure Reason*. However, he was often in equal measure criticized by the same people for the general picture of virtue and human agency they saw in his practical philosophy. The most memorable of these criticisms is probably found in what is known as Schiller's joke, where Schiller suggests that Kantian virtue requires one to despise one's friends:

Gladly I serve my friends, but alas I do it with pleasure. Hence I am plagued with doubt that I am not a virtuous person. Sure, your only resource is to try to despise them entirely, and then with aversion to do what your duty enjoins you.¹

Another criticism that has stuck in the minds of many is Hegel's claim that Kant reduces morality to an "empty formalism" which offers no helpful guidance to an agent concerning what she should do in a particular situation.² In Hegel's language, Kant provides "nothing but

* Quotations from inset: Plato, *Phaedrus* trans. R. Hackforth (Cambridge: Cambridge University Press, 1997), 279b (with the qualification that I have substituted "beautiful" for what is "fine" in Hackforth's translation, according to Stephen Scully's translation: Plato, *Phaedrus* trans. Stephen Scully (Newburyport: Focus Publishing, 2003), 279b); Cicero, *On Obligations* trans. P.G. Walsh (Oxford: Oxford University Press, 2008), II.17.

¹ Friedrich Schiller, *Über die Grundlage der Moral*, section 6, trans A.B. Bullock; quoted in H.J. Paton, *The Categorical Imperative* (London: University of Pennsylvania Press, 1971), 48. Also cited in Keith Simmons, "Kant on Moral Worth" in *History of Philosophy Quarterly* 6:1 (1989), 85.

² G.W.F. Hegel, *Outlines of the Philosophy of Right*, trans. T.M. Knox (Oxford: Oxford University Press, 2008), section 135. Alasdair MacIntyre attributes a similar problem to Kant in Alasdair MacIntyre, *A Short History of Ethics 2nd Edition* (London: Routledge, 1998), 197-198.

the establishment of abstract indeterminacy” according to which “no immanent doctrine of duties is possible”.³ Another early criticism, reflected in the language of recent philosophers, comes from Schopenhauer who describes Kantian commitment to the categorical imperative as a “fetish” for the cold of heart.⁴ He echoes Schiller in suggesting that, for Kant, “the value of the character begins only when a man who has no sympathy in his heart and is cold and indifferent to the sufferings of others, and not really born to be a philanthropist, nevertheless carries out good deeds merely for the sake of pitiful, miserable duty.”⁵

Unfortunately, this general outlook has persisted to the present day. Michael Stocker and Susan Wolf criticize Kantian moral theory for being inimical to the good life more generally. Stocker suggests that if we were to take Kantian ethics seriously and lead our lives on the basis of the reasons it provides, we would fall victim to a kind of moral schizophrenia. In using the term moral schizophrenia Stocker means to refer to a split between an agent’s generally healthy, natural desires and her reasons for moral action – again, along similar lines to those suggested by Schiller. To this charge of maladapted moral psychology (something to be avoided in itself) he adds that for similar reasons it would be impossible for the Kantian agent to secure the goods inherent in friendship and other loving relationships with others.⁶ Susan Wolf argues that in its demand for personal moral rigor and the performance of good deeds, Kantian ethics tends to eclipse the notion of non-moral value altogether. To her, the

³ Hegel, *Outlines of the Philosophy of Right*, 197-198.

⁴ Arthur Schopenhauer, *On the Basis of Morality* trans. E.F.J. Payne (Indianapolis: Hackett, 1995), 66. Interestingly, Schopenhauer also anticipates Anscombe’s main point in her paper “Modern Moral Philosophy” when he says: “Kant simply took over from theological morals the *imperative form* of ethics and hence the concepts of obligation, law, and duty, whereas he had to leave behind that which in theological morals alone lends force and significance to those concepts.” (Ibid., 65.) I will not be dealing with this objection in this thesis. For Anscombe’s paper, cf. G.E.M. Anscombe, “Modern Moral Philosophy” in *Philosophy* 33:124 (1958): 1-19.

⁵ Ibid., 65-66, italics redacted.

⁶ Michael Stocker, “The schizophrenia of modern ethical theories” in *Journal of Philosophy* 73 (1976): 453-466. Stocker’s argument is directed toward what he generally refers to as “modern ethical theories”, a term undoubtedly meant to capture Kantian ethics in particular, among other theories.

serious follower of Kantian ethics looks like a kind of moral saint who leads an unhealthy, uninteresting, and generally undesirable life. She argues that Kantian ethics should be rejected on account of it being incompatible with a reasonable, common sense account of the good life.⁷ Bernard Williams offers sustained criticism of Kantian ethics along very similar lines: it might be argued that much of his work aims at dismantling a broadly Kantian moral outlook and providing an alternative to it. He claims Kant has a crude view of emotion and wrongly denigrates the value of affection generally, as well as its role in the motivation of action and the leading of a good human life.⁸

Each of these thinkers contends there is something deeply wrong with Kant's understanding of morality, especially as it pertains to the sorts of things many believe to be valuable in human life. They are suggesting that Kant has a counterintuitive and incorrect, or at least incomplete account of moral psychology. They believe it fails to integrate all those aspects of the human person that it should, and in the right kind of way. In one form or other, these are matters of moral character.

The term 'character' has a wide range of meaning in the English language. The *Oxford English Dictionary* provides fifteen distinct definitions of the term, the meanings of which are associated with a wide range of contexts. These include the abstract and impersonal (e.g., distinctive symbols, tokens, and features of mathematical entities) as well

⁷ Susan Wolf, "Moral Saints" in *Journal of Philosophy* 79 (1982): 419-439.

⁸ See Bernard Williams, *Problems of the Self* (Cambridge: Cambridge University Press, 1973), 226-227. See also, for example, his "one thought too many" objection in Bernard Williams, *Moral Luck* (Cambridge: Cambridge University Press, 1981), 18.

as the political and the artistic (e.g., an individual's rank, a society's reputation, and a person portrayed in a work of fiction).⁹ Among these many definitions, the one most relevant to moral philosophy describes character as the “sum of the moral and mental qualities which distinguish an individual or a people, viewed as a homogeneous whole; a person's or group's individuality deriving from environment, culture, experience, etc.; mental or moral constitution, personality.”¹⁰ Under this description, character is meant to identify those features that mark out what is most distinctive about an individual (or group). That it picks out one's mental or moral constitution suggests that it aims to describe the deep features of one's personality – the sorts of traits that determine who one is and how one can be relied on to act; a person's deep-seated habits, motivations, and thoughts.¹¹

Aristotle's definition of character from the *Nicomachean Ethics* conveys a similar idea. As Terence Irwin points out, the “actions appropriate to a person's character are those said to be ‘proper to him’ or those he ‘is the sort’ to do”.¹² A state of character, or *hexis*, literally means ‘having’ or ‘possessing’ and includes a tendency to do something as a result of habituation or education over time.¹³ Virtuous character, or *ēthos*, is defined by the states (*hexis*) that result from the early habituation of desire in accordance with the correct use of rational deliberation.¹⁴ So, according to Aristotle, good moral character includes having the

⁹ “Character” in *Oxford English Dictionary* ([http://www.oed.com/viewdictionaryentry: showfullentry/true?isAdvanced=false&result=1&rskey=B5ca5n&t:ac=Entry/30639](http://www.oed.com/viewdictionaryentry?showfullentry/true?isAdvanced=false&result=1&rskey=B5ca5n&t:ac=Entry/30639)).

¹⁰ *Ibid.*

¹¹ This overlaps with Julia Annas's definition of virtue as “a *deep* feature of the person”. Julia Annas, *Intelligent Virtue* (Oxford: Oxford University Press, 2011), 9.

¹² See the entry for “character” in the glossary that accompanies Terence Irwin's edition of the *Nicomachean Ethics*. Aristotle, *Nicomachean Ethics* 2nd edition, trans. Terence Irwin (Indianapolis: Hackett, 1999), 319.

¹³ *Ibid.*, 1103b.

¹⁴ In fact, the term *ēthos* is derived from *ethos*, the word for habit (David Bostock, *Aristotle's Ethics* (Oxford: Oxford University Press, 2000), 36).

right kinds of desires and making the right kinds of decisions.¹⁵ Together with the relevant variation of the English term, this more or less captures what might be called the standard or commonsense conception of moral character. It presents the ideal moral agent as one whose desires are in harmony with the good and it presents the good as something that is to be in some sense continuous with human desire. It also describes the ideal moral agent as having the correct thoughts and making the correct decisions concerning morality. It is a picture of personal integration.

It is this picture of integration that the above critics of Kant believe to be incompatible with his moral philosophy. Many of them no doubt take Kant's portrayal of the philanthropist of *GR I* to be his image of the ideal moral agent (Schiller alludes to this philanthropist in the passage above). This philanthropist is plagued by a grief so strong that it renders him emotionally numb and so incapable of acting on the basis of warmer human sentiments. Kant contrasts this state with a previous state in which the same philanthropist found inner satisfaction and delight in meeting the needs of others. He says the philanthropist's action (of helping others) has moral worth only when he is in the later state, which guarantees he is acting on the basis of duty, because he is incapable of acting on the basis of sentiment.¹⁶ Understood in this way, this picture appears to eschew the importance of moral character altogether in favor of a concept of moral worth that requires nothing more of an agent than the exercise of sheer will.

But this is not correct, nor is it the whole picture. Kant's account of moral character and, by implication, moral worth is much richer and more expansive than this picture from the *GR* would suggest. It will be the burden of this thesis to demonstrate this.

¹⁵ For this reason, it is somewhat misleading to translate *hexis* or describe Aristotle's understanding of it as mere habit or disposition. See Aristotle, *Nicomachean Ethics*, 1105a; see also Irwin's entry for "state" (*ibid.*, 349).

¹⁶ 4:398.

In this thesis I provide a systematic reconstruction of Kant's concept of moral character. Moral worth and moral character are deeply connected for Kant, since he frequently employs the concept of moral worth to appraise the value of action and, after all, character without moral worth would not be moral character but something else. So a theory of moral character for Kant will involve a theory of moral worth. Nevertheless, the close connection between these concepts does not imply the sort of truncated account described by the critics above: I argue contrary to these critics that Kant's account of moral character is not as far from the standard, commonsense conception as many have thought. It is not far from this account because Kant has a number of very useful things to say about education, habit, and inclination, for example, and how these things may come together in the life of a virtuous agent.¹⁷ Of course, Kant's conception is not identical to the standard conception. Moral worth is ultimately not a matter of having the right inclinations but of acting on the right principles. Still, his account of moral worth is not inimical to the good life as standardly conceived, once his account is properly understood. I argue instead that these two strands in his thought – of reasonable reflection on moral character on the one hand and a seemingly strenuous standard of moral worth on the other – come together in an intuitively compelling,

¹⁷ I am not the first to take this general position. Several pioneers in this field include Marcia Baron, Richard Henson, Barbara Herman, and Nancy Sherman. For some of their related work, see: Marcia Baron, *Kantian Ethics Almost Without Apology* (Ithaca: Cornell University Press, 1995); Richard Henson, "What Kant Might Have Said: Moral Worth and the Over-determination of Dutiful Action" in *Philosophy Review* 88:1 (1979): 39-54; Barbara Herman, *The Practice of Moral Judgment* (Cambridge, MA: Harvard University Press, 1993); and Nancy Sherman, *Making a Necessity of Virtue* (Cambridge: Cambridge University Press, 1997). I will discuss some of this work in what follows. Further, Anne Margaret Baxley has more recently developed a similar account of Kantian moral worth and character in her book, *Kant's Theory of Virtue: The Value of Autocracy*, which I have read and find myself in considerable agreement with. (Anne Margaret Baxley, *Kant's Theory of Virtue: The Value of Autocracy* (Cambridge: Cambridge University Press, 2010).)

intellectualist account of moral character and moral worth. This account is intuitively compelling because it incorporates the most important aspects of the standard account (e.g., the role of education, habit, and inclination in forming good character, as mentioned). It is an intellectualist account because it emphasizes the primacy of intellectual capacities like belief, cognition, and judgment in the constitution of good moral character. In short, it requires that an agent have sound judgment and correct beliefs concerning both general moral principles and particular actions in order to attain an ascription of full moral worth.

In chapter one, I argue that Kant has a lot to say about character in the standard sense. I examine the *LP* and *AN*, where I argue he has an account of moral education that stresses the importance of bringing up children with virtues of a broadly Aristotelian kind. This account of moral education involves the use and cultivation of inclination and non-rational dispositions to prepare one to acquire an understanding of the moral law. Once one understands the moral law, one can act on it and one's actions can have moral worth. This is what Kant calls the grounding of character. I argue the proper formation of one's inclinations is highly desirable for it. This argument leads me to reconstruct Kant's account of habit and moral aptitude, where I further argue Kant recognizes four types of habit, two of which have indirect moral value and one of which has moral worth in his special sense of the term. I conclude by briefly outlining how inclination might feature in a Kantian account of moral action, given the way Kant understands personal autonomy.

In chapter two, I define Kant's duty of sympathy from the *MM* and its place in his practical philosophy. I argue it is an important feature of his moral psychology, a primarily cognitive disposition the purpose of which is to provide knowledge of what duty requires in a given circumstance. It is a form of practical wisdom and occupies a privileged place as a kind

of master virtue in the Kantian system, analogous to the place of *phronēsis* in the Aristotelian system. I further argue that an important purpose of the *AN* is to provide knowledge of other human beings, so that one can effectively perform one's duty toward them, and that this is a part of sympathetic perception, or Kantian practical wisdom. I examine the puzzle of whether one can simultaneously have a good will and bad moral character (evident in Kant's discussion of the change of heart in the *REL*) and I advance an exegetical and logical solution which draws on this account of sympathy.

In chapter three, I examine Kant's use of the German term *Charakter*, which is roughly analogous in meaning to the English term. This is instructive because it helps to situate the general notion of character in Kant's system and give a sense for how far it might resemble the standard conception. I then examine Kant's discussion of moral worth from the *GR*, where most interpretive attention concerning Kant's account of moral worth has been focused. I argue his definition of moral worth does not imply that duty must be an agent's only motive, but that it must be the grounding motive and must act as a constraint against bad action. This definition from the *GR* does not preclude the involvement of inclination in morally worthy action, contrary to what many have thought. Since, as will become apparent, maxims lie at the heart of moral character and reflective action on Kant's account, I examine what a maxim is and argue it may be understood in a way that is consistent with the notion of character from chapters one and two.

I then reconstruct a unified theory of moral worth that draws together the work of chapters one and two, and the work until that point in chapter three. This theory integrates Kant's account of moral worth with his broader account of moral character and allows that morally worthy action can proceed from character in something akin to but not identical to

the standard sense. In brief, I define moral worth as a category of praiseworthiness attributed to actions that proceed from a morally good disposition and sympathy. In more detail: (a) moral worth can be attributed to an action that proceeds from inclination, so long as the moral law would restrain an agent if that inclination is about to lead her to act contrary to the moral law; (b) moral worth can be attributed to an action if that action were morally required and would have been performed anyway, even if other motives had not been present; and (c) for an agent to receive an unequivocal attribution of moral worth, i.e., full moral worth, she must both have and exercise a well-developed disposition of sympathy; that is, she must possess and exercise good judgment in matters concerning morality and its application. Further, I suggest that moral worth comes in degrees and that we may draw a distinction between full moral worth and mitigated moral worth. This is the final condition in this account of moral worth: (d) an agent may receive an attribution of mitigated moral worth when she exercises the best moral judgment that can reasonably be expected of her, even though it may not meet the standard for full moral worth, as stated in (c). This final condition accommodates a common intuitive response to agents like Mark Twain's Huckleberry Finn, who manage to do the right thing against their better judgment, and it maintains the intellectualist nature of this account of moral worth evident in conditions (a) through (c).¹⁸

In the fourth and final chapter I respond to objections from Michael Stocker, Susan Wolf, and Bernard Williams concerning Kant's (explicit and implicit) account of the good life. These objections include that Kantian ethics is psychologically unhealthy, too demanding, and inconsistent with the enjoyment of human relationships and other forms of non-moral good. I argue these objections rest on an incorrect understanding of Kant and

¹⁸ The moral status of Huck Finn and the challenge he poses to intellectualist accounts of moral worth was (as far as I am aware) first discussed by Jonathan Bennett in his paper, "The conscience of Huckleberry Finn" *Philosophy* 49 (1974): 123-134.

dissolve in light of the account of character provided in this thesis. Finally, I turn to Kant's account of the highest good, the idea of a world in which happiness is distributed in proportion to personal virtue. Here I address two objections, one being that it is rigidly and undesirably moralistic; the second being that it has little or no practical import given its appeal to a divine ruler and an afterlife to bring it about. I argue these objections are not ultimately successful.

Nietzsche once remarked, "Kant wanted to prove, in a way that would dumbfound the whole world, that the whole world was right: that was the secret joke of this soul."¹⁹ Unlike Schiller's joke, this one expresses something true about Kant. Kant is not known for having a facility with written language; perhaps for this reason it is not surprising that the more sensible, standardly intuitive features of his practical philosophy are themselves not widely known. But they are there and it is worth finding them out. Once these features are found, they may be assembled alongside those seemingly more difficult features of Kant's philosophy in a coherent and compelling theory of moral worth and moral character. Or so this thesis will argue.

¹⁹ Friedrich Nietzsche, *The Gay Science* trans. Josefine Nauckhoff (Cambridge: Cambridge University Press, 2001), 40. Also quoted by Joshua Greene in Joshua Greene, "The Secret Joke of Kant's Soul" in *Moral Psychology Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders and Development* ed. Walter Sinnott-Armstrong (Boston: MIT Press, 2007), 35. However, Greene idiosyncratically translates *alle Welt* (what Nauckhoff aptly translates into "whole world") into "common man".

1. Education & Habit

*The soul of the student needs to have been prepared by habits for enjoying and hating finely,
like ground that is to nourish seed.*

~ Aristotle²⁰

Contrary to what has often been thought, Kant has a lot to say about character in the ordinary sense in which the Aristotelian virtues and vices are features of character. His *LP* and *AN* evince that he is attentive to the formation of such character and its moral importance. The *LP* provides an account of how beings that are not fully rational (children) may over time be brought to understand the moral law and enabled to act on it. Kant maintains that this preparation for moral action requires the use and cultivation of non-rational dispositions: in this way, he stresses the importance of bringing up children with virtues of a broadly Aristotelian kind – what Aristotle would call virtues of character. Kant seems to think that bringing up children in something like these virtues of character will prepare them to acquire a sense of duty in due course, much in the way Aristotle thinks they are prepared to acquire practical wisdom (*phronēsis*) and with it, complete virtue.²¹

²⁰ Aristotle, *Nicomachean Ethics 2nd Edition* trans Terence Irwin (Indianapolis: Hackett, 1999), 1179b25-27. From here on I will refer to the *Nicomachean Ethics* in shorthand, as *NE*.

²¹ Regarding Aristotle, I have in mind his idea that full virtue requires *phronēsis* in addition to virtue of character, and that virtue of character requires a good education. The latter point is discussed, for example, in *NE* 2.1, where he emphasizes the importance of acquiring the right sort of habits from youth (Aristotle, *NE*, 1103b22-25). See also *NE* 10.9, where he says, “someone who is to be good must be finely brought up and habituated” (*NE*, 1180a15-16). The former point is discussed in *NE* 6.13 where he says, “each of us seems to

Similar themes may be found in the *AN*. Kant says in the *GR* that a metaphysics of morals “needs anthropology for its *application* to human beings”.²² By this he means that knowledge of empirical human beings – their inclinations, behavior, and psychology – is necessary to enable people to consistently act on the basis of the moral law. He is saying two things here. First, he is saying that people need to understand what in their own nature both aids and hinders moral action. In what follows, I will focus on this point with respect to the development of children, though there are obvious implications for how adults can develop themselves if they have not been fully developed as children.²³ Secondly, he is saying that human beings need to understand the characteristics, habits, and desires of other human beings in order to know how to act morally toward them – a point I will discuss in chapter two. With these objectives in mind, the *AN* details the character of human beings from a number of perspectives. Alongside the *LP*, it helps us to achieve a more complete understanding of Kant’s conception of character and its role in moral formation and action.

I will focus on both of these texts in this chapter. I will first examine Kant’s account of discipline, physical education, shame, and the inclinations to honor, to argue that he

possess his type of character to some extent by nature; for in fact we are just, brave, prone to temperance, or have another feature, immediately from birth. But still we look for some further condition to be full goodness, and we expect to possess these features in another way. For these natural states belong to children and to beasts as well, but without understanding they are evidently harmful” (*NE*, 1144b4-10). He then says of *phronēsis*, “we cannot be fully good without *phronēsis*, or possess *phronēsis* without virtue of character” (*NE*, 1144b32). Of course, there remain important differences between Aristotle’s and Kant’s conceptions of virtue and character. On Kant’s side, this difference most notably involves the relation between inclination and moral worth, since he allows that moral worth can be attributed to someone with inclinations that do not harmonize with the morality, as in the case of the philanthropist of *GR* I, who acts strictly from duty without any inclination to do the good (4:398). Though this is a departure from Aristotle, since Aristotle would say this person lacks complete moral character, I will later argue (in chapter three) that it is not necessarily a departure from common sense or what may be considered intuitive. Some of these differences along with the unique features of Kant’s conception of virtue (and habit) will be clear by the end of this chapter; others will be clarified in chapter three where I discuss his idea of moral worth in detail.

²² 4:412.

²³ When Kant discusses the grounding of character in the *AN*, he says there are few people who have attempted to ground their character (something that happens in a moment of choice, the making of a solemn “vow to oneself”) before the age of thirty; so, the formation of one’s character is clearly something that can continue into adulthood (7:294).

believes these inclinations are important for moral formation and to explain how he believes this to be true. His general idea is to give the child the propensity to do the right thing whenever possible with the belief that when she has acquired the habit of doing the right thing, she will come to see through the exercise of practical reason that the moral law unites all these right actions under a common principle. Once she sees this, she will be able to act on the moral law for the first time. She will also be able to see in many cases what the moral law requires of her, because she was trained to respond to certain cases in a certain way, so her moral perception will have been developed (though, of course, not without need for further reflection and experience). With this trajectory in mind, I will then examine what Kant refers to as the grounding of character in both the *LP* and *AN*, seek to establish what it means for Kant, and argue that the proper formation of one's inclinations is highly desirable for it.²⁴ Finally, in light of this I will reconstruct Kant's account of habit (*Angewohnheit*) and moral aptitude (*Fertigkeit*) in an attempt to clarify his understanding of these features of human nature. This will support my argument that Kant believes the cultivation of something like Aristotelian virtues is crucial to preparing an individual for moral action and in due course will help them to acquire a sense of duty similar to practical wisdom, or *phronēsis*. It will therefore provide evidence that, contrary to what for example Bernard Williams has argued, Kant does not view the emotions as predominantly capricious products of natural causation that are experienced only passively.²⁵ It will also raise the question of whether, given the importance it places on the formation of inclinations, such a notion of moral character can be reconciled to the Kantian account of autonomy, which appears to present moral decisions as a clear choice between equally weighted motivations, with inclinations on

²⁴ For the grounding of character, see both 7:294 and 9:481.

²⁵ Bernard Williams, *Problems of the Self*, 226.

the one hand and the moral law on the other. In concluding this chapter, I will make some observations about this final point and suggest an understanding of Kantian autonomy that can be reconciled to this notion of moral character.

The Moral Education of Children

The organization of the published lectures on pedagogy (the *LP*) makes it difficult to understand the general structure of Kant's theory of education. However, their organization should not be taken to reflect a particular theory or understanding on Kant's part, since they are a product of Theodor Rink's editorial arrangement of the material.²⁶ Kant's central contention is that the educator must train the child to become disciplined, cultivated in skillfulness, prudent, and well-formed in preparation for morality (he enumerates each of these points for emphasis).²⁷ This appears to set a purpose or end for education, a set of skills and dispositions to be acquired in ascending order of value, beginning with discipline and ending with morality. In education, discipline – “a taming of savagery” – must be implemented before skillfulness in “the carrying out of whatever purpose” can be cultivated.²⁸ Then, prudence, or familiarity with human society and culture, must be acquired before it is possible to act on the basis of morality.²⁹ Kant refers to this final stage as moralization and it is central to the highest goal of education.³⁰ The progressive acquisition

²⁶ Rink was a student of Kant and assembled these lectures on Kant's behalf. See: Robert Louden, “Introduction to Kant's *Lectures on Pedagogy*” in *Anthropology, History, and Education*, eds. Robert Louden and Günter Zöller (Cambridge: CUP, 2007), 434.

²⁷ 9:449-50.

²⁸ 9:449.

²⁹ 9:450.

³⁰ *Ibid.*

of these skills and dispositions will provide the most helpful conditions to enable the child to become a free moral agent.

However, this remains something only the child can bring herself to become. It is not something that can be forced upon her for the obvious reason that this would contradict her being free to choose morality and hence, her being a free moral agent. In addition, the moral law is something the child must discover for herself, since it is an a priori principle of (practical) reason and must somehow be grasped by an individual, if that individual is to be said to understand it.³¹ This requires the educator to attain an appropriate balance between cultivating a child's freedom and constraining it, since restraint and redirection of desire is required to act on the basis of the moral law. Kant describes this as the most difficult challenge in education:

How do I cultivate freedom under constraint? I shall accustom my pupil to tolerate a constraint of his freedom, and I shall at the same time lead him to make good use of his freedom. Without this everything is a mere mechanism, and the pupil who is released from education does not know how to use his freedom.³²

Kant accepts this challenge and suggests a number of ways desire and thought might be restrained and redirected in the child, while encouraging the child's freedom.

The most basic way he suggests this be done is through discipline and what he calls physical education. Discipline is the first stage in the four-part process of education outlined above; it "prevents animality from doing damage to humanity" and tames what Kant calls "savagery".³³ "Savagery is independence from laws"; "[t]hrough discipline the human being is submitted to the laws of humanity and is first made to feel their constraint".³⁴ This is

³¹ For the moral law as an a priori principle or "fact" of reason; see: 5:31.

³² 9:453.

³³ 9:449.

³⁴ 9:442.

consistent with what he says in the *REL*, when he describes the vices of animal nature as vices of savagery, which generally entail “wild lawlessness (in relation to other human beings)”.³⁵ At early stages, discipline does not directly involve the cognitive faculties of the child – since these will be limited in power – and is instead an obvious paternalistic intervention on the part of the child’s educators. It is warranted because it prevents this lawlessness from damaging the child’s capacity to grow towards morality and to think and act on the basis of laws generally, a precondition for action that has moral worth. Discipline shapes the child’s habits and general disposition so that she will be prepared to grasp the principle of morality and act upon it when the time comes, because it better enables her to see what moral action looks like.

Physical education is a broader concept than discipline and represents one part of a distinction Kant sometimes makes between two general types of education:

Physical education is the part of education which the human being has in common with animals, or maintenance. *Practical* or *moral* education is the education by which the human being is to be formed so that he can live as a freely acting being.³⁶

Physical education is passive, since it involves obedience to a parent or teacher according to what they deem best for the child, whereas he describes practical education as education towards personality and the establishment of “a freely acting being who can support itself and be a member of society”.³⁷ This distinction is somewhat misleading, however, since both are meant to raise the child up to be an independent agent capable of acting on the basis of the moral law; both also involve the agency of someone else, such as a parent or teacher, who

³⁵ 6:26. Humanity’s vices he calls vices of culture and they involve the drive of human beings to gain ascendancy over one another (6:27). The disposition to personality has no vice, since it refers to the ideal moral disposition itself (6:27-8).

³⁶ 9:455, slight alteration of the translation on my part.

³⁷ 9:475 and 9:455. This progression from physical to practical education parallels the dispositions of the human being that he outlines in the *REL* (animality, humanity, and personality, as above), (6:26).

actively sets up the environment for the child; and throughout both, the child is meant to acquire certain habits and propensities, then reflect on and assimilate them to a moral disposition. So there is not a clear boundary between these two types of education. Although what Kant describes as physical education tends to deal more with bodily habit and less with the cognitive aspects of the human being, both types of education are moral. It is in this way that he understands physical education as an initial, perhaps limited, and yet very important stage of education that prepares the child for knowledge of the moral law and action on the basis of it.

Though physical education in Kant's sense is not limited to activity that involves physical exercise, as we might be tempted to think, it does include it. Kant discusses a number of games that children do and should play, including, for example, gymnastics. He explains that they teach skillfulness and commitment, since for the sake of them the child will "deny himself other needs, and thus learn little by little to do without other things as well".³⁸ These games must have "intention and final purpose" so that the child becomes habituated to control her needs and inclinations for the sake of attaining a goal.³⁹ Morality requires that one be able to control one's inclinations and subordinate them to the moral law; if the child is able to do this when it is required of her in games, she will more likely be able to do this when it is required of her for the sake of morality. Physical education also includes the training of one's mental powers or cognitive faculties.⁴⁰ For example, Kant instructs that memory be cultivated early in a child's education by remembering the names in stories,

³⁸ 9:468.

³⁹ Ibid.

⁴⁰ "Culture consists particularly in the exercise of one's mental powers. Therefore, parents must give their children opportunity for such exercise" (9:466).

reading and writing, and the acquisition of languages other than one's native tongue.⁴¹

Though these examples do not constitute moral education in the more specific sense in which Kant uses the term here, their most important purpose is moral formation.

Kant also discusses emotions associated with an individual's standing amongst their community, focusing in particular on "the inclinations to be honored and loved", which are "aids to morality" and "are to be preserved as far as possible".⁴² In the *AN*, he says, "*love of honor* is the constant companion of virtue", suggesting that this inclination is to remain in the mature moral agent.⁴³ With this in mind, he encourages the educator to make careful use of shame as a tool for guiding these positive inclinations.

He points to several features of the use of shame in education. One is the condition that it only be introduced in the years of adolescence, because "it can only occur when the concept of honor has already taken root."⁴⁴ Since shame is the feeling of one's lowered esteem in the eyes of others, it requires that one is capable of a healthy understanding of the esteem of others in the first place. Kant here assumes that the eyes of these others are worth attending to and that the community of which the adolescent is a part is a good one, with norms that are not immoral. He says elsewhere in the *LP* that a young child has "nothing to be ashamed of and should not feel ashamed".⁴⁵ He believes that to be responsible for doing something wrong requires that there be a reason supporting the judgment that the action is wrong and that the agent in question have a capacity to understand that reason. Children

⁴¹ 9:474.

⁴² 9:482.

⁴³ 7:257.

⁴⁴ 9:484.

⁴⁵ 9:465.

below a certain age would not have this capacity to a degree sufficient for full responsibility, hence his requirement that they not be made to feel ashamed.⁴⁶

A further reason in support of this is that, without the requisite capacity to understand the reason for shame, the child may become embarrassed and hide from others as a result of experiencing the gestures and attitudes that accompany shame. This transitions into Kant's further, cautionary point concerning shame – that it can have a destructive effect on a child's personality, including both her personality in the ordinary sense and in Kant's special, moral sense of the word. If educators cause a pupil to feel shame often, Kant believes the child will not “dare to ask for anything”; rather this child “conceals its disposition and always appears different from what it is, instead of being allowed to say everything frankly.”⁴⁷ The acquisition of moral personality, in one sense, happens gradually: it requires that a pupil grasp and articulate reasons for action until she attains a proficiency in acting independently on the basis of reasons.⁴⁸ Because reasons for action are not externally observable, this learning process requires the child to share her reasons with her teacher honestly and openly;

⁴⁶ Kant says at one point that shame is something given to the person by nature “so that he betrays himself as soon as he lies” (9:478). As is well known, Kant is preoccupied with lying as a fundamental wrong. So we might expect he would say something like this, which, taken explicitly, would restrict the use of shame to cases of lying. Nevertheless, it would be difficult to believe that he thought no serious violation of other rules could be subject to shame. Further, concerning knowledge of the moral law, shame would seem appropriate at certain points before the child grasps the moral law: for example, if she were to violate the rule ‘do not steal from other children’ while understanding that it is a rule, though not that it proceeds from the moral law. Additionally, the love of honor has broader application than lying, and shame is closely associated with honor, so it would seem shame could be more broadly applied as well.

⁴⁷ 9:465.

⁴⁸ There is a tension here. On one hand, the *cognitive facility* that must be in place to support moral personality may be acquired gradually. Yet, on the other hand, there may be a point in time at which moral personality is acquired and before which the agent did not have it (cf. the change of heart in 6:46-7 or the explosion and resolution expressed in 7:294). I will discuss this in further detail in the final section of the second chapter, on the change of heart and the puzzle of dual character.

otherwise progress toward mature moral agency will not be made.⁴⁹ In this way, Kant believes that people require help from one another to reach their capacity as human beings.⁵⁰

In summary, shame only makes sense in connection with the inclinations to be honored and loved by others in one's community; careful and limited use of shame in education provides a way to structure and direct these inclinations. The goal toward which these inclinations are to be directed is moralization, the formation of a disposition to act on the basis of the moral law. This being the case, shame must remain open to the guidance of reason and may only be employed in ways that enable the development of a child's rationality and autonomy.

Pleasure and not its absence, is also a feature of the moral life on the Kantian picture and is to be cultivated in the child. In the *LP*, he tells us that the "upright" person "is one who takes pleasure in willing."⁵¹ He goes on to say that children

must learn to have inner aversion replace the external aversion to human beings or to divine punishment, to have self-estimation and inner worth replace the opinions of human beings, to have the inner value of actions and deeds replace words and emotions, understanding replace feeling, and joy and piety with good humor replace morose, timid, and gloomy devotion.⁵²

The child is to be placed on a trajectory that will eventually enable her to take pleasure in a life structured by and committed to morality. Since an understanding of the moral law is necessary to this type of pleasure, to enjoy this the child must be able to act on the basis of

⁴⁹ I here set aside the problem of the unknowability of maxims and assume the pupil may at least approximate knowledge of her reasons for action to an effective degree.

⁵⁰ "Animals fulfill their vocation automatically and unknowingly. The human being must first seek to reach his, but this cannot happen if he does not even have a concept of his vocation. It is also completely impossible for the individual to reach the vocation [*Bestimmung*]" (9:445). Destiny, purpose, and assignment are more literal translations of the German *die Bestimmung*, translated by Louden as "vocation", which I also take here to imply "capacity".

⁵¹ 9:486.

⁵² 9:493. The sort of pleasure he has in mind here might be better understood by looking to what he says in the *MM*: "[p]leasure that must precede one's observance of the law in order for one to act in conformity with the law is pathological and one's conduct follows the *order of nature*; but pleasure that must be *preceded* by the law in order to be felt is in the *moral order*" (6:378). The idea of moral pleasure here is itself a development of the earlier idea of moral feeling from the *GR* (4:460).

the moral law. This is not something the child can be made to do because it is something she can only do for herself; nevertheless, it will be beneficial if she is trained up to have emotions that incline her to take pleasure in action that conforms to morality. At an early stage, these sorts of desires and emotions will not yet have moral content, but they will be more likely than others to be replaced by desires and emotions that do.

Finally, pleasure and friendship are integral features of Kant's conception of the good life, contrary to the tradition of interpretation of Kant familiar in Schiller's joke mentioned in the introduction to this thesis. There seem to be two points here: one is that Kant places an emphasis on enjoying oneself when that is compatible with duty and the other is that there is pleasure to be gained from doing one's duty.⁵³ With regard to the first, he advises that the child build friendships and prepare herself "for the sweetest enjoyment of life" (I will develop this idea further, saying more about pleasure and friendship in chapter three).⁵⁴ With the second point in mind, he says that children "must be openhearted too, and as cheerful as the sun in their expressions. The cheerful heart alone is capable of rejoicing in the good."⁵⁵ He adds that a "religion which makes the human heart gloomy is false" for the child must do her duty "with a cheerful heart and not out of constraint."⁵⁶ Cheerfulness – that the child takes pleasure in willing the good – appears to be a sign that she is en route to acquiring moral personality.⁵⁷ Kant also discusses cheerfulness in the *MM*, stating "what is not done

⁵³ The first point is also supported by the fact that Kant allows for the enjoyment of things indifferent to morality (e.g., enjoying either wine or beer with dinner, according to one's taste); he says that to act as though all such actions were relevant to morality would be "fantastically virtuous" and "turn the government of virtue into tyranny" (6:409).

⁵⁴ 9:484-5.

⁵⁵ 9:485.

⁵⁶ *Ibid.*

⁵⁷ Loudon says the sense of cheerfulness Kant is invoking here is "a strong one, and it is clearly more than emotional window-dressing". He adds that cheerfulness has a transcendental dimension because it results from the awareness of one's freedom having been restored from a state of dependence on natural impulse. (Robert Loudon, *Kant's Impure Ethics* (Oxford: Oxford University Press, 2000), 52.)

with pleasure but merely as compulsory service has no inner worth for one who attends to his duty in this way”.⁵⁸

The Grounding of Character

So far I have argued that Kant’s account of the moral education of children takes seriously the formation of non-rational dispositions: he seems to think that bringing up children with virtues of a broadly Aristotelian kind will best prepare them to understand and act upon the moral law. This is a kind of character formation in the ordinary sense. Yet Kant explicitly describes the ‘grounding of character’ as something different from this. In what follows, I will explain what he means by the grounding of character and how it fits with the account of character formation described above.

Kant refers to the grounding of character in both the *LP* and *AN*; the idea has a slightly different meaning in each case.⁵⁹ In the *LP*, he says:

The first effort in moral education is the grounding of character. Character [*Charakter*] consists in the aptitude [*Fertigkeit*] of acting according to maxims. In the beginning these are school maxims and later maxims of humanity. In the beginning the child obeys laws. Maxims too are laws, but subjective ones; they originate from the human being’s own understanding.⁶⁰

⁵⁸ 6:484. Louden quotes the same passage in his discussion. Of course, this looks to be in tension with Kant’s discussion of moral worth and the example of the philanthropist in the *GR*. I will look at this more closely in chapter three, where I will argue this tension follows from a misunderstanding of Kant’s views.

⁵⁹ He also covers it in Part One of the *REL* (see 6:44-52 in particular), where it is described in terms of a moral commitment, or “change of heart”.

⁶⁰ 9:481.

Here he is claiming that character consists in the ability to act on the basis of maxims, and it appears that these maxims do not have to be moral. Acting from maxims could mean acting on a kind of policy that is meant to apply in certain similar cases, rather than simply reacting to circumstance in a way that one does not think of as applying to other cases. What is important at this stage is that the child secures this ability and develops it into a kind of cognitive skill. This cognitive skill, or aptitude, is a cultivated proficiency in thinking and acting on the basis of rational maxims. Kant goes on to say, “if one wishes to form a character in children, it is very important to draw their attention to a certain plan in all things”.⁶¹ Elsewhere, he says that rules “must be found in everything that is to cultivate the understanding”.⁶² The acquisition of this aptitude requires practice and therefore has to be gradual.⁶³ The point of having these rules and drawing the child’s attention to them is to provide an environment ideally suited to the development of her rational faculties and facility in acting on the basis of maxims.

But how does an awareness of rules encourage this development? The rules Kant has in mind are those that admit of some form of rational justification, even if that rationality is at times only instrumental and comes in the form of a hypothetical imperative, since he regards hypothetical imperatives as belonging to practical reason.⁶⁴ This is likely because hypothetical imperatives involve some assessment from the point of view of an agent about how best to achieve a given end. A given practice that is based on rules, along with the ends it will entail, will initially be set by the parent or educator; however, the justification for this

⁶¹ Ibid.

⁶² 9:474-5.

⁶³ In one sense, though admittedly not in another, cf. note 48 above.

⁶⁴ Before introducing hypothetical and categorical imperatives as the two forms a practical rule may take, he says a “practical rule is always a product of reason because it prescribes action as a means to an effect, which is its purpose” (5:20).

practice will be within the rational grasp of the child once she reaches a requisite stage of cognitive maturity. An example of such a rule-based practice might include something as simple as speaking to others in a certain manner and tone of voice, because it demonstrates respect. This introduces the child to a certain end (showing respect for others); once they grasp this end, it will be possible for them to reflect on the various circumstances it might be relevant in, as well as other ways it might be achieved. This happens much in the way that a child may learn mathematics by rote, before grasping the basic mathematical principles involved and going on to develop new proofs. As mentioned above, other features of education (for example, games and what Kant calls physical education) provide the opportunity for the child to develop the sort of strength and resolution required to overcome immediate inclinations in order to achieve certain ends. These are skills necessary for and transferable to morality.

This is consistent with Kant's description of character in the *AN* as a kind of resoluteness and ability to consistently act on the basis of principles, whether good or otherwise.

But simply to have a character signifies that property of the will by which the subject binds himself to definite practical principles that he has prescribed to himself irrevocably by his own reason. Although these principles may sometimes indeed be false and incorrect, nevertheless the formal element of the will in general, to act accordingly to firm principles (not to fly off hither and yon, like a swarm of gnats) has something precious and admirable in it.⁶⁵

Here, Kant is indeed praising the ability to simply stick to one's principles, whether they are good or bad. He goes on to commend Sulla's "firm maxims" and "strength of soul", despite the fact that he "arouses disgust" through the violence of his maxims.⁶⁶ This praise is qualified, however. He attributes to Sulla "*strength of soul*" not "*goodness of soul*", and

⁶⁵ 7:292.

⁶⁶ 7:293.

claims that both must be found together to constitute “*greatness of soul*”.⁶⁷ He goes on to say that “character requires maxims that proceed from reason and morally-practical principles.”⁶⁸ He is not entirely consistent in his usage of the term ‘character’ (in German, *Charakter*: I will say more about his understanding of this term in chapter three). Nevertheless, it is not difficult to make sense of what he means against the backdrop of themes he consistently emphasizes throughout his practical philosophy. As in the *LP*, the goal of education is the formation of a disposition toward morality, or: rational, free, and moral agency. Resoluteness is important and praiseworthy because it is a necessary condition for this, since morality requires an agent to understand and articulate maxims, and then restrain and redirect her desires in order to act on them.

Kant seems to have something like this in mind when he describes the grounding of character the second time, in the *AN*.

One may also assume that the grounding of character is like a kind of rebirth, a certain solemnity of making a vow to oneself [...]. Education, examples, and teaching generally cannot bring about this firmness and persistence in principles *gradually*, but only, as it were, by an explosion [...]; the grounding of character, however, is absolute unity of inner principle of conduct as such.⁶⁹

In the context of this paragraph, the grounding of character entails a decisive commitment to the moral law and self-perfection. Unlike the account in the *LP*, it has a distinctly moral flavor. In this way, it closely resembles the phenomenon Kant describes in the *REL* as the change of heart, and I suggest it amounts to the same thing. (I will examine the change of heart in further detail in chapter two.) This change of heart, or acquiring of a fundamental predisposition to the good, involves an ability to know and accurately judge the moral worth

⁶⁷ Ibid.

⁶⁸ Ibid.

⁶⁹ 7:294.

of maxims. Kant says in the *REL*, that “[t]his predisposition to the good is cultivated in no better way than by just adducing the example of good people [...] and by allowing our apprentices in morality to judge the impurity of certain maxims on the basis of the incentives actually behind their actions. And so the predisposition gradually becomes an attitude of mind”.⁷⁰

This is how Kant approaches the education of children in the moral catechism of the Doctrine of the Method of Ethics, toward the end of the *MM*. Here, the “teacher elicits from the pupil’s reason, by questioning, what he wants to teach him”.⁷¹ By asking a series of careful questions and suggesting possible answers, the teacher encourages the pupil to see what her duty is and that dutiful action is what makes one worthy of happiness. Similar to the way in which encouraging people to think clearly may help them to see how conclusions follows from premises, the teacher makes the child aware of and attentive to an innate “rule and instruction” that lies in “*reason alone*”.⁷² The child is meant to see that practical reason demands purity of maxims and that the moral law is the principle on which all pure (moral) maxims are based.⁷³ This complements the way that good habituation is meant to prepare the child for the change of heart, as argued above, and it encourages the child to stick to the correct, moral principles. She is encouraged to develop self-control and take pleasure in willing the right sort of things, making it more probable that she will in time undergo the

⁷⁰ 6:48. Kant says this in the *REL*, immediately after saying that “a human being’s moral education must begin, not with an improvement of mores, but with the transformation of his attitude of mind and the establishment of a character [...]” (ibid.). This might appear to be at odds with the idea that this disposition can be cultivated; however, he concludes this sentence by saying: “[...] although it is customary to proceed otherwise and to fight vices individually, while leaving their universal root undisturbed” (ibid.). The point here seems to be that education should be directed toward establishing this character, and this is something a teacher can prepare her pupil for by helping the pupil understand what a pure maxim looks like and in doing so, what kind of maxim constitutes good character.

⁷¹ 6:480.

⁷² 6:481.

⁷³ See: “Fragment of a moral catechism” (6:480-484).

change of heart and choose to act on the moral law. Similarly, the method of the moral catechism brings her to understand the demands of practical reason and to separate maxims of morality from maxims of self-love. If she understands the nature of morality, it will be more probable that she will in time undergo the change of heart and choose to act on the moral law, since moral understanding is a necessary condition for moral action.

Of course, as Kant says, no amount of education can ultimately bring about this change; it is a commitment made on the part of the agent herself. The purpose of education is to place an individual in an ideal position to make this commitment by encouraging her to think, control her feelings, understand maxims, and act consistently on the basis of maxims. This is to be done through discipline and physical education, the careful use of shame and the inclinations to honor, and the cultivation of pleasure in willing, among other things. These things facilitate the grounding of character in the second, moral sense: the better an individual can think, control her feelings, understand maxims, and act on the basis of them, the more likely she will be to first, grasp the moral law, and then make the decision to act on it.

The Moral Value of Habit

It is helpful to reflect on what Kant says about habit and moral aptitude in light of this account of moral education and the grounding of character. He typically uses one of three terms to refer to the idea of habit at various points in his practical philosophy: *Angewohnheit*, *Fertigkeit*, and *Denkungsart*. In what follows, I focus on the first two terms and offer a

reconstruction of his account of habit and moral aptitude in an attempt to clarify his picture of these features of moral psychology.⁷⁴

In the *MM*, he contrasts habit with moral aptitude, stipulating that only the latter may count as virtuous.

An *aptitude* [*Fertigkeit*] (*habitus*) is a facility in acting and a subjective perfection of choice. – But not every such *facility* is a *free* aptitude (*habitus libertatis*); for if it is a *habit* [*Angewohnheit*] (*assuetudo*), that is, a uniformity in action that has become a *necessity* through frequent repetition, it is not one that proceeds from freedom, and therefore not a moral aptitude [*moralische Fertigkeit*]. [...] Only such an aptitude can be counted as virtue.⁷⁵

The reason why he says here that only aptitude, and not habit, can count as virtuous is that only aptitude involves choice on the part of the agent. Here habit is identified as an unthinking pattern of action resulting from frequent repetition of that action. He believes that habits of this kind typically restrict one's freedom: in the *AN* he says, "habit [*Angewohnheit*] deprives even good actions of their moral worth because it impairs the freedom of the mind".⁷⁶ In the *LP*, he even says that children should not be allowed to develop any habits.

The more habits [*Angewohnheiten*] one has, the less he is free and independent. It is the same with the human being as with all the other animals: they always retain a certain propensity for that to which they were accustomed early. The child must therefore be prevented from getting accustomed to anything; he must not be allowed to develop any habits.⁷⁷

This would seem to suggest that Kant is opposed to the development of even virtuous habits understood in the broadly Aristotelian sense. However, he is likely thinking that the more

⁷⁴ For the purposes of this chapter, I consider *Denkungsart* as a kind of aptitude and do not go into further detail, since I will be discussing this in chapter three. In short, the two are very similar. *Denkungsart* is typically translated as 'way of thinking', 'manner of thinking', 'habit in the way of thinking', or something similar. It is by itself morally neutral and can have good and bad manifestations. The most significant difference, perhaps, between *Fertigkeit* and *Denkungsart* is that the former denotes a kind of skill or facility and hence itself already implies a term of praise, whereas *Denkungsart* tends to be neutral in this respect.

⁷⁵ 6:407.

⁷⁶ 7:149.

⁷⁷ 9:463.

propensities one has that lead one to behave in ways one has not consciously considered, the less one is in control of one's behavior. It is also important to consider that what reads as 'habit' in English is a translation of the German '*Angewohnheit*', which connotes a sense of artificiality, as in a mannerism or unthinking routine. Neither the English nor the German term typically express one's inner character as a free and conscious moral agent, and these features are central to Kant's understanding of moral agency. I call this *unthinking habit*: it is Kant's first conception of habit and it is not conducive to the formation of moral character.

However, this does not mean that all habits or *Angewohnheiten* are bad. As I have argued above, discipline and physical education of the child are about a kind of habit formation and are highly desirable according to Kant. Furthermore, the parent is to encourage the healthy development of certain dispositions in the child, such as the inclination to honor and a tendency to take pleasure in willing the right things. These will be a product of the choice of an external agent, the child's parent or teacher, and as such do not entail a conscious choice on the part of the child to shape herself in a certain way. So we could not praise a child who does what is good merely as a result of this education. Nevertheless, the purpose of this education is to prepare the child to think and act independently and eventually morally, by preventing inclinations damaging to this from cropping up, as well as by encouraging the development of her mental faculties and self-control. Since self-control and a certain mental acuity are necessary for moral agency, insofar as certain habits encourage these things, they are of positive moral value for Kant. I call this *external moral habit*: it is Kant's second conception of habit and it is conducive to the formation of moral character. However, it is not sufficient for attributions of moral worth, since the child does not yet understand the moral law or choose it as the ground of her actions.

Kant has a third conception of habit (here using the term ‘habit’ broadly), which I call *mere aptitude*, since it involves a cultivated facility in acting on the basis of maxims given to an agent by herself, though those maxims are not yet moral. ‘Aptitude’ is an untroubled translation of the German ‘*Fertigkeit*’ and implies a skill or proficiency, a typically positive capability often consciously acquired. In addition to the above passage from the *MM*, Kant uses it in the earlier cited passage on the grounding of character from the *LP*, where he says the first effort in moral education is the grounding of character: “[c]haracter consists in the aptitude [*Fertigkeit*] of acting according to maxims. In the beginning these are school maxims and later maxims of humanity”.⁷⁸ Here aptitude is necessary for the acquisition of character even in the basic, not yet moral sense. Character cannot be attained through a process of unthinking habituation: it lies in the ability to think and act according to maxims. Mere aptitude is a proficiency in thinking and acting on the basis of maxims in general, namely hypothetical imperatives and not necessarily the categorical imperative, or moral law. So, while mere aptitude is a necessary condition for moral aptitude, it does not have independent moral value.

Moral aptitude (*moralische Fertigkeit*) is Kant’s fourth conception of habit and is necessary for moral character in the full sense. Kant defines moral aptitude as something that can be counted as virtuous because it proceeds from freedom.⁷⁹ Here, he has in mind how moral maxims are grounded in the moral law and yet subjective because they are products of an agent’s autonomy in the sense described in the *GR*.⁸⁰ Moral aptitude is a cultivated proficiency in thinking and acting according to rational and moral maxims. It is different

⁷⁸ 9:481. The passage from the *MM* that I am referring to is found in 6:407.

⁷⁹ 6:407.

⁸⁰ “Autonomy of the will is the property of the will by which it is a law to itself (independently of any property of the objects of volition). This principle of autonomy is, therefore: to choose only in such a way that the maxims of your choice are included as universal law in the same volition” (4:440).

from both unthinking habit and external moral habit because it includes an active cognitive component on the part of the agent. It is different from mere aptitude because it is grounded in moral maxims. So actions proceeding from moral aptitude qualify for ascriptions of moral worth.

These four conceptions of habit trace a path of steady increase in thoughtfulness, from a state of unthinking, to principles given to one by another (for example, one's parent), to principles given to one by oneself, and finally to principles given by the moral law. It seems reasonable to assume that when a parent encourages the development of external moral habit, the intention would be that these dispositions remain in the child through to and likely throughout her adult years. Kant may have this in mind when he says in the *AN* that “*love of honor* is the constant companion of virtue”.⁸¹ So this habit formation should have a lasting impact that is conducive to moral action.

Though this habit formation should have a lasting impact conducive to moral action, Kant does not spell out precisely how this might work, after the individual has attained the ability to think and act on the basis of moral maxims. One way it might work, which is consistent with his idea of moral worth, is that it would make it less difficult to act on the basis of the moral law and therefore more probable that an agent would consistently do so. Kant's definition of moral worth requires that an agent make a decision to act on the basis of the moral law. Further, as I will specify later, for an action to have moral worth the moral law must be a sufficient condition for the agent to perform that action. So the agent would have to be willing and capable of acting on the basis of the moral law under even difficult circumstances. This would rule out actions that proceed from inclination from ascriptions of moral worth, including even habits that incline one to act in a way that conforms to morality,

⁸¹ 7:257.

unless those actions are ultimately grounded in a moral maxim. However, if an agent could know in advance that cultivating certain inclinations and working to root out or diminish the force of others could make it more probable that she would choose to act on the moral law in future circumstances, it would be a good thing to do this. Even if this meant she were acting on inclination in these future circumstances and so not in a way that merits moral worth, it would, all things considered, be better if she were inclined to act in a way that is consistent with morality than in a way that is not (though, of course, thoughtful moral action remains the ideal). In this sense, there are two types of agent: here we would be dealing with an agent who may have undergone the change of heart, but not permanently, in contrast to an agent who has undergone a permanent change of heart. Since the latter (perhaps hypothetical) agent would always act on the correct, moral principle, she would not need to worry about taking these kinds of precautions. (Of course, it would seem unwise to assume that one could never be subject to such a lapse, or reversion of character.) This seems to fall under the duty of one's own perfection when Kant says in the *MM* that “[n]atural perfection is the *cultivation* of any *capacities* whatever for furthering ends set forth by reason”.⁸² He says also to “cultivate your powers of mind and body so that they are fit to realize any ends you might encounter”.⁸³ So it looks as though the habituation of one's early life retains positive moral value in the life of the adult, even after an individual has grasped the moral law and become capable of acting on the basis of it. This is also in spite of the fact that action proceeding from inclination does not have moral worth.

⁸² 6:392.

⁸³ *Ibid.*

Kantian Autonomy and the Influence of Inclinations

I began this chapter by setting out to understand Kant's account of the moral education of children and the grounding of character, primarily as discussed in the *LP* and *AN*. I argued that in this account Kant places great importance on the formation of character and the development of a variety of non-cognitive habits in the pupil. The predominant purpose of developing these habits is to encourage the child to think and control her feelings, so that she may be best prepared to understand and choose to act on the moral law. The commitment to a life of self-perfection based on the moral law is understood by Kant as the change of heart or grounding of character; it requires an adeptness in understanding and acting on maxims, and in particular, self-control and resolve to carry through on one's commitment. This self-perfection is partially constituted and supported by habit. In light of this, I revisited Kant's comments on habit (*Angewohnheit*) and aptitude (*Fertigkeit*), offering a reconstruction of these concepts that I believe to be more consistent with dominant features of his work, particularly with what he says about moral education and the grounding of character. It turns out that Kant has a lot to say about moral character in the ordinary sense and places significant moral value on the role of habits and other non-rational features of human nature in his moral psychology.

This should be considered alongside the idea one at times gets from Kant that in the context of moral decision making, the agent is always or at least regularly making a choice between inclination and the moral law as though they were equally weighted motives. For example, in the *REL* he says: "whatever his previous behavior may have been, whatever the natural causes influencing him, whether they are inside or outside them, his action is yet free

and not determined through any of these causes; hence the action can and must always be judged as an *original* exercise of his power of choice.”⁸⁴ I would like to suggest that this idea does not accurately capture what Kant has in mind in passages like this.

Given the evidence I have presented above, it seems clear that the training of one’s inclinations has positive moral value for Kant. In light of this, I suggest that he understands the moral law as available to the agent as a possible motive constant throughout all circumstances, whereas the strength of various inclinations as possible motives will vary depending on the experience of the agent. On Kant’s account, the agent will be free to act on either the moral law or inclination in any circumstance and is held responsible accordingly. However, the strength of these possible motives relative to one another will differ. The agent may exercise further self-mastery (autonomy) here in advance by modifying the strength of various inclinations that lead away from or toward action that conforms to the moral law. This means that the project of character formation, which entails the cultivation of one’s inclinations to conform with moral ends, is possible on the Kantian picture. Though it may not directly increase the strength of the moral law as a possible motive, it can adjust the strength of contrary inclinations. And as I have argued, this can make moral decisions and hence the moral life less difficult (and more pleasant) for the agent. It can also make it more likely that an agent’s future action will conform to morality, which is still a (comparatively) good thing even when it means that action does not come with an ascription of moral worth. Some inclinations or feelings may also serve an epistemic function and enable the moral law to be manifest to an agent more clearly in the context of decision, which I will say more

⁸⁴ 6:41.

about shortly. These are some of the reasons why Kant sees the cultivation of one's faculties as a duty.⁸⁵

⁸⁵ 6:391-392.

2. Sympathy

Our passions do not live apart in locked chambers, but, dressed in their small wardrobe of notions, bring their provisions to a common table and mess together, feeding out of the common store according to their appetite.

~ George Eliot⁸⁶

Many men would have arrived at wisdom if they had not fancied that they had already arrived, if they had not dissembled about certain traits in their character and passed by others with their eyes shut.

~ Seneca⁸⁷

In the previous chapter, I provided a reconstruction of Kant's account of habit and moral aptitude, arguing that four different conceptions of habit may be found in his work. Each of these four conceptions – unthinking habit, external habit, mere aptitude, and moral aptitude – indicate an increase in the level of personal thought involved in performing an action from that habit. All four can be conducive to moral action in the way described in chapter one, though only the fourth type – moral aptitude – qualifies for ascriptions of moral worth. As has been said, moral aptitude is a cultivated proficiency in thinking and acting according to rational and moral maxims: it includes an active cognitive component on the part of the agent and is grounded in the moral law. So far, this describes a type of virtue on the Kantian conception rather than a particular virtue in the way that, for example, Aristotle

⁸⁶ George Eliot, *Middlemarch* (Oxford: Oxford University Press, 2008), 156.

⁸⁷ Seneca, "On Tranquility of Mind" in *Seneca: Moral Essays Volume II* trans. John Basore (Cambridge, MA: Harvard University Press, 2006), 211.

describes the virtues of courage, liberality, pride, or truthfulness. Though one can find several lists of particular virtues in the *AN*, *LP*, and *MM*, for example, these lists seem rushed and provide little or no justification for the particular virtues they include. This does not mean they are meaningless or provide no insight into Kant's ideal of moral character; however, it does make it difficult to believe they are central to his conception of moral character. In this chapter, I will argue that there is a particular virtue central to Kant's conception of moral character: the virtue of sympathy.

One might ask whether and why there is only one such virtue. I will argue that sympathy occupies a privileged place among Kantian virtues because it is a kind of master virtue, arguably similar to Aristotelian *phronēsis*, which involves cognitive and may involve non-cognitive aspects of the agent's character, and functions to attune the agent to the deepest value in Kantian ethics – the dignity and well-being of humanity. In making this argument, I will examine Kant's extended discussion of sympathy in the *MM* as well as what he says about sympathy elsewhere. I will begin by providing some background to the duty of sympathy, explaining its more general place and function in Kant's ethics. Following that, I will outline the nature of the sympathetic disposition, explaining its most important, cognitive aspect as well as how it might be integrated with emotion. I will then explain how the moral anthropology of the *AN* contributes to the cognitive aspect of sympathy by providing knowledge of other human beings, the purpose of which is to enable the agent to more effectively perform one's duty toward them.

This account of sympathy connects with a broader question concerning the nature of moral character for Kant. In the *REL*, among other places, he asserts that an agent can have a change of heart such that she immediately acquires good character in the noumenal sense,

whereas good character in the empirical sense must be acquired gradually. This is puzzling. How can an agent be motivated by the moral law and yet still be inclined to do what is wrong, as this seems to suggest? Can a good will exist alongside inappropriate motivations or bad moral character? A positive answer to the latter question would seem counterintuitive, for we do not typically allow, for example, a motivation to cheat or harm another person alongside ascriptions of good character. However, if change in character could not be gradual, this would be equally strange, as it would seem to negate the purpose and value of character development and indeed, moral character itself. On the surface of things, Kant's discussion of moral worth in *GR* I might seem to do this – it might seem to negate the purpose and value of developing good moral character – but what I have argued so far in this thesis suggests otherwise. I will provide a complete response to these concerns in chapter three. However, I will in this chapter first examine the puzzle that surrounds the existence of a good will alongside bad inclinations, or bad empirical character. I will argue that one way bad empirical character can remain after the change of heart is a failure to carry through in developing the sympathies one has a duty to cultivate. This is fundamentally a matter of particular practical knowledge: because the inferential basis an agent has to act on does not change with the change of heart and this basis is important for moral action, she must develop this basis. In this way, a Kantian agent might possess a good will in the noumenal sense, be consistently motivated by the moral law, and yet still need to engage in the lengthy task of developing good empirical character. The general idea will be that full moral character in the Kantian sense is constituted by a good will in one's noumenal character and well-developed sympathy in one's empirical character. This avoids the counterintuitive conclusion that such an agent of good will simultaneously has both moral and immoral

motivations, thus deflating this puzzle and rendering Kant's account of moral character increasingly plausible.

Kantian Sympathy

Background to the Duty of Sympathy

Kant's most focused discussion of sympathy is found in the Doctrine of Virtue, which composes the second half of the *MM*. Here, Kant argues that the duty of sympathy follows from the more general duty of love to human beings. As he says, the general "duty of love for one's neighbor can [...] also be expressed as the duty to make others' *ends* my own (provided only that these are not immoral)."⁸⁸ When he introduces this duty, he emphasizes the aspect of constraint that it implies, in contrast with the feelings of pleasure or delight that often accompany love for another person. He says, "**love** is not to be understood as *feeling*, that is, as pleasure in the perfection of others; love is not to be understood as *delight* in them [...]. It must rather be thought as the maxim of *benevolence* (practical love, which results in beneficence)."⁸⁹ Though he is describing something – love – that at times involves the experience of very intense emotions, Kant is concerned that the good actions of love are not made dependent on an agent's experience of those emotions.

That the benevolent action of love should not be based on pleasure also follows from the idea stated earlier in the *MM* that there is no direct duty to seek one's own happiness, since the concept of duty entails constraint (for the purposes of morality) and people do not typically need be constrained to pursue their own happiness, pleasure, or delight, since Kant

⁸⁸ 6:450.

⁸⁹ 6:449.

says we already do this “unavoidably”.⁹⁰ Kant’s definition of happiness as “the satisfaction of all our inclinations” is perhaps simplistic, and his assertion that “[w]hat everyone already wants unavoidably, of his own accord, does not come under the concept of *duty*” may be misleading, since it overlooks that there seem to be many cases in which people want to pursue the happiness or good of others even to their own detriment.⁹¹ It is also worth noting that he holds in the *GR* that there can be a duty to obtain one’s own happiness, in the case that a want of happiness can be a temptation to wrongdoing.⁹² So his broader view is slightly more subtle, if not entirely consistent.

In any case, his statement in the *MM* could be corrected to say that people *typically* pursue their own happiness and do not need to be constrained to do so, which seems to be true. What Kant leaves unsaid here is important: love may and perhaps ought (in a non-moral sense) to involve delight in others and we do not typically need duty to tell us to pursue that delight. Within the constraints of morality there is freedom to pursue this delight when and how we wish. Instead, what Kant is emphasizing here is that we have a moral duty to act beneficently for the sake of others. It is helpful to note that the sort of love he is describing is not necessarily intimate love, but something like a general love of other people. This makes it easier to understand why he says love should not be thought of as delight, for it may perhaps be rare that we positively delight in those toward whom we have duties. Yet if even intimate love, which we would expect to involve delight and deep emotion, did not act beneficently toward the other in the temporary absence of those emotions, it would seem less than complete.

⁹⁰ 6:386.

⁹¹ A806/B834 and 6:386, respectively.

⁹² “To assure one’s own happiness is a duty (at least indirectly); for, want of satisfaction with one’s condition under pressure from many anxieties and amid unsatisfied needs, could easily become a great *temptation to transgression of duty*” (4:399).

Aside from the happiness of the agent whose duty it is to love, for Kant the duty to love another person follows from the more general duty to share the ends of others, which means to contribute to their happiness. This is explained in the Introduction to the Doctrine of Virtue. Kant's idea is that a doctrine of virtue, where virtue involves fortitude in one's moral disposition (*Gesinnung*), contains two ends that are also duties: one's own perfection and the happiness of others.⁹³ The argument seems to be that all action involving choice is directed toward some end, so if we grant that freely chosen moral action is possible (action based on the concept of duty), then those actions that are moral must be directed toward some end or ends, and those ends must be duties.⁹⁴ Why one's own perfection is a duty seems obvious, for it enhances (and in the case of perfection would entail) one's ability to consistently act morally. Kant explicitly suggests that it is a duty to contribute to the happiness of others because happiness is conducive to moral behavior, so contributing to others' happiness enhances their ability to consistently act morally. "Adversity, pain, and want are great temptations to violate one's duty. [... And] happiness is merely a means for removing obstacles to [one's] morality."⁹⁵

It might be tempting to conclude that Kant believes increasing the happiness of others is in this way only ever instrumental. However, as above, we do not typically need to be constrained to act toward this end, so Kant believes that to speak of a duty to do so is in most cases going to be meaningless.⁹⁶ He continues to explain that it is for each to decide what they count as belonging to their happiness.⁹⁷ Since it is in general a duty for me to also make

⁹³ 6:380 and 6:385.

⁹⁴ See argument in 6:384-385. Kant also says in *GR II* that "[r]ational nature is distinguished from the rest of nature by this, that it sets itself an end" (4:437).

⁹⁵ 6:388.

⁹⁶ 6:387.

⁹⁷ 6:388.

others' ends my own, and the only strict condition on their end is that it cannot be immoral, this affords a significant degree of latitude to individuals in their definition and pursuit of happiness. It also affirms that happiness has a place in the moral life, according to Kant.⁹⁸ Elsewhere, in the *CPrR*, Kant acknowledges a rational kind of self-love. "Pure practical reason merely *infringes upon* self-love; inasmuch as it only restricts it, as natural and active in us even prior to the moral law, to the condition of agreement with this law, and then it is called *rational self-love*."⁹⁹ So, again, self-love, which would involve the pursuit of happiness, is not impermissible for a rational agent provided it is constrained by the moral law. On the whole, it appears to be an at times tacit, though important assumption of Kant's work that happiness is an integral part of human life and indeed of the moral life.¹⁰⁰ In any case, to contribute to the happiness of others is a duty. This duty is constituted by the duties of love and respect for others, and the duties of love are divided into duties of beneficence, gratitude, and sympathy.¹⁰¹

For Kant, gratitude consists in "*honoring* a person because of a benefit he has rendered us"; "the feeling connected with this judgment is respect for the benefactor".¹⁰² The justification for this duty is unclear. Houston Smit and Mark Timmons point out that Kant does not attempt to derive this duty from the moral law.¹⁰³ The closest Kant comes to

⁹⁸ 6:450.

⁹⁹ 5:73. Stephen Engstrom discusses this and agrees with this point in Stephen Engstrom, "The Concept of the Highest Good in Kant's Moral Theory" in *Philosophy and Phenomenological Research* 52:4 (December 1992), 752-755.

¹⁰⁰ One problem surrounding happiness or self-love is the unbridled pursuit of it; the second problem, to which Kant is acutely attuned, is the place that it occupies amongst other maxims. The Doctrine of Right deals with the first problem and the Doctrine of Virtue with the second. With regard to the second, specifically moral problem, self-love is permitted into the life of even the ideal moral agent, so long as that agent is fundamentally committed to the moral law. I will respond to this and associated challenges for Kant's position in chapters three and four.

¹⁰¹ 6:449-452.

¹⁰² 6:454.

¹⁰³ Houston Smit and Mark Timmons, "The moral significance of gratitude in Kant's ethics" in *The Southern Journal of Philosophy* 39:4 (2011): 318-319.

providing a justification for it seems to be his claim that its violation “can destroy the moral incentive to beneficence in its very principle”.¹⁰⁴ Though he speaks of principle here, it is difficult to see how a response of ingratitude would deter the beneficent *agent* who acts on the basis of moral principle. For such an agent could not be steered toward moral action by any emotions she might experience at the grateful or ungrateful response of recipients of her beneficence: this would run contrary to the definition of beneficence in the *MM* as well as the upshot of Kant’s familiar discussion of the philanthropist in *GR* I.¹⁰⁵ It might be that Kant is offering a consequentialist justification of sorts, by suggesting that ingratitude would generally discourage action that is at least externally benevolent, while gratitude would encourage it. This is roughly what Smit and Timmons suggest in the first instance.¹⁰⁶ However, this goes well beyond the text and would be to exclude matters concerning moral motivation, that is, motivation on the basis of the moral principle, which Kant specifically references in the passage quoted above.

In the second instance, Smit and Timmons suggest the duty of gratitude could be grounded in respect for the humanity of the beneficent person, as an application of the formula of humanity.¹⁰⁷ This connects with the idea that gratitude involves a feeling of respect for the benefactor, but it does not connect with the claim that ingratitude can destroy the moral incentive to beneficence. Instead, perhaps the duty of gratitude is best understood as a demonstration of respect for action on the basis of the moral law, or moral principle,

¹⁰⁴ 6:455.

¹⁰⁵ Regarding the definition of beneficence, Kant says: “[t]o be beneficent, that is, to promote according to one’s means the happiness of others in need, without hoping for something in return, is everyone’s duty” (6:453). He also cautions against the kind of beneficence that comes at no cost to oneself, since it is a “way of reveling in moral feelings” (ibid). Again, for his discussion of the philanthropist from the *GR*, see 4:398-399.

¹⁰⁶ They offer what they call a “role based justification of the virtue of gratitude”, suggesting that the role of gratitude is to promote the general happiness of a society and the social bonds of respect that hold it together (Smit and Timmons, “The moral significance of gratitude in Kant’s ethics,” 318-319).

¹⁰⁷ Ibid., 318.

such that in demonstrating respect for others who act on the moral law, one is both respecting their humanity and indirectly respecting the moral law. Respect of this sort could then be understood to preserve and reinforce the moral principle in the *recipient* of beneficence, rather than the benefactor. Furthermore, even if a benefactor were not acting on the basis of the moral law, visible expressions of gratitude toward her could have the effect of preserving and reinforcing the moral principle in observers. For the observation of ingratitude could provide a disincentive to beneficence for one who has not yet undergone the change of heart and grounding of character. As I argued in chapter one, inclinations toward moral action are of moral value, though not yet of moral worth; correspondingly, it would be of moral value to not damage such inclinations or encourage disinclinations to moral action.

It is in the context of these duties of love that we find the duty of sympathy. In the section heading concerning this specific duty Kant asserts that “[s]ympathetic feeling is generally a duty” (“*Teilnehmende Empfindung ist überhaupt Pflicht*”).¹⁰⁸ On the face of it, this is a rather puzzling statement, since it would seem either difficult or impossible to make oneself have or experience any kind of feelings. He continues,

“[s]ympathetic joy and sadness (*sympathia moralis*) are sensible feelings of pleasure or displeasure [...] at another’s state of joy or pain (shared feeling, sympathetic feeling). Nature has already implanted in human beings receptivity to these feelings. But to use this as a means to promoting active and rational benevolence is still a particular, though only a conditional, duty.”¹⁰⁹

On its surface, this statement may seem to contradict his account of moral worth from *GR I* in the ways I mentioned in chapter one and will deal with in more detail in chapter three. As I have said, I do not believe that because an action involves emotion it cannot be understood to have moral worth in a Kantian sense. In short, I believe apparent puzzles like this dissolve

¹⁰⁸ 6:456.

¹⁰⁹ Ibid.

under a clear and comprehensive understanding of Kant's project. Nevertheless, these words may seem inconsistent with Kant's requirement that morally worthy action be motivated by the moral law, so either he is being inconsistent or things are not as they initially seem and what he is saying about sympathy is compatible with his account of moral worth.¹¹⁰ If he understands sympathy as something other than mere feelings of pity or sorrow for another's condition, it does not have to contradict his idea of moral worth.¹¹¹ I argue that this is indeed the case: Kant understands sympathy as a cognitive disposition meant to serve the practical-epistemic function of informing an agent about the needs of others, so that an agent can better perform her duty toward them. I am therefore arguing that Kantian sympathy is a moral aptitude of the kind outlined in chapter one. Once sympathy is understood in this way, we might say there is a duty to encourage sympathetic feelings and use them in exercising this disposition, but not a direct duty to have them in the strict sense, since feelings are not directly under anyone's control.¹¹² This is likely why he says sympathetic feeling is only a

¹¹⁰ Since Kant himself edited this work, this section heading must be attributed to him. This is unlike other late work of Kant's which was edited by others, like Benjamin Jäsche and Theodor Rink, including the *Logic* of 1801, the *Physical Geography* of 1802, and the *LP* of 1803. (See Manfred Kuehn, "Kant's *Metaphysics of Morals*: the history and significance of its deferral" in *Kant's Metaphysics of Morals: A Critical Guide*, ed. Lara Denis (Cambridge: Cambridge University Press, 2010), 9.)

¹¹¹ One could translate Kant's words here – *Teilnehmende Empfindung* – into something other than 'sympathy.' For example, 'participatory perception' would be a more literal translation of the German and would more easily suggest the cognitive disposition (or a disposition more open to being shaped by cognitive processes) that I argue Kant has in mind here. 'Moral perception' might also be used: though it would be a more creative translation, it too could arguably better convey what Kant has in mind. In conversation, Paul Guyer suggested 'sympathy' is the best translation, since the English 'sympathy' from Adam Smith's *The Theory of Moral Sentiments* was translated into the German '*Teilnehmende Empfindung*' during Kant's time (Paul Guyer, personal conversation with Carl Hildebrand, Istanbul, June 2016). It is worth noting that for Smith, too, sympathy was something more robust than the mere feelings of pity or sorrow we typically take it to refer to nowadays. Though it retains important differences from Kant's conception, it had an important cognitive component meant to serve similar purposes to those envisioned by Kant and suggested in this thesis. For example, Smith says: "[t]he compassion of the spectator must arise altogether from the consideration of what he himself would feel if he was reduced to the same unhappy situation [as one he observes suffering], and, what perhaps is impossible, was at the same time able to regard it with his present reason and judgment." (Adam Smith, *The Theory of Moral Sentiments* (London: Penguin Books, 2009), I.1.) I have chosen here to follow the conventional translation by Mary Gregor and use the term 'sympathy.' Though I do this with some reservation, because I take it to mean something different (more) than the current, common understanding would suggest.

¹¹² Kant says: "one cannot be put under obligation to have feelings" (6:449).

conditional duty. In what follows, I will outline the Kantian disposition of sympathy and explain how it contributes positively to Kant's moral psychology and conception of moral character. I will first suggest several ways that sympathetic feeling might aid moral understanding. This should demonstrate the kind of value emotion has and the function it can perform on a Kantian (or any other intellectualist) account of moral character. I will then explain how the disposition of sympathy fits with the categorical imperative (or moral law) and enables an agent to more effectively fulfill both her wide and narrow duties. I will conclude by looking at two examples of what might be considered a failure in sympathy, one from history and one from literature.

The Disposition of Sympathy

Though sympathy is a moral aptitude and so a predominantly cognitive disposition, feelings or emotions may be useful to it in several important ways, since they can help us to perceive and come to understand the world around us, including its moral features. Kant seems to agree with this when he suggests that sympathetic feelings are a useful means to promoting rational benevolence.¹¹³ First, they may help us to understand what brings happiness to others in a general sense, thereby enabling us to contribute to the happiness of others. It seems difficult to deny that emotions contribute to happiness or human flourishing. Kant defines happiness as the satisfaction of one's inclinations (as above) or "satisfaction with one's state" and I submit that most of us would think that the happy or good life would to some degree involve both positive feelings and the lack of negative feelings.¹¹⁴ For example, we would likely not say that someone who has strong relationships, is financially

¹¹³ 6:456.

¹¹⁴ 6:387.

well off, physically fit, and yet suffers from severe depression has a happy life. Their life may have many good qualities and they may experience pleasure as a result of these qualities, but they would nevertheless hope to eliminate their depression (which we could imagine entails negative feelings, a lack of positive feelings, or both) and until it were eliminated it would be difficult to say their life is a happy one. Perhaps it is for concerns related to this that Kant adds to his definition of happiness that it is “satisfaction with one’s state, so long as one is assured of its lasting”.¹¹⁵ In any case, the increase of positive emotions and the elimination of negative ones will contribute to happiness, and understanding the emotions of others will be crucial in helping an agent to understand how she can contribute to the happiness of others. This would require an agent to understand what state various emotions might signify when observed in another person. It would also require that it is in most cases possible to perceive another’s emotions. Assuming these things are possible, emotions could serve as a way of communicating to an agent what brings happiness to another. The disposition of sympathy might therefore require a sensitivity to facts concerning the well-being of others as communicated by such emotions – it would be emotionally outward looking in this way.

Then, an agent might come to understand the facts or meaning behind another’s emotions by analogy to their own experience of such emotions, for example, by imagining how one might experience the pain of hunger and feelings of dislocation or possibly despair at having to live homeless. This could be what Kant has in mind when he says, “it is a duty not to avoid the places where the poor who lack the most basic necessities are to be found but rather to seek them out, and not to shun sickrooms or debtors’ prisons and so forth in order to

¹¹⁵ Ibid. This is in line with what Kant says about happiness in the *GR*, namely that “for the idea of happiness there is required an absolute whole, a maximum of well-being in my present condition and in every future condition” (4:418). As Aristotle says, one swallow does not make a summer’s day (*NE*, 1098a18).

avoid sharing painful feelings one may not be able to resist.”¹¹⁶ So it seems sympathy would also require an understanding of emotion inwardly, as it were, of one’s own emotions and likelihood to feel a certain way in certain circumstances. In this way, emotions would enable an agent to perceive the moral fact that others are beings whose happiness is of equal objective value to her own. This is another reason why Kant says the cultivation of sympathy is a particular yet only conditional duty; for it is not a direct duty in itself but is a duty insofar as it enables an agent to better fulfill the duties of self-perfection and of contributing to the happiness of others.

Second, in addition to communicating what generally constitutes happiness (or pain) for another, emotion might further help an agent to understand what particular circumstances call for action.¹¹⁷ This is especially important in light of the fact that the same actions will not always have the same effect on different people (will not cause the same amount of pleasure or relief from pain), so an agent must be attentive to the needs of particular individuals as conveyed by emotion, as well as the effect of her actions on particular individuals – something that emotion can also help her to do. For example, not everyone desires the same kind of response from a friend amidst grief caused by a broken relationship. Sometimes it will be appropriate for a friend to speak encouraging words or remind one of facts about the relationship that make the prospect of life without it more appealing, while other times or with other people it will be appropriate to not say anything at all. An ability to read and

¹¹⁶ 6:457.

¹¹⁷ Sherman similarly suggests that emotions help us track what is morally salient in our circumstances, disclosing to us ways in which we might serve the well being of others: “[f]irst, emotions serve as *modes of attention* that help us to track what is morally salient as morally salient in our circumstances, and thus locate possible moments for morally permissible and required actions. Sympathy, for example, draws us to occasions of distress or need. We attend in a charged and alert way, taking in what detached reason or perception might miss.” (Nancy Sherman, *Making a Necessity of Virtue*, 145-146.) I hope to have added some detail to her account.

respond to these situations as they unfold and in a way that contributes to the well-being of the other will require suitably tuned emotional dispositions.

Thirdly, emotions can contribute to the provision of an index for prioritizing action according to the intensity of the needs of others to which we might respond. In other words, they can help to communicate the magnitude of another's suffering so that we can understand, measure, and prioritize fulfillment of wide duty in actions toward others, especially given that we have limited personal resources. Appropriately tuned feelings of sympathy may help an agent understand that she ought to respond first to the person lying injured and helpless on the street before she responds to someone in a less worse-off state.

Emotion will not be an infallible guide, nor will the guidelines it provides be infallible: it might sometimes distort the relative importance of values and at other times entirely misrepresent them, producing false positives and negatives. This is why Kant insists that emotion must always be structured by and subordinate to the moral law. I will explain what this means in more detail shortly. In the interim, it is worth noting a puzzle that arises concerning the role of emotion, including sympathy, in moral motivation. Kant says in concluding his discussion of the duty of sympathy that it is "one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish."¹¹⁸ This seems to suggest that sympathetic feeling may play a motivational role in the agent's performance of good actions. Could this be correct? In recent years, a number of interpretations have been suggested to support the idea that an agent might be motivated by emotion as well as the moral law, so that actions motivated (at least in part) by emotion

¹¹⁸ 6:457.

might still have moral worth for Kant.¹¹⁹ I will present my own interpretation of the interaction between emotion and moral motivation in chapter three. Before I do so, I would briefly like to suggest that emotion's involvement in moral action does not prevent the moral law from taking up a dominant role. The moral law may provide a kind of rational oversight for action, preventing us from being misguided by emotion and ensuring we do not perform actions contrary to the moral law. It could thus act as a brake on actions of this kind, so that its involvement is only obvious when the agent is inclined to act in a way that contradicts it. This might mean that, if an agent has trained her emotions effectively, the involvement of the moral law would hardly be noticeable at all to her on a regular basis. Of course, the moral law will also remain the standard by which some emotions, like sympathy, will be strengthened while others, like envy or schadenfreude, will need to be weakened or expunged.¹²⁰

Interestingly, this account of the usefulness of emotion shows how emotion can aid an agent in the fulfillment of her duty, but it does not show that these sorts of emotion have an independent value of their own. It could be possible for Kantian agents to value these things if they choose to do so, yet there seems to be no ground for granting them value independent of this on the Kantian system. And as I have pointed out, Kant clearly holds that the value of such feelings is only conditional on their support of the moral law. He mentions this point more than once, emphasizing that the affective elements of sympathy (*mitleidige natürliche (ästhetische) Gefühle*) are of subordinate value to sympathy as active participation in the fate of others (*tätige Teilnahme an ihrem Schicksale*):

¹¹⁹ For example, see the work of Richard Henson, Barbara Herman, Barcia Baron, and Nancy Sherman, as cited in the introduction to this thesis (cf. note 17).

¹²⁰ Kant provides a list of vices and virtues at 9:492 and includes envy and schadenfreude among the vices.

while it is not in itself a duty to share the sufferings (as well as the joys) [*Mitleid (und so auch Mitfreude)*] of others, it is a duty to sympathize actively in their fate [*tätige Teilnahme an ihrem Schicksale*]; and to this end it is therefore an indirect duty to cultivate the compassionate natural (aesthetic) feelings [*mitleidige natürliche (ästhetische) Gefühle*] in us, and to make use of them as so many means to sympathy [*Teilnehmung*] based on moral principles and the feeling [*Gefühl*] appropriate to them.¹²¹

Aside from this, he says there is no duty to share in such feelings of *Mitleid* or *Mitfreude*. He also offers a reason against such feelings, saying that when they do not result in moral action they only increase suffering in the world and that is a bad thing: “[i]n fact, when another suffers and, although I cannot help him, I let myself be infected by his pain (through my imagination), then two of us suffer, though the trouble really (in nature) affects only *one*.”¹²² So it seems that Kant is fairly consistent in ruling out the independent value of sympathetic feeling, and rendering it subordinate to the moral use, or excluding it from moral value on grounds independent of the moral law. The idea is that if sympathetic feeling had value independent of the moral law, it might follow that in some (perhaps many) cases it would be considered good for an agent to sympathize and hence to some degree suffer emotionally when they become aware of the suffering of others. This might remain the case even when the agent could not do anything to alleviate the suffering they are aware of. The outcome of this would be an increase in overall suffering in the world, which does not seem desirable. However, it is worth pointing out that the distinction between sympathy that serves this practical-epistemic (moral) function and sympathy that brings unnecessary suffering to its possessor may not be as sharp as Kant envisions.

Further, it must be granted that there will be cases in which moral sympathy as Kant understands it will involve little or no experience of emotion. We might imagine the case of a

¹²¹ 6:457.

¹²² Ibid.

surgeon who has a good will, performing surgeries with the purpose of healing patients, a keen ability to recognize pain in her patients and respond to alleviate that pain as best she can, and yet does not suffer along with her patients when they experience pain. (Perhaps this is because she routinely performs surgery that is painful due to the absence or ineffectiveness of painkillers, so has become inured to seeing patients suffer.)

Yet most instances of sympathy are not like this and do not allow for as clear a separation between the belief that another is suffering and the experience of at least some suffering in the person who recognizes that suffering.¹²³ Most of us tend to respond to the recognition of suffering with a feeling of pain or sadness; this is often expressed when we say, for example, that we ‘feel’ for someone in difficulty or that ‘our heart goes out’ to them. As Williams has argued, it may be more desirable on the whole that we experience these emotions, and we might regard someone who does not experience them as morally impoverished.¹²⁴ This may be why Kant at times appears caught in general ambiguity or tension in this discussion: he may not entirely believe that sympathetic feelings have no moral value or that this distinction is so sharp.¹²⁵ Perhaps part of him holds on to the more commonsense or humane view that these feelings do have some independent value. Nevertheless, this is not his official position and he frequently emphasizes sympathy’s role in aiding the performance of duty and its consequent instrumental value.

¹²³ Further, in the case of the surgeon it is difficult to imagine she could develop such a keen ability to perceive and respond to patients’ suffering without at some earlier point registering that suffering emotionally, i.e., feeling some pain or sorrow in herself as a result of observing it in others and by analogy with her own experience.

¹²⁴ See chapter one (“Person’s, character, and morality”) in Bernard Williams, *Moral Luck* and chapter thirteen (“Morality and the emotions”) in Bernard Williams, *Problems of the Self*.

¹²⁵ For example, the title of this section says that “[s]ympathetic feeling” (*teilnehmende empfindung*) is a duty, yet, as noted, he says elsewhere that one cannot be put under obligation to have feelings (6:449) and that sympathetic feelings on their own (here referred to as compassion) do not have moral value.

I will now look in more detail at how the disposition of sympathy fits with the moral law and at how it enables an agent to better fulfill those duties determined by the moral law. As discussed in the *GR*, the categorical imperative, or moral law, is understood to have at least four formulations: the formula of universal law, of ends in themselves, of the kingdom of ends, and of autonomy. At one point, Kant discusses the relation between the first three formulas by drawing an analogy with the class of ‘quantity’ from the categories of judgment in the *CPR*.¹²⁶ He says that the formula of universal law provides the form of the moral maxim (i.e., moral law), the formula of ends in themselves provides the matter, and the formula of the kingdom of ends provides the complete determination. Of most immediate significance to the moral agent and hence to moral psychology, are the first two formulations, for the agent must at least at some point be conscious of the dignity and equality of rational beings in order to take those things as reasons and act on the basis of the moral law. (It does not *prima facie* seem necessary that she is aware that her actions are autonomous in Kant’s sense or that they harmonize with a possible kingdom of ends in order to be motivated by the moral law.)

The formula of universal law requires that one act with a coherent set of maxims (a maxim is a reason for performing an action).¹²⁷ In *GR* II, Kant reviews four cases in which a moral question is posed and the moral law is employed to provide an answer.¹²⁸ One such example concerns a man who borrows money and promises to repay it without intending to fulfill that promise. This man would be acting on the maxim, “when I believe myself to be in

¹²⁶ 4:436.

¹²⁷ Kant’s clearest definition of a maxim is in the *CPR*. “Practical laws insofar as they are at the same time subjective grounds of actions, i.e., subjective principles, are called **maxims**. The **judgment** of morality concerning its purity and consequences takes place in accordance with **ideas**, the observance of its laws, in accordance with **maxims**” (A812/B840).

¹²⁸ See 4:421-423.

need of money I shall borrow money and promise to repay it, even though I know that this will never happen.”¹²⁹ Kant argues that if this man were to ask himself what would happen if his maxim were to become a universal law, he would discover that he is willing an incoherent or contradictory set of maxims. For he would simultaneously be willing that a lender lend him money on the condition that he pay him back, and also that he (in reality) not pay him back. This practical contradiction reveals that the man is being deceitful and unfair – making an exception of himself – as there is a disparity between how others are expected to act and how he expects himself to act.¹³⁰ As Christine Korsgaard helpfully points out, intuitively speaking, universalizing one’s maxims identifies such cases of unfairness, deception, and cheating.¹³¹ As Kant suggests, the condition that one’s maxims be universalizable provides a formal constraint on action and in this way makes up the formal aspect of the moral law. However, this account is incomplete taken on its own. One might respond to the above scenario by asking why we should care about these things: why should it matter if we make an exception of ourselves in this way? A material condition is required to enable agents to identify objects for this equal treatment. In the analogy with the class of quality, Kant says that the “three ways of representing the principle of morality are at bottom only so many formulae of the very same law, and any one of them of itself unites the other two in it.”¹³² So one formula will not be complete without the others.

¹²⁹ 4:422.

¹³⁰ I am here drawing on elements of both Onora O’Neill’s and Christine Korsgaard’s interpretation of the formula for universal law; see chapter 5 of O’Neill (Nell)’s, *Acting on Principle* (New York: Cambridge University Press, 1975) and chapter 3 of Korsgaard’s *Creating the Kingdom of Ends* (Cambridge: Cambridge University Press, 1996).

¹³¹ Korsgaard, *Creating the Kingdom of Ends*, 92.

¹³² 4:436.

The formula of ends in themselves, or humanity, is therefore necessary to identify what, or more properly, *who* is to be the object of the equal treatment described above. It provides the necessary material aspect of the moral law, and the formula of universal law cannot be usefully employed without it. It holds that all rational beings (humanity) have dignity and one therefore ought always to act in such a way that one treats humanity as an end and never as a mere means.¹³³ It adds content to the universal law formulation by identifying the proper objects of moral attention or respect, namely rational, or we might say human, beings. It shows Hegel's charge of "empty formalism" to be misconceived.¹³⁴ The idea is that each being endowed with even the most basic capacity of choice is capable of understanding its choices (or the ends for which it acts) as good, and so can recognize itself as a kind of authority on what is good in general. This recognition entails respect for oneself in virtue of this capability. From this it follows that any being with such a capability deserves similar respect.¹³⁵ Such beings stand on equal footing with one another, which is why Kant believes making an exception of oneself (as in cases of unfairness, deception, and cheating) is wrong: it violates the formula of humanity. The dignity of all persons and action towards them that corresponds with this dignity is foundational for Kantian ethics. It is a practical analog of the principle of non-contradiction in classical logic and a kind of rational intuition, or a "fact of reason" as he says in the *CPrR*.¹³⁶

Sympathy attunes us to this value – the dignity of human beings, other human beings in particular – so that we are able to more effectively fulfill our duties toward them. As

¹³³ 4:429.

¹³⁴ Cf. G.W.F. Hegel, *Outlines of the Philosophy of Right*, section 135.

¹³⁵ This follows Allan Wood's interpretation in his book, *Kantian Ethics* (Cambridge: Cambridge University Press, 2008), 90-93.

¹³⁶ 5:43. For the idea of the moral law as a rational intuition, see Ralph Walker, "The Rational Imperative: Kant Against Hume" in *Proceedings of the British Academy* LXXIV (1988), 113-133.

explained in the *GR* and elsewhere, our duties are of either a narrow or wide nature. Though this distinction cannot be a sharp one, narrow duties typically involve specific actions and are “strict” or “unremitting”, which means that they admit of no exceptions.¹³⁷ The above example of a man who lies about repayment in order to secure a loan is an instance of a violation of a narrow duty. Kant believes lying is never permissible because the kind of maxim lying involves “necessarily contradict itself” and “make[s] use of another human being *merely as a means*”.¹³⁸ Among his examples in the *GR*, he places suicide in the same category for what he believes are the same reasons.¹³⁹ Wide duties, on the other hand, involve exhortations to action and do admit of exceptions in the sense that not every opportunity to fulfill a wide duty must be acted upon. This is because an agent may, for example, fail to help someone in need without willing a contradictory maxim or directly violating another’s dignity. (In the *GR*, Kant discusses the wide duties of developing one’s talents and helping others in need, which he believes apply to all rational beings.¹⁴⁰ In a slightly different form, it is these two duties that he later develops in detail in the *MM*.) Nevertheless, an agent must in at least some circumstances help others who are in need; it is just that there is room for discretion regarding exactly when and in what circumstances they do this.¹⁴¹ As Onora O’Neill explains with regard to the wide duty of developing one’s talents,

Kant points out that there is no contradiction in the idea of a system of nature in which men entirely neglect their talents. But he claims that there is a

¹³⁷ 4:424.

¹³⁸ 4:422 and 4:429, respectively.

¹³⁹ Ibid.

¹⁴⁰ See 4:422-423 and 4:430.

¹⁴¹ Though discretion is permitted (presumably so that an agent is not so rigorously engaged in moral action that it prevents them from also acting for the sake of their own goals, life projects, and well-being – a point I will discuss further in chapter four). The degree to which an agent acts to fulfill these wide duties might tell us something about the general praiseworthiness of an agent. This could be why Kant describes wide duties as “meritorious” (4:424).

contradiction in a rational agent's willing to be a part of such a system. Rational beings have ends; they also will some sufficient means to these ends, and the development of some talents is a part of the means to ends of any sort. So rational agents must make it an end to develop at least some of their own talents. If they do not have this end, they cannot claim to have any other end.¹⁴²

The cultivation of a sympathetic disposition in oneself is, in one sense, already a fulfillment of wide duty, since it is a response to the duty of self-perfection.¹⁴³ Yet, in another sense and perhaps more importantly, it serves as an important aid elsewhere in the fulfillment of both types of duty, narrow and wide.

Since Aristotelian virtues of character consist in more or less specific dispositions (such as courage, liberality, or truthfulness), they are relevant to specific kinds of facts and encourage specific kinds of action. For example, an agent who possesses the virtue of courage and not liberality may be courageous in battle yet not responsive to the needs of fellow citizens who live in poverty. Unlike these Aristotelian virtues, Kantian sympathy, which we may think of as a kind of virtue, in its ideal form entails an awareness of the much more general moral law (in the way described above) as well as a sensitivity to any perceivable moral facts relevant to the dignity and equality of others. An agent who possesses the ideal disposition of sympathy will therefore perceive all facts that are relevant to the moral duties under which she is obligated. Kantian sympathy is in this way similar to Aristotle's intellectual virtue of *phronēsis*, since it consists in a moral perception sensitive to the appropriate moral facts an agent requires in order to properly fulfill her duty. Aristotle holds that all virtues share in this one cognitive capability or knowledge component: "full

¹⁴² O'Neill, *Acting on Principle*, 88.

¹⁴³ Interestingly, this is in spite of the fact that Kant places it under the duty of love, indicating it is meant to serve the happiness of others. It could be argued that it fulfills both of these duties outlined in the *MM*, including self-perfection, because Kant says that the duty of one's perfection "can therefore consist only in *cultivating one's faculties* (or natural predispositions)", including "the cultivation of one's *will* (moral cast of mind), so as to satisfy all the requirements of duty" (6:387).

virtue cannot be acquired without *phronēsis*” and “one has all the virtues if and only if one has *phronēsis*, which is a single state.”¹⁴⁴ As John Cooper explains, “the ethical virtues actually ‘cluster’ around practical wisdom [*phronēsis*] because *it* is the virtuous condition that gives direction to, and shapes and controls the workings of, these other virtues, the virtues of feeling.”¹⁴⁵

Sympathy operates in a very similar way for Kant. It is itself conceptually or necessarily tied to the foundational Kantian value of the moral law, which provides its moral content. Given the sort of empirical beings we are, we need some degree of sympathy in order to understand what the moral law requires of us in particular circumstances: as Kant says, it is “one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish.”¹⁴⁶ A significant difference is that for Kant it is not a disposition to choose but a disposition that informs choice by providing information that makes the morally right action perspicuous to the agent and the right decision therefore possible. Another difference is that Kant pays much less attention to individual virtues of character. Sympathy performs the role of shaping and directing further feelings and dispositions, ensuring that they remain under the direction of the moral law.

I will now look at how the disposition of sympathy enables an agent to better fulfill both narrow and wide duties, developing some of my earlier points concerning the value of

¹⁴⁴ *NE*, 1144b31-33 and 1145a1-3, respectively. Aristotle presents this position in contrast with the view that Socrates presents in the *Protagoras*, which he rejects. Immediately following the statement that full virtue cannot be acquired without *phronēsis* he continues, “[t]hat is why some say that all the virtues are [instances of] *phronēsis*, and why the inquiries Socrates used to undertake were in one way correct, and in another way in error. For insofar as he thought that all the virtues are *phronēsis*, he was in error; but insofar as he thought they all require *phronēsis*, what he used to say was right” (*NE*, 1144b16-23).

¹⁴⁵ John Cooper, “The Unity of Virtue” in *Social Philosophy and Policy* 15:1 (1998): 266.

¹⁴⁶ 6:457. Kant says that sympathy is “only a conditional duty” (6:456). As I have explained above, I take it this means it is conditional on the moral law. I also take it to be conditional on the fact that we are embodied, empirical beings, as I have also explained. A community of divine or holy wills, for example, with perfect knowledge of themselves and one another’s condition, would not need the disposition of sympathy, since their knowledge is perfect and their “volition is of itself necessarily in accord with the law” (4:414). So, as Kant understands it, sympathy is a human virtue.

sympathetic feeling to moral understanding. Sympathy aids in the fulfillment of wide duty because it enables the agent to perceive opportunities for acting she would otherwise not perceive, if she had not developed the capacity to share in an understanding of the suffering of others. As suggested above, sympathetic feeling may enable an agent to perceive moral facts and circumstances for action she might otherwise be blind to. However, this enhanced perception does not rely on feeling and may be developed in its absence. The point is that the agent develop this sympathetic perception, not that she experience sympathetic feelings that may be conducive to it (and in some cases, will not be conducive to it). At an early point in the process the agent must realize that other rational beings deserve the same moral treatment as herself. Once that realization has dawned, sympathy can help her to realize what the moral law requires in the particular circumstances she might find herself in. For example, if an agent were to choose to fulfill the duty to develop sympathy by volunteering in the palliative care ward of her local hospital, she might realize she has a duty to somehow support her neighbor who suffers from terminal cancer (or alternatively, say, give money to a cause that supports people suffering from terminal illness).¹⁴⁷ Without having previously volunteered, she might callously neglect her neighbor, not perceiving the gravity of the neighbor's situation. Her neglect, in this case, would stem from a failure to fulfill the duty to develop sympathetic feeling and she would be guilty of a moral failure according to Kant.¹⁴⁸

¹⁴⁷ Kant is careful to note that sympathy is not the same as pity and pity is to be avoided (6:457).

¹⁴⁸ Because the duty to cultivate one's sympathy is a wide duty, it may be done in a number of different ways. We do not all have to work in the palliative care ward or pursue every possible form of activity that could develop sympathy. Kant is not saying that we must be moral saints, but that we must cultivate morally appropriate sympathy in some way. Presumably, I could develop sympathy as a result of one activity (e.g., being a parent) and it would increase my sympathy in other domains (e.g., being a friend), but this gets into more complicated discussions about trait-consistency that cannot be covered in the scope of this thesis. Here, my point is that, given that the duty to cultivate sympathy is a wide duty, Kant leaves it open to us to determine how we might go about doing this.

We can imagine a variety of similar scenarios in which an agent takes initiative to make herself aware of the suffering of others in order to aid them directly as well as to understand how she might aid yet others in other circumstances. In summary, we may enumerate at least four ways that sympathy may aid an agent in the fulfillment of wide duty. First, as mentioned, the development of a sympathetic disposition is itself a response to and partial fulfillment of the wide duty of self-perfection. Second, in the process of developing this disposition by, for example, as Kant suggests, seeking out “where the poor who lack the most basic necessities are to be found”, sickrooms, and debtor’s prisons, the agent will likely be provided with opportunities to fulfill wide duties in this initial encounter with those in these circumstances. Third, having become aware of how those in these circumstances suffer, the sympathetic agent will seek out further circumstances beyond this and beyond those to which she would regularly be exposed, in which she will be able to contribute to the welfare of others. Fourth, she will have an enhanced capacity to perceive that suffering to which she may regularly be exposed but otherwise blind to, as in the case of the suffering neighbor above. Or, to develop Kant’s example, in the way that someone who has spent time in a debtor’s prison may better understand the circumstances of a colleague or friend who has recently run into severe financial trouble. This understanding might encourage her to come alongside this person and aid them in locating resources, legal counsel, and personal counseling, thus fulfilling a wide duty. Without the practical knowledge that sympathy affords, she may have otherwise done nothing and thus forfeited opportunities to fulfill wide duties.¹⁴⁹

¹⁴⁹ There is a problem here in identifying where an individual agent’s wide duties end, often referred to as the problem of moral demandingness. I will respond to this in more detail in chapter four.

I would also like to argue that sympathy aids in the fulfillment of narrow duty. Unlike wide duties, which allow many possible actions to count in their fulfillment, narrow duties are very specific and either require or forbid very specific actions. Because sympathy attunes us to the suffering of others and violations of narrow duties result in such suffering, if an agent has a well-developed sympathetic disposition and is motivated by the moral law, then she will also better perceive possibilities for specific violations of the dignity and well-being of others, and will avoid these violations. It is useful in both major and minor violations of narrow duty alike. A comparatively minor (though inexcusable) violation might include actively disrespecting someone because of the type of work they do, in a way that violates this person's dignity. A disposition of sympathy might enable an agent to anticipate the feeling of belittlement and shame such disrespect would bring to the other person, thus signaling to this agent that disrespect would be a violation of that person's dignity. This agent might then decide to respect that person's dignity by refraining from disrespect, thereby not violating the moral law. (This raises the question of whether the fully sympathetic agent would have such an impulse in the first place, an objection I will discuss in more detail toward the end of this chapter.)

Sympathy may also aid in the prevention of major, even gross violations of narrow duty. Sympathy's place among moral duties also suggests an answer to the question of how Kant might respond to honest mistakes in moral perception. An extreme example that touches on both of these points might be the (allegedly) honest failure of Nazi war criminals like Adolf Eichmann to realize that an entire category of people is deserving of the same moral treatment as oneself. This is, of course, an extremely difficult thing to understand and a skeptical response to a claim like this is appropriate. Standard intuition would say that figures

like this, even if their mistakes are honest, could not be acting with moral worth and would clearly not deserve praise for their actions, instead deserving blame. Eichmann's mistake is so egregious it is difficult to take him at his word.¹⁵⁰ Though we could reasonably expect better of someone, we might speculate that Eichmann was so deeply enmeshed in a regime and culture that incessantly and forcefully claimed the non-humanity of Jews that, given a weak mind and easily manipulated personality, he succumbed to believe this was in fact the case. In scenarios like this, a disposition of sympathy would attune such individuals to the basic fact that their decisions and actions involve the well-being of other people who deserve the same respect and treatment as themselves. The sympathetic person would recognize that there are no justifiable grounds to assert the non-humanity of Nazi victims. The sympathetic person would then think into the other person's place (the place of the victim), think consistently (whether they like to be treated the way the victims are to be treated), and think for themselves (in this case, contrary to an ideology that denies the victim's humanity), something Kant advocates elsewhere, in the *AN*.¹⁵¹

This is similar to the appeal made to Shylock in Shakespeare's *Merchant of Venice*, namely to recognize another's humanity and to think in this manner. When Shylock appears to the court to demand the pound of flesh from Antonio, the court is baffled at his apparent inability to imagine himself in Antonio's place. The Duke appeals to Shylock's courtesy,

¹⁵⁰ In chapter three I will draw a distinction between full and mitigated moral worth, specifying that agents who are morally motivated and make honest mistakes may still qualify for a mitigated ascription of moral worth, provided they are doing the intellectual best that can reasonably be expected of them in assessing how best to act morally in the situation they find themselves in. A figure like Eichmann does not qualify for full moral worth, because he has not acquired appropriate moral sympathy (as I will explain further in chapter three); he also does not qualify for mitigated moral worth, because it is reasonable to expect better of him even given the epistemic challenges of his situation. We might add that he deserves blame because the consequences of his actions, whatever their motivation, are so deeply detrimental to human dignity and well-being. In general, Kant does not deal in reasonable expectations; nevertheless, I will suggest that we may do differently, while developing his general position in this direction, provided we are careful to not attribute to him things he did not or could not have believed.

¹⁵¹ 7:200

pity, and human gentleness, asking “[h]ow shalt thou hope for mercy, rend’ring none?”¹⁵² Yet Shylock remains strangely blind to the suffering the extraction of his payment would cause to Antonio. To be switched off, as it were, to another’s suffering to the degree that figures like Eichmann and Shylock are, indicates more than a severe lack of sympathy. For Kant, this kind of mistake is a violation of one’s duty, not only as we might see it, but as how these individuals ought to see it.

The mistake of being blind to another’s suffering is more common than we might think; it can involve itself in our lives in mundane ways and on a scale much smaller than that of Shylock or Eichmann. It can be a sin of omission rather than of commission. We often overlook the suffering of others and our duty to respond to it appropriately, and we often cause suffering to others out of ignorance. It is easy to slip into actions and patterns of living by which we treat those who are different from us (or simply those that are not us) as somehow not deserving of equal consideration as us. In that way, we treat them as somehow less than fully human, or not in keeping with their dignity. This might happen in response to familiar political questions around immigration, or in the pursuit of one’s own career at all other costs. In some respects, we may be closer to Eichmann than we realize, and perhaps some of us are lucky for reduced opportunity to do wrong due to our location in time and historical context. Across these diverse circumstances, given that the moral law provides us with a reason to see the value of other people, we have an obligation, also grounded in the moral law, to shape our faculties and emotions so that they will help us to perceive this value. A failure to shape our faculties so that we are readily able to perceive the value of others is a failure to fulfill a duty of sympathy and hence a moral failure for which we are responsible.

¹⁵² William Shakespeare, “Merchant of Venice” in *The Oxford Shakespeare: The Complete Works 2nd Edition*, eds. John Jowett et. al. (Oxford: Oxford University Press, 2005), 4.1.

Moral Anthropology's Contribution to Sympathy

A concern to avoid this kind of failure and fulfill the duty of sympathy is one of the most significant motivations behind Kant's project in the *AN*. In the *GR*, he explains that a metaphysics of morals will have two parts: a rational part that is based on a priori principles and an empirical part that is based on grounds of experience.¹⁵³ He gives the name "practical anthropology" to the latter.¹⁵⁴ He says in the *GR* that a metaphysics of morals "needs anthropology for its *application* to human beings", meaning that knowledge of empirical human beings is necessary to enable people to consistently act on the basis of the moral law.¹⁵⁵ As I have suggested, this can be taken to imply two things: first, that an agent needs to understand what in their own nature both aids and hinders moral action (as discussed in chapter one), and second, that an agent needs to understand the characteristics, habits, and desires of other people in order to know how to fulfill their duties toward them. In the examples above, I illustrated how a well-developed disposition of sympathy can alert an agent to the happiness and particularly the suffering of others. The enhanced perception that sympathy affords enables an agent to understand how they may act to alleviate the suffering of others and contribute to their happiness, hence helping them to fulfill obligations of wide duty; it also enables them to understand what would violate another person's dignity, hence helping them to avoid such violations and so fulfill narrow duty. Kant's project in the *AN* contributes to the development of the same enhanced, moral perception, which is why I argue it contributes to an understanding of the Kantian virtue of sympathetic perception.

¹⁵³ 4:388.

¹⁵⁴ Ibid.

¹⁵⁵ 4:412.

This is apparent in the preface to the *AN*, where Kant asserts that the goal of cultural progress is the acquisition of the kind of knowledge and skill that can be put to use in the world. The most important object in the world for whose sake this knowledge is to be used is the human being, since the human being constitutes the only end in itself of which we are aware through our senses:

All cultural progress, by means of which the human being advances his education, has the goal of applying this acquired knowledge and skill for the world's use. But the most important object in the world to which he can apply them is the human being: because the human being is his own final end. – Therefore to know the human being according to his species as an earthly being endowed with reason especially deserves to be called *knowledge of the world*, even though he constitutes only one part of the creatures on earth.¹⁵⁶

He goes on to explain that knowledge of the human being is of two kinds, physiological and pragmatic. Physiological knowledge “concerns the investigation of what *nature* makes of the human being; pragmatic [knowledge], the investigation of what *he* as a free-acting being makes of himself, or can and should make of himself.”¹⁵⁷ The focus of the *AN*, as its title makes obvious (*Anthropology from a Pragmatic Point of View*), is the attainment of this pragmatic knowledge. In Part II of the *AN*, this pragmatic knowledge is developed along two dimensions corresponding with the meaning of anthropology mentioned above, knowledge of people as individuals (what we would nowadays probably refer to as psychology) and knowledge of people in the collective sense, as various groups (what we would nowadays probably refer to as anthropology). It is with this in mind that Kant says that “world history, biographies, even plays and novels” are “aids” to anthropology.¹⁵⁸ Interestingly, this anticipates Martha Nussbaum's point that the “mainstream realist novel [...] constructs compassion in readers, positioning them as people who care intensely about the sufferings

¹⁵⁶ 7:119.

¹⁵⁷ Ibid.

¹⁵⁸ 7:121.

and bad luck of others, and who identify with them in ways that show possibilities for themselves.”¹⁵⁹ Both philosophers believe that literature provides the possibility for emotional identification with and knowledge of individuals and groups, which contributes to the psychology and moral behavior of readers. As Kant makes the general point clear, both dimensions of pragmatic knowledge are to contribute to what a free human being can and should make of herself – what we would call moral character. Given that sympathy is central to the Kantian concept of moral character and its primary purpose is to serve a practical-epistemic function as I have argued above, this pragmatic knowledge is primarily a contribution to the Kantian virtue of sympathy.

This is clear through each section of Part II, where Kant discusses, in turn, the character of the person, sexes, peoples, races, and species. Among this group, his discussion of the character of the person and the species are the most nuanced and fine-grained, likely because these categories deal in a high degree of abstraction, being concerned with features of moral psychology relevant to understanding particular individuals and the human species, respectively. This means that Kant, the philosopher, is on familiar terrain; it also leaves less room for the kind of speculation that reflects the biases of one’s particular social and historical location. Such biases are indeed evident in the *AN* in particular, compared to Kant’s other work; however, Kant generally strives to overcome biases of which he might be aware in order to attain a nuanced and fair understanding of people different from himself.¹⁶⁰ Even

¹⁵⁹ See Martha Nussbaum, *Poetic Justice: The Literary Imagination and Public Life* (Boston: Beacon Press, 1995), 66. It is worth noting that though I am agreeing with much of what she says, I am disagreeing with her when she says that it is difficult for Kant to motivate beneficence consistently (*ibid.*, 65). See also, Martha Nussbaum, *Love’s Knowledge: Essays on Philosophy and Literature* (Oxford: Oxford University Press, 1990).

¹⁶⁰ For example, in *TPP*, we observe a respect for the dignity of others very different cultures that transcends such biases in his condemnation of the “horrifying” tendency of occupying people’s to count local inhabitants as “nothing”: “[i]n the East Indies (Hindustan), they brought in foreign soldiers under the pretext of merely proposing to set up trading posts, but with them oppression of the inhabitants, incitement of the various Indian states to widespread wars, famine, rebellions, treachery, and the whole litany of troubles that oppress the human

his use of physiognomy, a problematic endeavor nowadays understood as a pseudoscience, “is for the art of cultivating taste, and to be more precise not taste in things but in morals, manners, and customs, in order to promote human relations and knowledge of human beings generally”.¹⁶¹ At times, there is content to his physiognomy that suggests he is engaging in what we might consider to be the different and plausible endeavor of art criticism, as when he suggests that the facial features of Greek statues, cameos, and intaglios indicate certain aesthetic and cultural ideals.¹⁶² In any case, the primary purpose of physiognomy in Kant’s eyes is to enhance an agent’s understanding for pragmatic, moral purposes.

His remarks on the character of the sexes reveal the same purpose. He discusses to a limited extent the power differentials between men and women as rational beings, and how certain social pressures motivate certain forms of behavior. To acquire a sense of how he handles some of these issues, it is worth reading in full one of his observations.

In civil society the woman does not give herself up to the man’s desire without marriage, and indeed *monogamous* marriage. Where civilization has not yet ascended to feminine freedom in *gallantry* (where a woman openly has lovers other than her husband), the man punishes his wife if she threatens him with a rival. But when gallantry has become the fashion and jealousy ridiculous (as never fails to happen in a time of luxury), the feminine character reveals itself: by extending favors toward men, woman lays claim to freedom and, at the same time, to the conquest of the entire male sex. – This inclination, though it indeed stands in ill repute under the name of coquetry, is nevertheless not without a real basis of justification. For a young wife is always in danger of becoming a widow, and this causes her to extend her charms over all men whose fortunate circumstances make them marriageable; so that, should this situation occur, she would not be lacking in suitors.¹⁶³

It is noteworthy that he believes an inclination that stands in ill repute admits a real basis of justification. His observations, albeit anachronistic, are not without nuance and sympathy.

race” (8:358-359). Demonstrating a degree of cultural self-reflexivity, he says that although such powers “make much ado of their piety [...] they drink wrongfulness like water” (8:359).

¹⁶¹ 7:297.

¹⁶² Ibid.

¹⁶³ 7:304-305.

Kant's discussion of the character of peoples, by which he means "the *number* of beings united in a region, insofar as they constitute a whole", proceeds in a similar fashion.¹⁶⁴ He embarks on a description of the nature of various European nations with the point of providing practical knowledge and encouraging the disposition of sympathy as defined above, for the benefit of all parties who might encounter and engage with one another. He says:

In an anthropology from a pragmatic point of view, however, the only thing that matters to us is to present the character of both [a present people and their ancestors], as they are now, in some examples, and, as far as possible, systematically; which makes it possible to judge what each can expect from the other and how each could use the other to its own advantage.¹⁶⁵

To describe a situation involving two parties in which each uses the other to its own advantage might resonate as egoistic, or mercenary. However, it is worth noting that knowledge, or "prudence in virtue of which one is able to use all human beings for one's own final purposes", is constitutive of the process that Kant refers to in the *LP* as the "civilizing" of human beings.¹⁶⁶ As such, it is a precondition for moralization, which includes the choice of "nothing but good ends [...] which can be the simultaneous ends of everyone".¹⁶⁷ We also observe that Kant is describing very large groups of people (nations) with the aim of achieving a situation in which both groups flourish, since this is how the greatest amount of happiness may be achieved. Knowledge of the expectations and desires of the other people groups an individual will come in contact with, in addition to knowledge of one's own group, will help an individual to discern how she ought to act toward that other group. It will help

¹⁶⁴ 7:311.

¹⁶⁵ 7:312.

¹⁶⁶ 9:450. This is also referred to in 7:324 of the *AN*.

¹⁶⁷ *Ibid*.

her to understand what her duties toward them are and will therefore contribute to the end of serving their happiness, in addition to her own.

Kant's discussion of the character of the races is extremely brief and offers little that is of interest. His discussion of the character of the species, however, is detailed and interesting. Among other things, it situates his moral psychology in a greater, politically ambitious framework by asserting that there is a cosmopolitan dimension to pragmatic anthropology and hence to the virtue of sympathy, a theme otherwise implicit in the rest of the *AN*. Unlike in the previous sections of Part II, his aim here is to provide knowledge of how human beings actually function in society and what sort of society it is possible for them to achieve. He indicates that the capacity for freedom and self-perfection is what distinguishes human beings from the rest of animate nature: "in order to assign the human being his class in the system of animate nature, nothing remains for us than to say that he has a character, which he himself creates, insofar as he is capable of perfecting himself according to ends that he himself adopts."¹⁶⁸ He goes on to explain that although humankind naturally finds itself in a state of discord, reason impels it to preserve, educate, and govern itself on an individual and species-wide level.¹⁶⁹ The goal of this process is to bring about concord and "the perfection of the human being through progressive culture, although with some sacrifice of his pleasures of life."¹⁷⁰ He later continues:

The sum total of pragmatic anthropology, in respect to the vocation of the human being and the characteristic of his formation, is the following. The human being is destined by his reason to live in a society with human beings and in it to *cultivate* himself, to *civilize* himself, and to *moralize* himself by means of the arts and sciences. [...] The human being must therefore be *educated* to the good.¹⁷¹

¹⁶⁸ 7:321

¹⁶⁹ 7:321-322.

¹⁷⁰ 7:322.

¹⁷¹ 7:234-235.

Since civilization, as it is meant here, implies an understanding of the behavior, manners, and culture of human society, Kant may be understood as urging that people develop knowledge of these things in order to approximate the cosmopolitan ideal he has in mind.¹⁷² He later says that “with the advance of culture [people] feel ever more strongly the ill which they selfishly inflict on one another” and as a result are willing to subjugate individual interest to the interest of all.¹⁷³ An important effect of culture ought to be an increased comprehension of the suffering of others and an understanding of one’s own responsibility for this. A desire to restrain one’s self-interest for the sake of others – for the sake of fulfilling one’s duty toward them – is the appropriate and morally necessary response to this comprehension. In this way, Kant maintains that the knowledge implied in a sympathetic disposition is crucial to the successful pursuit of the cosmopolitan political ideal. The importance of acquiring this knowledge underwrites his project of pragmatic anthropology.

The Change of Heart and the Puzzle of Dual Character

I have suggested that Kant’s account of sympathy connects with a broader question concerning the nature of character in his ethics, in particular, concerning how an agent can be understood to have a good will, or good character in the noumenal sense, and still have bad character in the empirical sense. This puzzle further connects to Kant’s common identification of virtue with strength, which leaves the impression that virtue is only present when there is some contrary inclination for it to overthrow, an impression deepened by his

¹⁷² See what the *LP* says on civilization at 9:450 and as discussed above.

¹⁷³ 7:329.

account of moral worth from *GR I*. I will discuss Kant's idea of moral worth in the next chapter, but I would first like to suggest here that one way bad empirical character can remain after the change of heart is due to an agent's not having developed an adequate disposition of sympathy, or particular knowledge of herself and others. Given that on the Kantian conception, sympathy is primarily a matter of practical knowledge, an illuminating analogy may be drawn with one account of weakness of will, in which failure to act morally results from a kind of psychological impossibility or a certain lack of knowledge. Though for Kant, the problem is not of weakness in a good will but rather lack of thoroughgoing knowledge. I will explain how this analogy is illuminating. The general idea is that full moral character in the Kantian sense is constituted by a good will in one's noumenal character and well-developed sympathy and knowledge of self in one's empirical character. I will argue that Kant has a double conception of character according to which one acquires complete character at a later point than one has made the moral law one's fundamental motivation (*Gesinnung*), and that a proper understanding of this conception of character renders this puzzle much less troubling.

In the *REL*, Kant distinguishes between virtue that belongs to one's empirical character and virtue that belongs to one's intelligible character. The virtue of empirical character "has the abiding maxim of *lawful* actions, no matter whence one draws the incentives that the power of choice needs for such actions."¹⁷⁴ It is acquired incrementally through a gradual process of habituation. It results in a human agent having passed from "a propensity to vice to its opposite" and having developed a sense of stability in the maintenance of dutiful actions.¹⁷⁵ However, it does not imply that one's fundamental maxim

¹⁷⁴ 6:47.

¹⁷⁵ Ibid.

is moral: as Kant says, it does not imply a change of heart from bad to good but only a change of *mores* (*Sitten*).¹⁷⁶ As Kant points out, a liar might begin telling the truth for the sake of reputation and it does not seem natural to say such an individual has good character, even if they consistently tell the truth.¹⁷⁷ It is thus possible that one's external actions conform to duty while one's fundamental maxim or disposition remains evil (i.e., one makes self-love the condition for moral behavior). The self-interestedly honest person may be empirically virtuous because empirical character only picks out the external form of action, or its legality as Kant often says, and not the agent's deepest motivation. Virtue in one's empirical character does not entail virtue in one's intelligible character; this would require that one's fundamental maxim be moral/the moral law.¹⁷⁸

It is easy enough to understand how an individual may have virtue in the external or empirical sense and yet lack virtue in the internal or intelligible sense (i.e., not be motivated to do the virtuous thing because it is virtuous). Examples like Kant's self-interested honest man are familiar to us. Yet the dual nature of character found in this text allows the order of virtue to go in the other direction, too, and this is more difficult to understand, if it is correct. Once an individual undergoes the instantaneous change of heart, the transformation of her empirical character takes time: "a revolution is necessary in the mode of thought but a gradual reformation in the mode of sense [...] he is to this extent, by principle and attitude of mind, a subject receptive to the good; but he is a good human being only in incessant

¹⁷⁶ Ibid.

¹⁷⁷ Ibid.

¹⁷⁸ Kant holds that such a change in intelligible *Charakter*, and hence in one's fundamental maxim, cannot happen gradually as it may happen in one's empirical *Charakter*. This has to do with his idea of rigorism, which forbids that maxims be mixed and so requires that one's fundamental maxim be either good or evil (where the 'or' is taken in the exclusive sense). So there is no way for such a change to happen except suddenly, as if by a kind of "revolution" or "single and unalterable decision" (6:47-48). The result of this change of heart is that one "should become not merely *legally* good, but *morally* good (pleasing to God) i.e. virtuous according to the intelligible *Charakter*" where "for God, this is the same as actually being a good human being" (6:48).

laboring and becoming”.¹⁷⁹ This raises the question of how one who has undergone the change of heart and whose fundamental motivation is therefore good can still not get things right and so must engage in this process of continual self-improvement. Or, how can a person with good will have what appears to be bad character?

Because it operates at the level of an agent’s intelligible character, the change of heart involves the agent’s fundamental maxim, which describes her character in the deepest sense. Since it involves a maxim and so an underlying intention, the change of heart must involve the agent’s capacity for choice. As Kant says, an agent’s good or evil character must be “an effect of his free power of choice” otherwise her character could not be imputed to her.¹⁸⁰ In line with Kant’s earlier remarks in this chapter, the change of heart involves the movement from a fundamentally evil maxim of self-love to the good maxim of the moral law.¹⁸¹ The change of heart is a choice for the good, a choice to make the moral law the condition of one’s action in the sense that all other reasons for action, including self-love, are subordinate to it.¹⁸² When it is moral, Kant describes this fundamental maxim as a good character [*Charakter*] acquired by choice.¹⁸³ Elsewhere he describes this fundamental maxim or “first ground”, whether good or evil, as that by which the human being expresses the character [*Charakter*] of her species.¹⁸⁴ This fundamental maxim, also understood as intelligible character, represents the first, primary element in Kant’s double conception of moral

¹⁷⁹ 6:48.

¹⁸⁰ 6:44.

¹⁸¹ See his discussion of the disjunctive proposition and ‘rigorism’ in 6:22 and mentioned in note 178 above. Whether one is evil or good is then a matter of which maxim, self-love or the moral law, is made the condition of the other, as he says in 6:36.

¹⁸² 6:36.

¹⁸³ “But now this is possible only because the free power of choice incorporates moral feeling into its maxim: so a power of choice so constituted is a good character [*Charakter*], and this character [*Charakter*], as in general every character of the free power of choice, is something that can only be acquired” (6:27).

¹⁸⁴ 6:21.

character. It can be evaluated as good or bad, requires choice on the part of the agent, and is imputable to the agent.

The secondary element in Kant's double conception of moral character is, in this passage, introduced alongside the idea of empirical character. Once the change of heart has occurred, that one's fundamental maxim has been converted to the good (the moral law) entails that the individual must embark on a journey of moral self-formation to become an agent that is thoroughly good. As Kant says, "by principle and attitude of mind, [the change of heart makes him] a subject receptive to the good; but he is a good human being only in incessant laboring and becoming".¹⁸⁵ The change of heart may happen instantaneously, at one point in time; however, to become a good human being (*ein guter Mensch*) involves continual work over a period of time, indicating that one is not yet an entirely good human being after one has undergone the change of heart.

This indicates that good character is not complete until features beyond an agent's fundamental motivation (*Gesinnung*) have been shaped to align with this motivation, to enact and enhance its activity. Kant describes this formation in terms of moral education, which must begin not with an individual's improvement in *mores* (*Sitten*), but with the "transformation of his attitude of mind [*Denkungsart*] and the establishment of a character [*Charakter*]"¹⁸⁶ He goes on to say that a predisposition to the good "gradually becomes an attitude of mind [*Denkungsart*]"¹⁸⁷ This language resonates with his account of moral aptitude from the *MM*, discussed in chapter one.¹⁸⁸ The education he speaks of clearly takes place after the change of heart, so it might appear puzzling that he continues to speak of the

¹⁸⁵ 6:48.

¹⁸⁶ Ibid.

¹⁸⁷ Ibid.

¹⁸⁸ Recall that moral aptitude is "a facility for acting and a subjective perfection of *choice*" in which one "determines oneself to act through the thought of the law" (6:407).

establishment of a character. A plausible response to this is that here he is speaking of the establishment of empirical character, or character in its secondary sense, as I have suggested, which is something primarily cognitive. The transformation of attitude of mind, or *Denkungsart*, is a part of this establishment of character: indeed, a moral *Denkungsart* appears to constitute the establishment of character in this secondary sense. In chapter one I argued that Kant grasps four distinct types of habit. *Denkungsart*, the word he uses several times in this passage from the *REL*, falls under the category of aptitude – the same category to which *Fertigkeit* belongs. There are two types of aptitude, mere aptitude and moral aptitude, and as we saw in the character of Sulla, aptitude that is not grounded in the moral law might be grounded in evil character and so pernicious to morality. It makes sense to hold, as Kant does, that after a change of heart it takes time for an agent with a new predisposition to the good to obtain a firm moral aptitude or attitude of mind (*Denkungsart*). It makes sense to hold that this takes time at least partly because aptitudes or attitudes of mind acquired as the result of a previous, evil disposition (of self-love) must be appropriately reworked in order to adequately serve a new, good disposition. Aptitudes or attitudes of mind that linger in this way are essentially habits in the simpler, unthinking sense, because the agent is not aware of them. As any therapist would say, the reworking of one's patterns of thought is a process and therefore takes time.

These patterns of thought will not all be transparent to the agent at once and it will take time for her to discover them so that she can change them; the agent will in this way have to engage in a potentially lengthy process of cognitive rehabilitation. Throughout this process she might consciously choose to act on the moral law, yet unthinkingly act against what it requires due to the lingering presence of false beliefs that stem from a previous

disposition of self-love. As a variation on Kant's familiar example, we might imagine a wealthy philanthropist who regularly gives money to charities for self-interested reasons. Perhaps these acts give her a sense of worth precisely because they underscore her superior wealth and social status in contrast to that of her beneficiaries. Perhaps this philanthropist has been a generally proud and unkind person for some time, and her acts of giving are accompanied by remarks that are condescending, hurtful, and contrary to the dignity of those who benefit from her donations. She might one day realize these acts of giving are motivated by self-love and resolve to act for the sake of the moral law instead, thus undergoing what Kant would call a change of heart. She might then perform the same acts of giving for moral reasons and yet, without knowing this were contrary to their dignity, still refer to those who benefit from her donations in a condescending manner. She might be so used to speaking that way that she does not recognize it as an instance of moral failure; perhaps for a long time she has been surrounded by others who speak that way. She seems to be guilty of a moral failure, but a mitigated one, since it is a failure of knowledge and not motivation (she has not formed the belief that her remarks are harmful and is unaware that they register morally at all). On Kant's account, this would be a failure of moral character.

So the secondary element in Kant's double conception of moral character is represented by a moral attitude of mind (*Denkungsart* or *Fertigkeit*) that is the result of deliberate and continual work on the part of the agent. This secondary element of character is primarily cognitive. This is supported by the contrast Kant draws between non-cognitive habits (including unthinking habit and external moral habit) and cognitive aptitude (including mere aptitude and moral aptitude) outlined in chapter one, that he holds only moral aptitude qualifies an agent for ascriptions of moral worth, and his emphasis on facility in judgment as

a feature of moral aptitude. It is also supported by his definition in the *MM* of moral aptitude in contrast to mere habit.¹⁸⁹ It could further be added that one's empirical character can be brought in line with the predisposition (motivation) to the good by the exercise of judgment and the formation of an attitude of mind in ways similar to those meant to encourage the grounding of character, from chapter one. This deflates this apparent inconsistency and lines up with what he says elsewhere. This attitude of mind is a capacity and facility for moral judgment and action on the basis of that judgment; it requires choice and is imputable to the agent. Kant says an agent cannot be a "good human being" without it, so it would seem necessary for completely good, or complete moral character. If an agent were then to examine her own maxims in the same way Kant's apprentices in morality are to do, and that agent had undergone the change of heart, it would have to be said that she is looking out for habits, beliefs, or attitudes of mind that are the result of a previous disposition of self-love as described in the example of the wealthy philanthropist above.¹⁹⁰ This is because it could not be said that after the change of heart she is acting on the moral law and yet also acting on a maxim of self-love; so these sorts of acts would have to be unthinking and not involve maxims. Though Kant does not state here that this attitude of mind is to be identified with the virtue of sympathy, sympathy as I have described it above satisfies these criteria. If it is correct that this virtue is central to his conception of moral character, it would make sense that sympathy constitutes a necessary condition for full moral character.

Thus far is Kant's double conception of complete (good) moral character, as comprised of a fundamental intention (maxim) to enact the moral law and sympathy as practical knowledge, or a cognitive ability to understand what the moral law requires

¹⁸⁹ 6:407.

¹⁹⁰ For Kant's apprentices in morality, cf. chapter one and 6:48.

(involving a moral aptitude, attitude of mind, or *Denkungsart*). To return to the original puzzle, how can an agent be understood to have undergone the change of heart and yet be required to engage in a process of continual self-improvement, when this would imply that not all of her inclinations are in favor of acting for the sake of the moral law? The way of easing this puzzle involves understanding what good character in the secondary sense does and does not entail. Kant appears to be saying that good character on the whole cannot be had without a sympathetic disposition, which involves a cognitive skill or facility in discovering the corrupted state of one's attitude of mind and then, where that corruption lies, correcting it to line up with the good disposition. If this is the case, it is possible for an agent to have a fundamental motivation or intention to enact the moral law and consistently fulfill her moral duties in cases where she knows it is her duty to do so, yet due to previous habituation in the contrary direction, remain unaware of other cases in which she ought to do so, or unknowingly act against the moral law, and therefore fail to fulfill her moral duties in those cases. The agent's moral failure in such cases would: first, be an objective moral failure, for example, from the perspective of the moral law and an impartial observer, second, indicate a flawed or incomplete character on the part of the agent, and yet, third, could coexist with a knowledge of the moral law and the intention to act on it on the part of the agent. The final, third point is possible because moral action requires empirical knowledge and proper cognitive habituation, as it were – knowledge of what one's duty in a given situation is – in addition to general knowledge of the moral law and the intention to act on it. An agent in such a case of moral failure would not have acquired the moral attitude of mind or particular knowledge necessary to enable her to enact her fundamentally good intention, and so would fail to enact her good intention.

There is an analogy between this sort of case and one account of weakness of will. In both, the agent under consideration appears to have knowledge of what is good and yet fails to act on that knowledge in cases where they ought. Aristotle is held to have in some places explained weakness of will as an instance in which desire on one side, defeats reasoning on the other, by preventing that reasoning from coming to an agent's attention.¹⁹¹ If this were a straightforward matter of seeing two possible courses of action, one motivated by the moral law and the other by desire contrary to it, and then choosing to follow the latter, this would not bode well for Kant. It would suggest that one can have good moral character in the primary sense – a good will – and yet that good will can be defeated by contrary desires. This would seem absurd, since it would then not seem right to say the agent has a good will in the first place. However, if one can be consistently motivated by the good and yet fail to act accordingly due to a lack of knowledge of what the good requires, this concern would not arise because the agent would remain motivated by the good (not wayward desire) but incapable of following through on that motivation for intellectual reasons. This is in line with another account of weakness of will provided by R.M. Hare. Accordingly, the existence of such cases would render Kant's picture of character increasingly plausible. This is what a comparison to Hare account bears out.

According to Hare, the failure of agents to do what they say they ought typically results from either impossibility to do so (whether physical or psychological), or a use of 'ought' in an insincere or off-color way.¹⁹² If an agent intentionally does the latter, it is safe

¹⁹¹ See book 7 of Aristotle's *NE*, as well as David Bostock, *Aristotle's Ethics*, 133. Similarly, David Charles says Aristotle explains weakness of will as resulting from either (i) the non-activation of a minor premise of practical reasoning, or (ii) the possession of a conclusion but in an off-color way (cf. analogy to drunk people and students), in David Charles, *Aristotle's Philosophy of Action* (London: Duckworth, 1984), 127. See Aristotle, *NE*, VII.3 for this discussion.

¹⁹² For a summary of his position, see: R.M Hare, *Freedom and Reason* (Oxford: OUP, 1963), 82-83. As in note 191 above, Charles observes similar features in Aristotle's account.

to say her will is not sufficiently good since an insincere and intentional utterance along these lines belies that she does not mean ‘ought’ in the proper way. Hare is probably right on this point but it is not of obvious relevance to Kant. The prospect of impossibility, however, is helpful to the Kantian picture in at least two ways. First, Hare’s discussion of divided personality provides a way in which an agent may fail to do what she ought though her will remains fundamentally good. The explanation he provides is that one part of the agent’s personality issues a moral command to the other while the other is unable to obey this command because of a “recalcitrant lower nature”.¹⁹³ This explanation has the effect of retaining the strength of the command as well as the agent’s ability to endorse it in her ‘higher’ self, even though her ‘lower’ self is incapable of obeying it. Such an agent may have a good will or character in Kant’s primary sense of a fundamental intention to do the good, and yet not have good character in the secondary, complete sense because her recalcitrant lower nature makes this impossible for her, at least for the time being.¹⁹⁴ If Hare’s account is correct and such agents exist (and action in accordance with the good is impossible for them), it provides us with one explanation that may help to deflate the puzzle surrounding Kant’s dual conception of character. Though Kant believes this sort of change in character is possible, he acknowledges that it takes time and requires certain knowledge of self to proceed, so in this sense it may not be possible for an agent to alter this recalcitrant nature all at once. This explanation also does not effectively handle the fact that mastery of oneself and the acquisition of character in the secondary sense is primarily a cognitive task that concerns an agent’s knowledge. This is better handled in a further point concerning the notion of impossibility.

¹⁹³ Ibid., 81.

¹⁹⁴ It should be noted that Kant does not have room for the idea of a recalcitrant lower nature except insofar as it is not maxim governed and therefore not under an agent’s control.

This second point is that it may be that the agent of good character in the primary sense, whose fundamental intention is to do the good, is unaware of further patterns of thought that continue to linger on in her mind and self and direct her to not do the good, after she has undergone the change of heart. Kant's example of allowing "apprentices in morality to judge the impurity of certain maxims on the basis of incentives actually behind their actions" (discussed above) suggests something just like this.¹⁹⁵ It suggests that an agent must seek out and expunge beliefs, patterns of thought, and features of character that result from prior habituation in self-love, and that one who is inexperienced in morality cannot immediately detect these. Nevertheless, an agent can learn how to detect them and be successfully cured. Learning to detect such features of character and root them out is part of the task on which the agent who has undergone the change of heart is embarked. As such, it is concomitant with the Kantian virtue of sympathy that is also primarily cognitive in nature, since it involves the apprehension of facts concerning the well-being of others and one's duty toward them, or how to apply the moral law in particular circumstances. Knowledge of these features does not come immediately with the change of heart – it must be acquired through time – and adequate possession of this knowledge is necessary to make one's good character complete. Until knowledge of these features is acquired, the agent of good intention is ignorant of their existence and it is therefore impossible for her to eradicate them. In this way, it is possible that Kant's agent may have a good will, or good character in the primary sense, and not yet have good character in the secondary, complete sense.

This does not quite explain how on Kant's account, an agent might have knowledge of the good (moral law) and yet not, as it were, activate that knowledge in a particular case. It would seem there are such cases in which desire for some course of action that is contrary to

¹⁹⁵ 6:48.

the moral law distorts an agent's belief that that course of action in fact is contrary to the moral law (for example, the belief that one's partner will be happier if she doesn't know the truth about a gambling addiction, so it is better to fabricate reasons for lost income). It is difficult to imagine what Kant would say to this. For although this agent can in some sense be said to be acting in ignorance of the good, it seems she is culpable for this ignorance, which may make it more difficult to believe she has undergone a genuine change of heart. Though Kant may not provide a full enough explanation of how this might happen after the change of heart, he is attentive to the practical worry related to it and cautions that we cannot know when the change of heart actually takes place.

Assurance of this cannot of course be attained by the human being naturally, neither via immediate consciousness nor via the evidence of the life he has hitherto led, for the depths of his own heart (the subjective first ground of his maxims) are to him inscrutable. Yet he must be able to *hope* that, by the exertion of *his own* power, he will attain to the road that leads in that direction, as indicated to him by a fundamentally improved disposition.¹⁹⁶

There is never a reason to not be vigilant concerning the moral status of one's maxims, and the change of heart is something we hope in and strive to make evident by our effort toward moral action. In *EA*, Kant emphasizes that this effort is unending as far as an earthly life is concerned, saying that moral life even at its best "will always remain an ill compared with a better one", so the goal of moral satisfaction can only be posited beyond time and understanding.¹⁹⁷ In short, Kant is alert to worries around self-deception. He elsewhere warns

¹⁹⁶ 6:51.

¹⁹⁷ "Even assuming a person's moral-physical state here in life at its best – namely as a constant progression and approach to the highest good (marked out for him as a goal) – he still (even with a consciousness of the unalterability of his disposition) cannot combine it with the prospect of *satisfaction* in an eternally enduring alteration of his state (the moral as well as the physical). For the state in which he now is will always remain an ill compared to a better one which he always stands ready to enter; and the representation of an infinite progression toward the final end is nevertheless at the same time a prospect on an infinite series of ills which, even though they may be outweighed by a greater good, do not allow for the possibility of contentment; for he can think that only by supposing that the *final end* will at sometime be *attained*" (8:335). And, "this is a concept in which the understanding is simultaneously exhausted and all thinking itself has an end" (8:336).

that “one is never more easily deceived than in what promotes a good opinion of oneself.”¹⁹⁸ Perhaps he has this sort of thing in mind when he identifies virtue with fortitude and the “resolve to withstand a strong but unjust opponent.”¹⁹⁹ In any case, although it could remain conceptually possible that an agent may have a good will and yet allow an illicit desire to bend her perception of the moral law in some circumstances, Kant exhibits that he is both aware of and guarded against this possibility.

Yet it might also be that the change of heart can be episodic or involve a change for the worse and a reversion to evil character. Throughout the *REL*, Kant talks of the change of heart in terms of a conversion to the good, since he is generally trying to draw a parallel with the religious notion of conversion. However, there is nothing to prevent such a change from going in the opposite direction, and the fact that the change of heart occurs outside of time and is inscrutable makes it impossible to know. So it remains possible that an agent that has undergone the change of heart for the good may undergo a change of heart back to her previous, evil self, driven by self-love. This is one way Kant might be rescued from this puzzle, though perhaps too easily to be satisfying.

Nevertheless, I hope to have shown another way in which this puzzle concerning the change of heart can be eased and perhaps resolved. It may seem a bit strange, *prima facie*, that an agent can have a good will while not yet possessing good character in the complete sense. Though Kant’s agent begins from a position of disadvantage insofar as her mind and character are warped, bent away from the good, even though her desire and intention are to correct this. The change of heart entails an imperative to engage in this thoroughgoing work of correction – it signals the commencement of this project. That it takes time to get things

¹⁹⁸ 6:68.

¹⁹⁹ 6:380.

right and reform one's character completely should not be a surprise. Given the argument I have provided, there is no cause for worry at this juncture of Kant's conception of moral character.

3. Moral Worth

*Silver has no color while hidden
in the greedy earth. You are no lover of metal,
Crispus Sallustius, unless it shines, in tempered use.*

~ Horace²⁰⁰

*Love and desire are the soul's wings
to great deeds.*

~ Goethe²⁰¹

So far I have argued that Kant has a lot to say about moral character in the standard, broadly Aristotelian sense. In chapter one, I argued that inclinations contribute positively to the grounding of character and that Kant recognizes four distinct types of habit, two of which have moral value. In chapter two, I argued that for Kant sympathy is a kind of intellectual virtue, analogous to *phronēsis*, that makes use of emotion to help us know what serves the well-being of others so that we can fulfill our duty toward them. Taken together, the evidence I have provided shows that Kant's moral psychology is richer and more expansive than many

²⁰⁰ Horace, *The Complete Odes and Epodes* trans. David West (Oxford: Oxford University Press, 2008), II.2. Manfred Kuehn points out that Horace was among Kant's favorite poets: "Kant continued to think highly of the ancients, reading them throughout his life. Seneca and perhaps somewhat surprisingly Lucretius and Horace remained his favorites." (Manfred Kuehn, *Kant: A Biography* (Cambridge: Cambridge University Press, 2001), 49.) It may have been that Kant had this stanza in mind when he describes the good will as a "jewel [...] that] shine[s] by itself, as something that has its full worth in itself" (4:394), since he frequently read Horace and this particular ode is in praise of virtue, with the description of metal shining in use meant as an analogy to virtuous character in obvious action.

²⁰¹ Goethe, *Iphigenia auf Tauris*, Act II Scene 1, translation mine. Project Gutenberg: <http://www.gutenberg.org/cache/epub/2054/pg2054-images.html>.

have believed.²⁰² However, as I have mentioned, his account of moral worth appears to stand in tension with these more attractive features of his moral psychology. For example, Kant says in the *GR* that only when a man “wishes for death and yet preserves his life without loving it, not from inclination or fear but only from duty” does his maxim have “moral content.”²⁰³ Passages like this are striking and leave a lingering impression on the reader. It is sensible to say this is a case where the agent acted out of duty and against his prevailing inclinations; but the impression it leaves is that the agent’s action has moral worth precisely because it was performed out of duty and *against* his prevailing inclinations. If this were a condition on morally worthy action, it would undermine the account of moral character I have outlined so far: since action from moral character could not have moral worth, it would be confused to call this character moral. Further, this seems like a gross failure of personal integration – since it presents a deep incoherence between rational choice and personal desire – and it suggests a strange picture of the ideal moral agent as one for whom the emotions woven into everyday life are a hindrance rather than an aid, let alone somehow valuable. After all, we cannot wish away this part of our nature even if we wanted to, and it seems undesirable to do this anyway. In summary, this account of moral worth presents Kant with two problems: one to do with inconsistency and the other to do with implausibility.

In this chapter, I will argue that contrary to the above interpretation, moral worth may be ascribed to actions that proceed from moral character in the sense I have outlined so far. To do this, I will take a close look at Kant’s most direct discussion of moral worth in *GR* I, to establish exactly what this concept does and does not entail. The most important feature of action with moral worth is that it is performed on the basis of the right sort of maxim, namely

²⁰² For example, those critics mentioned in the introduction: Schiller, Hegel, Schopenhauer, Stocker, Wolf, and Williams.

²⁰³ 4:398.

the moral law, so I will also look at what a maxim is and what it means to act on the basis of one. However, before I do this I will examine Kant's broader understanding of character as indicated by his use of the term *Charakter* throughout the work of the Critical Period. This will take me beyond the *GR* and to some of Kant's later work, most notably the *REL*, where he says that an agent's maxim is ultimately what determines her character. Though it will not tell us everything we need to know about Kant's notion of moral character, this will be instructive. For the way he uses the term *Charakter* consistently points to the crucial role that an agent's maxim plays in determining the sort of person she fundamentally is – her character – and ultimately her moral worth. It therefore suggests it is reasonable to understand Kant's notion of moral worth against the backdrop of his theory of moral character. Since maxims are principles on which an agent acts, to have a maxim as well as to act on one requires a capacity for understanding and choice on the part of the agent. This means that on the Kantian conception, even prior to considering the virtue of sympathy, moral character necessarily has a cognitive or intellectual dimension. I will explain what a maxim is and how it features in cases of morally worthy action on the Kantian conception.

I will then unify these various features of his moral psychology in a general account of moral worth attributable to Kant. On this account, moral actions (actions from duty or the moral law) may involve participating inclinations (in other words, they may be overdetermined by inclination) and still have moral worth, so long as the agent's fundamental maxim is the moral law. This account allows that an action can have moral worth either if it would have been performed anyway, even if other motives had not been present, or if, although the agent did not have the moral law as a conscious motive, she acted out of considerations which the moral law would have constrained, if necessary. Others, including

most prominently Marcia Baron, Barbara Herman, and Nancy Sherman, have interpreted Kant along similar lines, saying he does not take the view that one must act on the moral law alone and without participating inclinations. However, his notion of moral worth has not been understood as a constitutive part of a broader, intellectualist theory of character in the way I argue it ought to be understood here. I will apply this account of moral worth to a hypothetical case concerning friendship, where I will suggest that Kant's position is not far from a common sense or broadly Aristotelian position. I will then bring this together with the account of character from chapter two and argue that full moral worth on the Kantian picture has a distinctive cognitive dimension that makes it a better theory than its main competitors.

The final definition of full moral worth I arrive at holds that unequivocal or full moral worth is a category of praiseworthiness attributed to actions that proceed from a morally good disposition and involve the exercise of well-developed sympathy, or good judgment in the practical application of moral principles. This definition encompasses the two points concerning participating inclinations (and the lack thereof) mentioned above. I will also outline a definition of what I call mitigated moral worth, reserved for agents who do not have ideal sympathy but are doing the best that can reasonably be expected of them.²⁰⁴ This will allow us to attribute a lesser degree of moral worth to agents who are committed to correct moral principles but do not possess proper sympathy or good judgment concerning the application of those principles, due to a failure for which they cannot be blamed.

²⁰⁴ In general, it seems wrong to attribute moral worth to an agent who has not acquired sympathy in Kant's sense of a cognitive facility for judging moral facts and situations correctly. Yet in some cases an agent may not be entirely to blame for her lack of good moral judgment and even her wrong moral beliefs. After all, it takes time to realize and change one's (false) beliefs and perhaps a given agent has made great progress when compared to her previous self. So sometimes an agent in this less than ideal condition will merit an attribution of moral worth, albeit a kind of mitigated moral worth, because this is the best that can reasonably be expected of her. This is what I will argue below.

Kant on *Charakter*

Kant uses the German term *Charakter* to describe the good will, which sets the context for his well-known discussion of moral worth in *GR I*.²⁰⁵ The *Duden* German dictionary defines *Charakter* in the first instance as the “distinctive quality of an individual person, as seen in their inherited and acquired qualities and expressed in their will and action”.²⁰⁶ This is very similar to the English term, the contemporary use of which is diffuse, but commonly describes the “sum of the moral and mental qualities which distinguish an individual”, or an individual’s “mental or moral constitution”.²⁰⁷ The English term ‘character’ has a wide semantic range and reflects multiple origins. It is partly a borrowing from French (*caractère*) and partly a borrowing from Latin (*character*), which is in turn a borrowing from Greek. The original Greek (*kharaktēr*) refers to a die or stamp that impresses a distinctive mark upon an object, such as a coin. This meaning was adopted into Latin metaphorically by Augustine and others to express the indelible quality that sacred rights like baptism impress upon the soul.²⁰⁸ Like the English, the current German definition retains this ancient idea of a distinctive mark or indelible impression, as shown by its use of *Gepräge*, which standardly refers to an imprint of this nature. In its metaphorical adaptation, this definition is meant to pick out the deepest and most distinctive features of a person and their mental or moral constitution: it aims to describe a person’s deepest and most enduring

²⁰⁵ “Understanding, wit, judgment and the like, whatever such *talents* of mind may be called, or courage, resolution, and perseverance in one’s many purposes, but they can also be extremely evil and harmful if the will which is to make use of these gifts of nature, and whose distinctive constitution [*Beschaffenheit*] is therefore called character (*Charakter*), is not good” (4:393).

²⁰⁶ In German: “*individuelles Gepräge eines Menschen durch ererbte u. erworbene Eigenschaften, wie es in seinem Wollen u. Handeln zum Ausdruck kommt*”. Translation mine. *Duden Deutsches Universalwörterbuch* (Mannheim: Dudenverlag, 2003).

²⁰⁷ “Character” in *Oxford English Dictionary* (<http://www.oed.com/viewdictionaryentry/showfullentry/true?isAdvanced=false&result=1&rskey=B5ca5n&t:ac=Entry/30639>).

²⁰⁸ *Ibid.*

qualities in a way similar to Augustine's reference to the enduring state of one's soul.²⁰⁹ A detailed definition of a person's character so understood might be formulated in response to a question asking 'what sort of person is A – what are her deepest motivations and how can she be counted on to act?' An answer to this question would consider the same sort of qualities that Kant's definition of moral worth would, since the latter requires that an action be performed because the moral law requires it, where that moral motive is sufficient to cause moral action and to prevent immoral action.

So, in his use of the term *Charakter* in the *GR*, Kant is drawing a contrast between a person's simple dispositions (talents, temperaments, and gifts of fortune) and the deeper constitution (*Beschaffenheit*) of the will that structures and directs those dispositions.²¹⁰ Kant uses this term (*Charakter*) to describe moral agency on a number of occasions throughout his practical philosophy, most markedly in the *REL*. Given its resonance with the English definition, it is helpful to examine his use of the term *Charakter* here to better understand its place in the Kantian system and whether it contributes to a coherent picture of moral character in something like the standard sense of chapters one and two. It turns out that for

²⁰⁹ As mentioned in the introduction, this idea of character as the deep features of a person overlaps with Julia Annas's general definition of virtue (Julia Annas, *Intelligent Virtue*, 9).

²¹⁰ *Beschaffenheit* is itself sometimes translated as character, but more frequently understood as composition, constitution, or nature. Further, the concept of *Charakter* is here bound up with the concept of the will, or *Wille* in German. Given that Kant has not yet fully developed the distinction between *Wille* and *Willkür* in the *GR* (he does this in his later work of the 1790s, namely the *REL* and *MM*), it is most reasonable to understand his use of *Wille* here in a broad sense that includes both the faculty of volition generally (which is the source of laws or imperatives) and the faculty or power of choice. (Henry Allison, *Kant's Theory of Freedom* (Cambridge: Cambridge University Press, 1990), 129-130.) Given that it bears the form of the moral law, in some contexts *Wille* can imply positive moral content. In combination with the lack of distinction in this text between *Wille* and *Willkür*, we might be led to believe that there is some ambiguity here as to whether he means to talk about the positive moral content provided by *Wille* or more simply, the power of choice. However, his example considers an agent with good dispositions, talents, and so forth, alongside an evil will: this rules out the idea that *Wille* refers here to only positive moral content and suggests that he is instead referring to a particular will that has become evil by choice. Such a will acts as a foil to the good will as the only unconditional good which shines "like a jewel" or like Horace's silver (4:393-394). In both positive and negative cases, Kant is describing the deep features of an agent's character, as defined by her fundamental maxim. This is his only use of the term *Charakter* in the *GR*. Though he does not use the German word in the positive case, the rest of the *GR* evinces that in both cases character implies choice and in the positive case, where one's *Charakter* is moral, that choice is constitutive of moral worth.

Kant, this term is most frequently used to describe an agent's fundamental and freely chosen moral disposition (at other times he refers to this disposition as a *Gesinnung*). This moral disposition indicates the distinctive constitution of an agent's will and so provides the ground for attributions of moral worth. What Kant says about *Charakter* corroborates and to some extent deepens what he says elsewhere about moral worth.

Charakter is used in the *REL* most frequently in Part One, in Kant's discussion of radical evil. He employs the term in his description of both intelligible and empirical character, as discussed in chapter two of this thesis. Recall that regarding empirical character, he says:

When the firm resolve to comply with one's duty has become a habit, it is called *virtue* also in a legal sense, in its *empirical character* [*Charakter*] (*virtus phaenomenon*). Virtue here has the abiding maxim of *lawful* actions, no matter whence one draws the incentives that the power of choice needs for such actions. Virtue, in this sense, is accordingly acquired *little by little*, and to some it means a long habituation (in the observance of the law), in virtue of which a human being, through gradual reformation of conduct and consolidation of his maxims, passes from a propensity to vice to its opposite. But not the slightest *change of heart* is necessary for this; only a change of *mores*. A human being here considers himself virtuous whenever he feels himself stable in his maxims of observance to duty – though not by virtue of the supreme ground of all maxims.²¹¹

We observe here that an agent has many maxims and that these maxims can be reformed and consolidated over time to form an outwardly virtuous character along the lines of external moral habit from chapter one. However, it is clear that virtue in empirical character alone is incomplete, since it does not include virtue in the supreme ground of one's maxims. Kant goes on to explain that a “*revolution* in the disposition [*Gesinnung*] of the human being” is necessary for an agent to become “not merely *legally* good, but *morally* good”, or “virtuous according to the intelligible character [*Charakter*] (*virtus noumenon*) and thus in need of no

²¹¹ 6:47.

other incentive to recognize a duty except the representation of duty itself”.²¹² This is his familiar notion of the change of heart. The way it is presented here reiterates that for an action to have moral worth, it must not merely conform to morality but be performed on the basis of the moral law. However, we here see that the maxim of duty may be one maxim among many, though for action to have moral worth, it must be the grounding maxim.

Earlier in the *REL*, Kant argues that an agent’s many maxims are ultimately grounded in only one maxim. Since there cannot be an infinite regression of maxims (e.g., reasons for adopting a particular maxim) and he accepts that personal freedom somehow obtains, Kant assumes an agent’s network of maxims forms a chain that bottoms out in a fundamental maxim or “first ground” (*ersten Grund*).²¹³ Kant further holds that this fundamental maxim or first ground of one’s action must be either the moral law or self-love. It is helpful to note that his “rigorism” (*Rigorismus*) – the belief that human nature and therefore maxims for action must be either good or evil and nothing in between – is a necessary premise here since this requires that one’s fundamental maxim be either good or evil.²¹⁴ Once we allow that only one of these two maxims may determine an agent’s incorporation of subsequent maxims and thus her action, Kant believes we need only accept that we sometimes reverse the moral order of our maxims by subordinating the moral law to self-love, to conclude that human nature is innately evil, or has a natural propensity to evil.²¹⁵

When he says the human being is by nature evil he means in the more general sense that evil applies to the human being considered as a species, not that it may be inferred from

²¹² Ibid.

²¹³ Ibid. He uses the expression *ersten Grund*, for example, on 6:21. His justification for freedom is complex and a full discussion of it is beyond the scope of this thesis.

²¹⁴ For discussion of this notion of rigorism, see 6:22 as well as notes 178 and 181 above. As stated immediately above, rigorism refers to the idea that an agent’s fundamental maxim is either good or evil; it does not refer to the idea (or rather, the accusation) that Kantian ethics allows only exceptionless maxims, which I will not discuss here.

²¹⁵ 6:29.

the concept of its species.²¹⁶ He suggests first that a propensity – a kind of inclination – toward evil is inherent in the human species, though in a way that the individual is still culpable for it.²¹⁷ Kant does not provide much in the way of an argument for the innateness of evil, perhaps because he believes we do not need much of an argument to be convinced of this point. He says we can spare ourselves the formal proof of innate evil by the “multitude of woeful examples that the experience of human *deeds* parades before us” and the “long melancholy litany of charges against humankind”.²¹⁸ What he provides instead is an account of how we might understand evil as innate, given that we already accept this as a fact, and so he provides a kind of backward-working explanation. The root of moral evil, he says, lies not in our dependence on the incentives of our sensuous nature, for he says these are innocent, as is the kind of self-love that this accords with. The root of moral evil instead lies in the reversal of the moral order of our incentives whereby we make the moral incentive subordinate to the incentive of self-love.²¹⁹ He has in mind here a pride or selfishness whereby the value of other rational beings is made subordinate to one’s self-interest and so carries no influence on one’s motivations. In such cases, actions may outwardly conform to morality and have good consequences etc. while being performed on the basis of reasons that are not moral. An adequate evaluation of character must include an evaluation of the agent’s reasons for action. This is precisely what Kant has in mind when he discusses moral worth in *GR I*.

²¹⁶ Ibid.

²¹⁷ Of this, he says: “to have one or the other disposition by nature as an innate characteristic does not mean here that the disposition has not been earned by the human being who harbors it, i.e. that he is not its author, but means rather that it has not been earned in time (that he has not been the one way or the other *always, from his youth on*). [...] This disposition too, however, must be adopted through the free power of choice, for otherwise it could not be imputed” (6:25). He goes on to say that there cannot be any cognition of this disposition: so he simply assumes that choice grounds this disposition, though how this happens is mysterious.

²¹⁸ 6:32-3.

²¹⁹ 6:36.

In this context, he employs the term *Charakter* to refer to the deep-rooted tendencies of human nature generally, which are revealed in this fundamental maxim, since by this first ground the human being “expresses at the same time the *Charakter* of his species.”²²⁰ Such *Charakter* is a feature that every agent possesses and the primary feature by which their actions are to be evaluated. Correspondingly, human nature is “the subjective ground wherever it may lie – of the exercise of the human being’s freedom in general (under objective moral laws) antecedent to every deed that falls within the scope of his senses”.²²¹ His point is that the evil found in human nature cannot be a product merely of inclinations, which we have not chosen, since we would only say a person is evil if they have chosen to be so. When we make such a judgment about a human being’s nature we are making reference to this first ground for the adoption of maxims.

An issue here is that we cannot often or reliably observe our maxims, so we may not say with certainty that a particular person (including oneself) is good or bad. As in the *CPR*, our maxims are ultimately grounded in our intelligible character, which it is beyond our ability to know with certainty. Kant seems to allow some capacity to judge a person’s maxims when he says that we call a human being evil not because his actions are evil but because his actions “allow the inference of evil maxims in him.”²²² However, this provides only a tentative sort of indication. It also does not assure us that we may infer good maxims on the basis of good actions (though it may seem more plausible to infer evil maxims on the basis of evil actions). Kant’s examples of morally worthy action in *GR I* are meant to at least help with this, but they do not guarantee insight into the fundamental ground of an agent’s

²²⁰ 6:21.

²²¹ 6:21.

²²² 6:20.

actions.²²³ Bearing this difficulty in mind, Kant warns us against actions that are in conformity with the law and appear good as far as an external observer might be able to discern, yet are performed for reasons to do with self-love.

In this reversal of incentives through a human being's maxim contrary to the moral order, actions can still turn out to be as much in conformity to the law as if they had originated from true principles [...]. For example, when adopted as a principle, truthfulness spares us the anxiety of maintaining consistency in our lies.²²⁴

In such cases, Kant says the “empirical *Charakter* is then good but the intelligible *Charakter* still evil.”²²⁵ His use of the term intelligible *Charakter* here is again meant to capture the agent's deepest motivation, or fundamental maxim – their moral disposition. In this case it is something that tends toward evil and corruption.

However, Kant also argues that human nature contains an original predisposition toward the good. A predisposition in this sense is a susceptibility to perform a certain kind of action. Kant outlines the predisposition toward the good in three stages that correspond to several aspects of the human person, namely animality, humanity, and personality.

The predisposition to good according to animality is manifest in self-preservation, the preservation of the species, and community with other human beings, all of which Kant takes to be a kind of self-love.²²⁶ Since Kant places these things under the category of a predisposition toward the *good*, he obviously does not understand them as bad, so self-love is not intrinsically bad or evil according to Kant. It only becomes evil when it is made the supreme maxim under which even the moral law is subsumed, as described above. Not unlike

²²³ Especially bearing in mind that, at one point in the *GR*, Kant wonders “whether any true virtue is to be found in the world” (4:407).

²²⁴ 6:36-7.

²²⁵ “[D]a dann der empirische *Charakter* gut, der intelligibele aber immer noch böse ist” (6:37).

²²⁶ 6:26. Note also Engstrom's discussion of rational self-love, mentioned in chapter two (cf. Stephen Engstrom, “The Concept of the Highest Good in Kant's Moral Theory,” 752-755).

the talents and temperaments of *GR I*, action on the basis of these predispositions ought to be constrained by the moral law so as to not bring about a lapse into vice. The particular vices to which the disposition of humanity tends are gluttony, lust, and lawlessness, which he terms vices of savagery.²²⁷

The predisposition toward the good according to humanity manifests itself in the comparison with others (which requires reason) by which one judges oneself to be happy. This too is a kind of self-love and not necessarily bad insofar as one strives only for *equal* worth in the eyes of others. The morally valuable type of shame discussed in chapter one makes use of this broad disposition, since “out of this self-love originates the inclination *to gain worth in the opinion of others*”.²²⁸ Despite its possible moral uses, this disposition too can give way to an unjust desire to acquire superiority over others and can engender the vices of envy, ingratitude, and *Schadenfreude*, which Kant terms vices of culture.²²⁹ So, in addition to the disposition to animality, it must be constrained by the moral law, to which the ideal moral agent is fundamentally committed.

The predisposition toward the good according to personality “is the susceptibility to respect for the moral law *as of itself a sufficient incentive to the power of choice*.”²³⁰ This susceptibility includes moral feeling insofar as moral feeling is caused by the rational concept of the moral law (for example, as spelled out in the *CPrR*) and provides an incentive for action on the basis of the moral law.²³¹ Kant further maintains that moral personality is

²²⁷ 6:27.

²²⁸ Ibid.

²²⁹ Ibid.

²³⁰ Ibid.

²³¹ In the *CPrR*, Kant says the moral law is a subjective determining ground, or incentive, “inasmuch as it has influence on the sensibility of the subject and effects a feeling conducive to the influence of the law upon the will.” He goes on to clarify that this moral feeling cannot be effected by sensibility, but is brought about when “the representation of moral law deprives self-love of its influence and self-conceit of its illusion, and thereby the hindrance to pure practical reason is lessened and the representation of the superiority of its objective law to

possible only in virtue of a free power of choice that can incorporate the moral law into its maxim. In this way, moral personality is the fullest expression of good human character. When it is achieved, it is the actualization of human freedom for the sake of the moral law, an ideal instantiation of rational choice. In this context, Kant employs the term *Charakter* to say: “a power of choice so constituted is a good *Charakter*, and this *Charakter*, as in general every *Charakter* of the free power of choice, is something that can only be acquired”.²³² This does not mean that *Charakter* is strictly identical to moral personality; it could not be since Kant uses the term elsewhere to describe features of the human being prone to evil. Instead, since *Charakter* is used to describe moral personality, it indicates that for Kant the term is reserved to describe the fundamental ground and enduring qualities of moral agency. At this level, he believes choice is operative, whether it leads one to acquire a good or evil character. So, for Kant, *Charakter* necessarily involves reference to an agent’s fundamental maxim and as such, is deeply associated with the exercise of individual will. This maxim is constitutive of an agent’s fundamental moral disposition, and it is this moral disposition that ultimately provides the ground for attributions of moral worth.

Moral Worth: the Locus Classicus

The *GR* contains Kant’s most detailed discussion of moral worth. At the outset of *GR* I, not long after introducing the good will as the only thing “that could be considered good without limitation”, Kant provides several examples of individuals who possess such a good will. In each of the examples, he emphasizes that an agent must act from duty (*aus Pflicht*)

the impulses of sensibility is produced and hence, by removal of the counterweight, the relative weightiness of the law (with regard to a will affected by impulses) in the judgment of reason” (5:75).

²³² 6:27.

for their action to have moral worth. The difficulty is that in each of these examples, the agent who is acting from duty and whose action therefore has moral worth does so either without additional inclinations that direct him towards the right action, or in the presence of inclinations that direct him against it. So, it looks as though for an action to have moral worth an agent cannot desire to perform the action other than for reasons that are strictly moral. In some cases, it looks as though the agent must desire not to perform the action for it to have moral worth. This makes the ideal moral agent out to be either some sort of rational machine or an emotional mess, neither of which are particularly desirable, or realistic.

It is worth looking at one of these examples in detail to understand why this interpretation might have been adopted and whether it is correct.

To be beneficent where one can is a duty, and besides there are many souls so sympathetically attuned that, without any other motive of vanity or self-interest they find an inner satisfaction in spreading joy around them and can take delight in the satisfaction of others so far as it is their own work. But I assert that in such a case an action of this kind, however it may conform with duty and however amiable it may be, has nevertheless no true moral worth but is on the same footing with other inclinations, for example, the inclination to honor, which, if it fortunately lights upon what is in fact in the common interest and in conformity with duty and hence honorable, deserves praise and encouragement but not esteem; for the maxim lacks moral content, namely that of doing such actions not from inclination but *from duty*. Suppose, then, that the mind of this philanthropist were overclouded by his own grief, which extinguished all sympathy with the fate of others, and that while he still had the means to benefit others in distress their troubles did not move him because he had enough to do with his own; and suppose that now, when no longer incited to it by any inclination, he nevertheless tears himself out of this deadly insensibility and does the action without any inclination, simply from duty; then the action first has its genuine moral worth.²³³

Given the contrast between the emotional state of the agent at earlier and later times, and the fact that only his later actions are said to have moral worth, it is not difficult to see how this passage might be taken to paint an unfavorable picture of the ideal moral agent. However,

²³³ 4:398.

several further things should be observed. First, Kant specifically aims to capture actions that have “true moral worth” and “moral content”, in contrast to actions that are “amiable” and further the good of others. For Kant, morality is broader than and meant to include more than just morally worthy action. Of course, the Doctrine of Right (which composes the first half of the *MM*) provides a systematic account of right action that is enforceable by others and is not performed for the sake of the moral law, but is performed in conformity with the moral law.²³⁴ For obvious reasons, this is an important category. However, the point here is that moral worth and not merely right action is under discussion in these examples. Second, moral worth is a matter of deserving “esteem” (praise) that goes beyond standard “praise and encouragement”. Action that benefits others and yet is performed on the basis of inclination may not have moral worth, yet it “deserves praise and encouragement”. So, actions without moral worth can still be praiseworthy and ought to be encouraged. This supports the idea that for Kant morality, and indeed morally praiseworthy action, is much broader than moral worth.

What sort of actions might be morally praiseworthy and yet not have moral worth in the strict sense? Actions that proceed from external moral habit, as defined in chapter one, are a good candidate. For they are not performed on the basis of the moral law, since they occur in the agent (for example, in a child) prior to the grounding of character and the decision to adopt the moral law as a fundamental maxim that this grounding of character entails. Nevertheless, external moral habit can help bring the agent closer to the point of making this decision, since it helps her to see that right actions follow a pattern, because they are based on a principle, namely the moral law. Further, actions from external moral habit

²³⁴ “Any action is *right* if it can coexist with everyone’s freedom in accordance with a universal law, or if on its maxim the freedom of choice of each can coexist with everyone’s freedom in accordance with a universal law” (6:230). Right action is a part of what Kant calls “external lawgiving” (6:229).

contribute to the happiness of others, which Kant clearly believes is a good thing because it is something we have a duty to do, as mentioned here in the *GR* and filled out in more detail in the *MM* (and as explained in chapter two).²³⁵ For both of these reasons, actions from external habit should be encouraged and indeed praised, even though they do not deserve the special kind of praise that actions from duty do and hence do not deserve an ascription of moral worth.

Third, we notice in the above example that actions that result from a certain kind of moral luck – something akin to what Thomas Nagel calls “constitutive luck” – in which an agent’s actions “fortunately light upon what is in fact in the common interest and in conformity with duty”, do not qualify for ascriptions of moral worth.²³⁶ As mentioned above, the moral content of an action comes from the maxim it is grounded in and morally worthy actions are ultimately grounded in duty or performed from duty. This feature of moral worth is apparent in Kant’s most focused discussion of the matter.

Thus the moral worth of an action does not lie in the effect expected from it and so too does not lie in any principle of action that needs to borrow its motive from this expected effect. For, all these effects (agreeableness of one’s condition, indeed even promotion of others’ happiness) could have been also brought about by other causes, so that there would have been no need, for this, of the will of a rational being, in which, however, the highest and unconditional good alone can be found. Hence nothing other than the *representation of the law* in itself, *which can of course occur only in a rational being*, insofar as it and not the hoped-for effect is the determining ground of the will, can constitute the preeminent good we call moral, which is already present in the person himself who acts in accordance with this representation and need not wait upon the effect of his action.²³⁷

²³⁵ It is reasonable to assume that a world in which people contributed to the happiness of others, though these contributions were not performed from a maxim of duty, would be better than a world in which people did not contribute to the happiness of others at all.

²³⁶ This includes luck in “the kind of person you are, where this is not just a question of what you deliberately do, but of your inclinations, capacities, and temperament.” See chapter three of Thomas Nagel, *Mortal Questions* (Cambridge: Cambridge University Press, 1979). Nagel also raises some challenges for Kant’s position, which I will not discuss here. The direct quotation is from Kant’s passage above.

²³⁷ 4:401.

This special praise of which morally worthy actions are deserving, can only be attributed to actions that are ultimately grounded in the moral law, and only rational beings are capable of so grounding their actions. So, for Kant moral worth necessarily involves two things: a rational agent and the moral law; and these two things must be related in a special way: the agent must not only know the moral law but act for the sake of the moral law. So far then, though we will need more detail to fill out what this means, we may define *moral worth as a special category of praiseworthiness attributed to an action performed by a rational agent because it is required by the moral law.*

This makes sense against the greater backdrop of Kant's project in the *GR*. It is also worth noting that the *GR* was written rather quickly and was meant to be a merely preparatory piece – a groundwork – for a more complete metaphysics of morals to be written later on.²³⁸ The task Kant sets for himself in the *GR* is to exhibit the supreme principle of morality in its purity, and so to establish a rational framework for morality.²³⁹ He begins by stating in the preface that rational cognition is divided into respective formal and material elements and it is best to handle any complex task, as in the trades, with a “division of labor [...] so as to be able to perform it most perfectly and with greater facility.”²⁴⁰ He asks rhetorically, “is it not thought to be of utmost necessity to work out for once a pure moral philosophy, completely cleansed of everything that may be only empirical and that belongs to anthropology?”²⁴¹ His answer to this is an obvious ‘yes’: he believes nothing empirical

²³⁸ Paul Guyer and Allen Wood say the *GR* “seems to have been almost a by-product of his plan to write a ‘metaphysics of morals’”. See their introduction to the *GR* in Immanuel Kant, *Practical Philosophy* eds. Paul Guyer and Allen Wood (Cambridge: Cambridge University Press, 1996), 39.

²³⁹ Ibid.

²⁴⁰ The formal element is logic, which has no empirical part and applies to all modes of cognition. The material element is concerned with some particular object and can be further divided into laws of nature and laws of freedom; rules of the former kind describe what happens and rules of the latter kind describe what ought to happen. (4:387.)

²⁴¹ 4:389.

should be allowed into a foundational framework for morality. He seems to take it as basic that the fundamental principle of morality is a priori in the sense that we ought to act on it, and that this is a convincing place to begin in an analysis of morality.²⁴² While still in the preface, he says: “[e]veryone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command ‘thou shalt not lie’ does not hold only for human beings, as if other rational beings did not have to heed it.”²⁴³ A key idea is that empirical rules cannot generate coherent ends because they are based on – or drawn up to satisfy – inclinations, and these inclinations often contradict one another and easily change. The inadequacy of empirical rules prompts Kant to look for something that delivers necessary and coherent commands amidst the instability of empirical conditions. He describes this ground in the will that is “good without limitation” and “corrects the whole principle of action and brings it into conformity with universal ends.”²⁴⁴ He then attempts to describe what the principle of this good will is. This leads him to an early description of the categorical imperative – “I ought never to act except in such a way that I could also will that my maxim should become a universal law” – before he establishes its various formulations in *GR* II.²⁴⁵

So, we see that the *GR* is engaged in the specific and limited project of establishing a rational foundation for ethics. Since this is the context in which Kant’s most well-known

²⁴² Kant does not provide a justification for the principle of the good will that lies at the basis of his account of duty because, much like Aristotle’s first principles (νοῦς) of practical deliberation, the categorical imperative that constitutes the principle of the good will is something grasped intuitively and taken as basic; it cannot be reached via discursive reasoning. For Aristotle, see *NE*, 1151a: “[I]n actions the end we act for is the principle, as the assumptions are the principles in mathematics. Reason does not teach the principles either in mathematics or in actions; [with actions] it is virtue, either natural or habituated, that teaches correct belief about the principle”. Regarding Aristotle, see also: Anthony Kenny, *Aristotle’s Theory of the Will* (London: Duckworth, 1979), 152. Regarding Kant, see 5:43, as well as: Ralph Walker, “The Rational Imperative: Kant Against Hume” in *Proceedings of the British Academy* LXXIV, 1988, 120-1.

²⁴³ 4:389.

²⁴⁴ 4:393.

²⁴⁵ 4:402, emphasis redacted. These formulations include the formula of universal law (4:421), humanity (4:429), autonomy (4:440), and the kingdom of ends (4:433), discussed in chapter two.

discussion of moral worth is found, it should be no surprise that that discussion reflects similar limitations when it examines what rational action looks like once empirical incentives have been stripped away. When Kant makes a similar argument about the foundations of morality and moral worth later on, in the *CPrR*, he compares the moral philosopher to a “chemist”, since “he can at any time set up an experiment with every human practical reason in order to distinguish the moral (pure) determining ground from the empirical”.²⁴⁶ So if the task is to ascertain the fundamental determining ground (maxim) of the ideal moral agent’s actions, it makes sense to remove other possible incentives to see whether this agent would still perform the right action. Then, if she performs the right action when she must act on the moral law and nothing else, we can be confident that she is so acting because the moral law requires it, so we can be confident that her action is grounded in the moral law and has moral worth.

This is precisely what Kant does in the earlier example of the philanthropist. In the first case, the philanthropist performs deeds of beneficence because of the “delight” he takes in “the satisfaction of others” so far as it is his own work.²⁴⁷ In the second case, when he is no longer capable of taking delight in the satisfaction of others (meaning there is no incentive other than the moral law for him to act beneficently) and he chooses to perform deeds of beneficence “from duty”, his action has moral worth for the first time.²⁴⁸ That this thought experiment has him experience no “sympathy with the fate of others” in the second instance simply serves to reveal the ground of his action; it does not imply that this kind of sympathy is incompatible with moral worth. The elimination of this feeling is a test designed to reveal

²⁴⁶ 5:92.

²⁴⁷ 4:398.

²⁴⁸ Ibid.

what the man's true character is.²⁴⁹ Further, the philanthropist's actions did not have moral worth in the prior instance because they were grounded in inclination and not the moral law; the problem was not that they involved inclination but that they depended on it. Kant does not necessarily exclude positive inclinations from all cases of morally worthy action. So we may add to the above, preliminary definition of moral worth that *an action that has moral worth would have been performed anyway, even if other motives had not been present to encourage the same action.*

When he introduces the idea of the good will at the outset of *GR I*, Kant extends this care to elucidate a distinctly rational ground for morality by disqualifying talents, temperament, and so-called "gifts of fortune" or constitutive luck from being good in themselves.²⁵⁰ The reason he cites for doing this is that "without the basic principles of a good will", these talents and temperaments "can become extremely evil".²⁵¹ In this way, their

²⁴⁹ Here it seems that Kant's thought experiment really is trying to rule out the influence of any feeling, including moral feeling, in motivating the philanthropist's action at this moment in time. He says this philanthropist's grief has "extinguished all sympathy with the fate of others" and he is no longer moved by "any inclination". Indeed, if someone is severely depressed, they may not have any meaningful feelings for the moral law to impinge on, thereby enabling them to experience the relative weightiness of the moral law (again, cf. *CPrR* 5:76 for an account of this moral feeling). Kant believes an agent must and can act on the moral law in even these conditions. For this reason, it does not quite fit to say, as Wood does, that the philanthropist's "tearing himself out of dead insensibility" would have to consist in his making himself actively susceptible to the feelings of love for those he helps that lie at the ground of his moral predisposition" (Wood, "The Final Form of Kant's Practical Philosophy," 19). It is not clear that there is "a *pleasure* taken in the rational representation of the dignity of the rational nature of the other, which *prompts* [the philanthropist] to treat the other as an end in itself" (*ibid.*, 18, italics mine). Perhaps there is a kind of rational desire at work here, but pleasure seems something different. Further, it is not clear exactly what Kant means when he says in the *MM* that love of human beings (*Menschenliebe*) must be presupposed in an agent for that agent to be receptive to the concept of duty: his remarks in this section are obscure and difficult to follow (6:401-402). For these reasons, I do not believe Wood is correct in saying, "the very possibility of our being under a duty to be beneficent to others presupposes that we have a predisposition to feel love for them" (Wood, "The Final Form of Kant's Practical Philosophy," 18). While an agent can be expected to attempt to cultivate certain feelings, she cannot strictly speaking be expected to force herself to have certain feelings; however, she can be expected to recognize the needs of others and act for the sake of them. For someone who is severely depressed, this alone would require a significant exertion of will, as Kant's language suggests. None of this rules out the possibility of such feeling being involved in moral action: the point is just that Kant holds moral action is possible without it, even if it rarely exists without it. In any case, Wood is correct that what Kant says here is not inconsistent with what he later says in the *MM*, as I am also arguing in this chapter.

²⁵⁰ 4:393.

²⁵¹ 4:394.

value is only instrumental compared to the good will, which he sees as “good in itself and, regarded for itself, is to be valued incomparably higher than all that could merely be brought about by it in favor of some inclination”.²⁵² Nevertheless, when he disqualifies these talents, temperaments, and inclinations from being good in themselves, he is cautious enough to acknowledge that they remain “good and desirable for many purposes”.²⁵³ He adds, “[s]ome qualities are even conducive to this good will itself and can make its work much easier”.²⁵⁴ A guiding point here is that such qualities may be used in good or bad ways; they rely on the guidance of a good will to correct them and bring them under the supervision of a universal, coherent, and morally necessary rule: the moral law. If any of these qualities were to stray by serving a purpose inconsistent with the moral law, the good will (insofar as it is possessed by the agent) would rein them in. Kant therefore envisions the good will as playing a crucial role in providing a kind of rational oversight and moral constraint for action. So, again, it is evident that Kant does not necessarily want to exclude positive inclinations from morally worthy action: here he is concerned that *in action that counts as morally worthy the moral law must restrain an agent, if an inclination is about to lead her to act contrary to its demands.*

Kant acknowledges that there are cases in which an agent may not be capable of overruling inclination with the moral law, though, of course, these would not occur with the ideal moral agent. In his discussion of affects and passions in Kant’s practical philosophy, Patrick Frierson argues it is “wrong to think that, for Kant, emotions are always capable of

²⁵² Ibid.

²⁵³ Ibid.

²⁵⁴ Ibid.

being overridden by practical reason.”²⁵⁵ He points out how affects in particular (as opposed to passions, using Kant’s special meaning for both terms) exert a sudden, surprising, and often strong force on an agent that makes reflection impossible for her.²⁵⁶ In this way, affects bypass “consultation with an agent’s higher faculties of cognition and desire”.²⁵⁷ Since choice is among these higher faculties, affects also bypass choice and cannot be incorporated into maxims. In these cases, an agent is acting directly on inclination “without the reflection required to be ‘capable of doing otherwise’ in any meaningful sense.”²⁵⁸ Since Frierson argues that character lies in the higher faculties, actions from affect look more like the absence of agency or character, rather than the misuse of it. He argues it is nevertheless an absence for which an agent is morally culpable and may be blamed.²⁵⁹ He points out that “we need to distinguish moral responsibility for actions motivated by affects, and responsibility for the affects themselves.”²⁶⁰

Frierson’s account is essentially correct (and illuminating). On the account of character presented in this thesis, these sorts of maximless actions may be understood in a way analogous to unthinking habit. As argued in chapter one, Kant holds that an agent should be in control of her behavior and so is generally against this agent behaving in a way she has not consciously considered. Affect and the sudden impulse it entails also undermines self-control, though it does this by overwhelming an agent’s cognitive faculties rather than proceeding past them undetected. Generally, an agent has a duty to seek out and expunge those features of her empirical character that lead to violations of the moral law, and affects

²⁵⁵ Patrick Frierson, “Affects and Passions” in *Kant’s Lectures on Anthropology: A Critical Guide* ed. Alix Cohen (Cambridge: Cambridge University Press, 2014), 106.

²⁵⁶ *Ibid.*, 101.

²⁵⁷ *Ibid.*, 103-105.

²⁵⁸ *Ibid.*, 105.

²⁵⁹ *Ibid.*, 111-112.

²⁶⁰ *Ibid.*, 112.

arise from those sorts of features. Toward the end of chapter two I argued that the cultivation of Kantian sympathy enables an agent to eventually overcome wrong actions that stem from something like unthinking habit. The acquisition of this sympathy can only happen gradually, through time. Something similar can be said concerning wrong actions resulting from affect: an agent must gradually bring these unconscious processes under conscious control, to prevent sudden outbursts that might overwhelm her cognitive faculties and power of choice. Kantian sympathy may be brought to bear on this task. Further, this task is to be achieved through the cultivation of a tranquil mind and what Kant refers to as “moral apathy”.²⁶¹ He says: “virtue necessarily presupposes apathy (regarded as strength)” and the “true strength of virtue is a *tranquil mind* with a considered and firm resolution to put the law of virtue into practice.”²⁶² So both sympathy and apathy, or tranquility of mind (as Kant defines these things) may work together to prevent an agent from being led astray by inclination manifested as affect.

With his example of the shopkeeper, Kant tries to show why even positive inclinations are not enough to guarantee that the right action is performed, and so need this constraint (and so need to be grounded in the moral law).

For example, it certainly conforms with duty that a shopkeeper not overcharge an inexperienced customer, and where there is a good deal of trade a prudent merchant does not overcharge but keeps a fixed general price for everyone, so that a child can buy from him as well as everyone else. People are thus served *honestly*; but this is not nearly enough for us to believe that the merchant acted in this way from duty and basic principles of honesty; his advantage required it; it cannot be assumed here that he had, besides, an immediate inclination toward his customers, so as from love, as it were, to give no one preference over another in the matter of price.²⁶³

²⁶¹ 6:408.

²⁶² 6:408-409 (the former quotation is a section title from the *MM*). Frierson also acknowledges this as the appropriate solution, in Patrick Frierson, “Affects and Passions,” 112.

²⁶³ 4:397.

In this example, we are meant to see that neither self-interest nor good inclinations, which might dispose the shopkeeper favorably toward an inexperienced customer, are enough to guarantee that he charges her a fair price. We can easily imagine circumstances in which a shopkeeper could get away with overcharging a customer or group of customers; for example, people from a foreign country who do not understand local customs and language. Kant is arguing that only if the shopkeeper charges a fair price because it is the right thing to do (or the moral law requires it, assuming of course that the moral law indeed does require it), can he be relied upon to charge a fair price under circumstances where it will be in his favor not to. In this way, if the agent is fundamentally committed to the moral law, the moral law will act as a constraint on her action. It will do this by preventing her from acting in a way that violates the moral law, if her inclinations direct her to do so.

In summary, Kant's discussion of moral worth in the *GR* aims to establish and illustrate the rational foundation of moral action, just as the rest of the *GR* seeks to establish the rational foundation of ethics generally. It does this by in a series of thought experiments stripping away any incentives an agent might have aside from the moral law in order to see what rational, moral action looks like in isolation or 'purity'. Through this we are meant to see that actions of moral worth are performed because they are required by the moral law. This will mean the agent has the moral law as her maxim, and Kant seems to take it for granted that moral praise is reserved for agents who act from a maxim. While Kant does not say a great deal in the *GR* about what a maxim is, he does indicate that having the moral law as one's maxim or one's ground does not exclude one from having other inclinations that accompany it, so long as the moral law restrains an agent, if her inclinations direct her to act in a way that violates the moral law. So, for the ideal moral agent, the moral law functions as

both a fundamental, sufficient motivating ground and a normative restraint on action generally. In this way, a fundamental commitment to the moral law as one's maxim for action is a central feature of moral character on the Kantian picture.

The Nature of Maxims

As the *GR* explains that *Charakter* is the distinctive constitution of the will, it also explains that in the case of the good will, the form it takes is identical to the form of the categorical imperative – “never to act except in such a way that I could also will that my maxim should become a universal law”.²⁶⁴ In this way, the categorical imperative provides the agent with a structural requirement for action and tells us what constitutes a truly good moral agent, namely an agent whose maxims conform to it. The *REL* deepens this picture by revealing that an agent does not just have one maxim at a time but has many maxims that form a complex, hierarchical structure that bottoms out in a fundamental maxim of either self-love or the moral law. I argue that acting on a maxim, properly understood, is consistent with the broadly Aristotelian conception of character from chapters one and two.

As Onora O’Neill suggests, Kantian maxims are “those underlying principles or intentions by which we guide and control our more specific intentions”.²⁶⁵ These underlying intentions frequently express the “larger and longer-term goals, policies, and aspirations of a life”.²⁶⁶ In addition to these underlying intentions, an agent must have a variety of surface

²⁶⁴ 3:402 and 421 respectively; emphasis redacted in the latter.

²⁶⁵ Onora O’Neill, *Constructions of Reason* (Cambridge: Cambridge University Press, 1989), 84, emphasis redacted.

²⁶⁶ *Ibid.*, 92.

intentions that make the fulfillment of these underlying intentions possible.²⁶⁷ In most cases an agent's action can be traced from more immediate surface intentions to a set of deeper intentions that are more truly descriptive of the kind of agent she is. These maxims constitute the most fundamental features of a person's (moral) character. I say these are fundamental maxims without intending to restrict the scope of a maxim to only these deeper intentions. Kant says a maxim is "the subjective principle of volition", which leaves the matter of whether a maxim is only a surface intention or a deeper intention ambiguous.²⁶⁸ Jens Timmermann highlights this ambiguity between what he calls "thin" and "thick" maxims (i.e., surface and deep intentions) in Kant's philosophy and argues that thin maxims are a precondition for one's having thick maxims.²⁶⁹ Here I am arguing only that, for Kant, so-called thick maxims or deep intentions are constitutive of one's character; for this to be true it does not matter whether surface intentions are also classified as maxims in the general sense. If one is to have moral character and one's actions are to have moral worth, then one's thick maxims must themselves be grounded in the moral law. On this conception, Kantian ethics allows for variability in the expression of individual character, so long as the plans etc. that constitute one's character are further grounded in and constrained by the moral law.

For example, a person might have the intention or goal to acquire the best experiences that life can offer and hold this goal alongside other goals, like being a good friend and being respected by colleagues. This person might have a simple understanding of pleasure and determine that the best experiences life can offer are ones in which the most basic types of physical pleasure are maximized. Say this is her most fundamental concrete goal (maxim)

²⁶⁷ Ibid.

²⁶⁸ 4:402.

²⁶⁹ Jens Timmermann, "Kant's Puzzling Ethics of Maxims" in *The Harvard Review of Philosophy* 8 (2000), 51. For a different discussion of thin and thick maxims, along Williams's understanding of the term, see: Adrian Moore, "Maxims as thick ethical concepts" in *Ratio* 11:6 (2006): 129-147.

that she gives priority over her other life goals. According to Kant, this fundamental concrete goal is further grounded in either self-love or the moral law, according to which principle ultimately controls an agent's maxims. (Evidently, the goal of maximizing pleasure for oneself alone could not be grounded in the moral law.) This goal of pleasure might then lead this agent to pursue a career in which, on balance, she can attain the most resources with adequate opportunities to direct them toward providing pleasurable activities. This would not prevent her from being respected by her colleagues. It may also not prevent her from being a good friend most of the time, though this would depend on luck.²⁷⁰ Her endorsement of this goal might often be tacit or unconscious but, along with the principle of self-love, it is all the same representative of what at the deepest level moves her to act. So, for example, if she were in a position where helping a friend with something important to that friend would foreclose an opportunity for great physical pleasure, she might not help her friend and instead pursue the opportunity for pleasure. Such a life goal might also lead her to construe personal relationships as opportunities to accrue more pleasurable experiences. When presented with an opportunity to seek out a variety of individuals at a social gathering, she would be likely to seek out someone who has the resources to provide more opportunities for physical pleasure (e.g., a minor celebrity or Gatsby-like character who throws lavish parties). In this way, an agent's deeper intentions or goals – her maxims – structure her life and action in very general as well as highly specific ways.

Importantly, for Kant these maxims are to be the focus of moral attention. For it is in virtue of them that an agent's character and actions are evaluated as moral or immoral, good or bad. Our reflective energy ought to be focused on developing these underlying maxims or

²⁷⁰ If she were lucky, she may never have to face a situation in which her maxims conflict at a fundamental level. But most of us are not that lucky and face such difficult choices at various points in our lives. These are opportunities for character formation in the sense I am outlining here.

goals and the further maxims that follow from them, in order to bring them in conformity with and ground them in the moral law. As one does this, one's goals and actions throughout the various domains of one's life can be gradually integrated under the principle of the moral law. This might be the sort of thing Kant has in mind in his reflections on natural perfection, where he says first, to "[c]ultivate your powers of mind and body so that they are fit to realize any ends you might encounter" and then to do the same sort of thing but for the sake of morality, in order to cultivate morality in oneself, since the "greatest perfection of a human being is to do his duty *from duty*".²⁷¹

This account of the nature of maxims is consistent with the notion of character from chapters one and two for several important reasons. The first is that it requires a decisive commitment to the moral law as a fundamental maxim, which is equivalent to the grounding of character discussed in chapter one. Without the moral law as a fundamental maxim, the agent's action can have no moral content. Second, it is consistent with the involvement of inclination (e.g., pleasure, shame, etc.) in moral action, so long as the particular action in question is not dependent on the involvement of inclination and would have been performed even if the agent did not have inclinations in favor of doing the same thing (i.e., the moral law is sufficient as a motivation). More specifically, this means an agent can also have inclinations to do the right thing, without those inclinations ruling out an ascription of moral worth. (Of course, it is implied that the moral law would restrain an agent if her inclinations directed her to act contrary to it.) This leaves room for the account of external moral habit, which does not qualify an agent for ascriptions of moral worth but does aid an agent, particularly children, in the acquisition of moral character. For then, once the grounding of character has occurred, the agent may retain these emotions and incorporate them in moral

²⁷¹ 6:392.

action (that is, action grounded in the moral law). She is not under some strange obligation to lose these emotions so as to act from a maxim of duty alone.

Third, this interpretation of the nature of maxims accommodates the virtue of sympathy outlined in chapter two. It does this by providing a plausible cognitive account of moral action and character more generally, and therefore a psychological backdrop that renders the virtue of sympathy increasingly intelligible. Kant's comment in the *REL* that claims an agent acquires virtue in empirical character "through gradual reformation of conduct and consolidation of his maxims" suggests the acquisition of empirical character is a complex matter involving the acquisition and consolidation of many maxims.²⁷² This would, of course, be the case even when that empirical character is properly grounded in the moral law. Further, the acquisition of sympathy could be understood as part of a life plan. At its deepest level, this life plan would be grounded in and directed by the moral law. It would also require the acquisition and refinement of a number of surface maxims. This set of surface maxims would have both an obligatory and permissive element: the agent would be obligated to acquire the sort of specific maxims that would enable an adequate knowledge of the needs of others, so that she could perform her duty toward them. The agent would also be permitted to acquire maxims that involve her own pleasure, such as the enjoyment of friendship, which has already been mentioned.

²⁷² 6:47.

A Unified Account of Moral Worth

Definition of Moral Worth

In summary, in the *GR* Kant holds that the moral worth of an action is determined by the maxim that forms the basis of that action. This is consistent with his understanding of character, as indicated by his discussion of the German term at important points in his practical philosophy, and his conception of a maxim. So, for Kant, moral worth and character are deeply linked, and both place central emphasis on an agent's maxim. Agents have many maxims and the relationship between an agent's many maxims is complex. An agent's character and moral worth will fundamentally be assessed on the basis of her fundamental maxim in any given action, which in turn is grounded in a maxim of either self-love or the moral law. Importantly, given the nature of maxims, this allows an agent to act from the moral law while simultaneously acting from surface-level maxims that incorporate inclinations such as pleasure, so long as these inclinations are constrained by the moral law. The actions of such an agent can be considered to have moral worth. As I have mentioned, Kant seems to take it for granted that moral praise is reserved for agents who act on a maxim. The importance of maxims to the Kantian conception of moral worth means that moral worth has a necessarily cognitive or intellectual dimension. However, in his explicit discussion of moral worth, Kant does not emphasize the importance of having good judgment, or beliefs that are also accurate and well-informed about the actual moral import of one's action – in other words, whether one's action, for example, really does respect the dignity of another person when that is what one intends. As I argued in chapter two, the possession of accurate information like this is entailed in Kant's notion of sympathy, which he recognizes as a kind

of privileged, intellectual virtue analogous to *phrōnesis*. And sympathy is crucial to Kant's conception of character and necessary for full moral character (in both the noumenal and empirical sense). I will argue it is also necessary for unequivocal or full moral worth. Though this might go beyond what Kant recognized in his explicit discussion of moral worth, it is a helpful distinction to make and is consistent with his understanding of moral character. An agent with full moral worth will then have the moral law as a fundamental maxim (in other words, will have a morally good disposition) and a well-developed disposition sympathy, or good judgment concerning the application of moral principles.

If we bring these various lines of thought together, we can provide a unified definition of moral worth that says: *full moral worth is a category of praiseworthiness attributed to actions that proceed from a morally good disposition and sympathy*. By way of clarification, this means that: (a) *moral worth can be attributed to an action that proceeds from inclination, so long as the moral law would restrain an agent if that inclination is about to lead her to act contrary to the moral law*, (b) *moral worth can be attributed to an action if that action were morally required and would have been performed anyway, even if other motives had not been present*, and (c) *for an agent to receive an unequivocal attribution of moral worth, i.e. full moral worth, she must both have and exercise a well-developed disposition of sympathy, that is, she must possess and exercise good judgment in matters concerning morality and its application.*²⁷³ Such an agent will also have taken the initiative to undertake an empirical inquiry of sorts to discover basic facts about others and what

²⁷³ There may be instances where an agent acts on the moral law, exercises such sympathy, and yet despite her best efforts the situation goes awry and the morally desirable outcome does not come about. In these situations, an agent like this will still merit an ascription of moral worth, provided the agent's disposition is good, since moral worth is not dependent on consequences. As Kant is known to believe, the disposition of an agent in this situation would "like a jewel [...] still shine by itself, as something that has its full worth in itself" (4:394).

contributes to their well-being, and so be able to judge and act rightly in moral situations.²⁷⁴ Strictly speaking, this is a process of information gathering for the purpose of developing sympathetic perception, rather than a direct exercise of sympathy itself. However, sympathy as defined in chapter two remains at the heart of this work. Moral worth in the mitigated sense might then be attributed to an agent that has a morally good disposition though she does not have well-developed sympathy, or good judgment concerning the application of moral principles (I will discuss this further below). So we may add to the above: (d) *an agent may receive an attribution of mitigated moral worth when she exercises the best moral judgment that can reasonably be expected of her, even though it may not meet the standard for full moral worth as stated in (c).*

The notion of mitigated moral worth, which I will outline in more detail in what follows, introduces some complexity into the notion of moral worth; I suggest this complexity fits well with the intuitive or everyday sense in which we think of moral agents, which I take to be a good thing. On the one hand, we can reasonably hold that individuals like Shylock and Eichmann do not deserve any attribution of moral worth for acting on what they allegedly understand to be right principles, due to a gross failure in judgment concerning the application of these principles: we would want to say they deserve blame. Good moral judgment here means the ability to discern the morally salient features of the empirical world and correctly apply the principle of the moral law to it, enabling one to correctly understand what one's duty is. Of course, in the case of a fictional Eichmann (that is, a case in which the real Eichmann could be taken at his word when he reported he believed he was following Kant's categorical imperative as he organized the Holocaust, the truth of which is highly

²⁷⁴ As Kant says: a priori laws of moral philosophy "no doubt still require a judgment sharpened by experience, partly to distinguish in what cases they are applicable and partly to provide them with access to the will of the human being and efficacy for his fulfillment of them" (4:389).

dubitable), his was not merely a failure of judgment but a failure to inquire into basic empirical facts concerning the humanity of Jewish victims. Sympathy would have aided in such an inquiry. On the other hand, we recognize that individuals like, say, Nelson Mandela deserve attributions of full moral worth for having the right principles and good judgment in the application of those principles.²⁷⁵

We might think that most agents fall somewhere along a continuum between these poles. However, for Kant, each agent is either acting on the moral law or self-love, and so an agent's action either has moral worth or does not. Aside from this moral worth, the comparative moral *value* of their actions can be evaluated by the consequences of those actions for human dignity and well-being.²⁷⁶ For example, we might say that Eichmann is worse than Shylock because his lack of good moral judgment brought about more human suffering. Likewise, for the same reason, Shylock might be worse than an otherwise benevolent slave owner from the antebellum Southern United States (or perhaps not, depending on the consequences of their respective ill judgment). Further, we might want to believe that some agents are at bottom fundamentally good and in fact committed to doing what is right, whether they judge that correctly or not. In some cases, they may be successful in doing the right thing, while in other cases they may fail. I will discuss two examples of an agent who in one case does not merit an attribution of moral worth and in another case, does merit an attribution of *mitigated* moral worth. Though his circumstances are somewhat remote from ours, he is still perhaps typical of the sort of agents we are: when many of us are

²⁷⁵ For example, amidst imprisonment and gross mistreatment by whites under the apartheid regime, Mandela reports: "I knew that people expected me to harbor anger toward whites. But I had none. In prison, my anger toward whites decreased, but my hatred for the system grew. I wanted South Africa to see that I loved even my enemies while I hated the system that turned us against one another." Nelson Mandela, *The Long Walk to Freedom* (London: Abacus, 1995), 669-753. What is relevant here is not Mandela's emotion itself, but the way he structures and directs it. He does not allow this mistreatment to bring him to hate his enemies or deny their humanity; instead, he loves them (and we might say, respects their humanity).

²⁷⁶ Of course, this does not include moral unworth, or those agents who do not have a moral disposition.

at our best, we are merely on the way to achieving good moral character and imperfect in our intellectual grasp of it. I suggest it makes sense to say that agents like this can be morally worthy, even if imperfectly so.

Points (a) and (b) of the account of moral worth I suggest is similar to the accounts suggested by several others, including Nancy Sherman and Barbara Herman insofar as it allows that inclination may participate in actions of moral worth, provided the moral law performs a regulatory function in those actions. For example, Sherman argues correctly that Kantian virtue is a hierarchically structured composite with the disposition to act from moral principles as its primary constituent, and cultivated natural capacities (i.e., inclinations) that support that disposition, as its secondary constituent.²⁷⁷ Herman points out that moral worth does not turn on the presence or absence of inclinations supporting an action required by duty, but on whether that inclination is included in the maxim as a determining ground for that action.²⁷⁸ She adds that Kantian motives are reasons for action and are not desires or causes; therefore, provided an agent takes the moral law as her reason for action, it does not matter whether she has further desires to perform that action, since these desires are not ultimately what control her.²⁷⁹ This is what it looks like to have the moral law as a primary motive. Herman continues to say that in cases where duty is a secondary motive, for example, in actions that are morally permissible but not morally required, the moral law provides limiting conditions for the way other motives can be acted on.²⁸⁰ In this way, it

²⁷⁷ Nancy Sherman, *Making a Necessity of Virtue*, 158.

²⁷⁸ Barbara Herman, *The Practice of Moral Judgment*, 11.

²⁷⁹ *Ibid.*, 12.

²⁸⁰ *Ibid.*, 14-15.

serves “to keep us free of the effects of temptations in ordinary situations that can suggest morally prohibited courses of action.”²⁸¹

Paul Guyer expresses skepticism about this general approach which allows the moral law to act as a restraint or a limiting condition. He believes treating duty in this manner as a “limiting condition” differs from Kant’s own position “by treating inclinations as potentially independently motivating factors in the first place, that is, factors that are themselves ever *sufficient* to bring about an action, which can be *allowed* to move the morally worthy agent as long as they are consistent with the requirements of duty”.²⁸² He believes this approach presents a problem because it does not take seriously enough the idea that inclinations never move an agent unless they have been incorporated into the agent’s maxim, and it further does not seem to capture the idea that “morally worthy action must arise *from* duty”.²⁸³ However, this overlooks that an agent can have the moral law as a fundamental maxim and yet incorporate inclination in a higher-level maxim: in short, Guyer does not here acknowledge that maxims can occur in a complex, hierarchical structure, as Sherman has argued and as I have argued above. According to the picture of maxims presented above, if we could trace the inclination along the hierarchy of maxims to its fundamental ground, we would find that it is ultimately embedded in a moral maxim, namely, the moral law. If we are to take Kant’s account of maxims in the *REL* seriously then, inclinations are never really independent motivating grounds, because they are always also grounded in either self-love or the moral law, which we might think of as second-order motivations.²⁸⁴ Contrary to Guyer’s understanding of it, this interpretation does not suppose that an agent can simultaneously be

²⁸¹ Ibid., 17.

²⁸² Paul Guyer, *Kant on Freedom, Law, and Happiness* (Cambridge: Cambridge University Press, 2000), 293.

²⁸³ Ibid. Guyer points to Henry Allison’s notion of the “incorporation thesis” to ground this claim; see Henry Allison, *Kant’s Theory of Freedom*, 5. See also my dispute of this point in note 210 above.

²⁸⁴ Of course, if the agent is motivated by self-love, Kant holds she is fundamentally motivated by inclination.

committed to the moral law and another fundamental maxim that might permit an action incompatible with it.²⁸⁵ This is because these maxims are not here equal and competing, but rather one is embedded in and conditioned by the other. In the case of the ideal moral agent, maxims involving inclination are only located higher up (or toward the surface) in the chain of maxims and never in the fundamental ground of that chain. There is therefore no contradiction here.

Further, it is not correct, as Guyer suggests, that understanding the moral law as a limiting condition implies that an agent is merely “a passive subject in whom feelings and inclinations simply occur although an active agent insofar as he can restrain the causal efficacy of those causes by virtue of the limiting condition of respect for duty.”²⁸⁶ The moral law remains operative in the ideal moral agent as her fundamental maxim and in that way is operative in all action, including a kind of action from inclination (where inclination forms a surface-level maxim that is ultimately grounded in the moral law), because it performs the role of permitting that action. The moral motive may not be noticeable or at the forefront of an agent’s consciousness at these times, yet it is all the while present and effective, in the way that a radar system can remain effective when there are no threats to detect. As long as the agent chooses the moral law as her ground, she is an active and not a passive subject, and so does not make of herself “a plaything of the mere inclinations”.²⁸⁷

Elsewhere, in what he describes as “the phenomenal etiology of morally worthy action”, Guyer suggests that “Kant’s claims about the noumenal determination of the will become irrelevant to his empirical story about the etiology of moral action, and can safely be ignored in figuring out the latter even if they are accepted as a metaphysical account of

²⁸⁵ Paul Guyer, *Kant on Freedom, Law, and Happiness*, 295.

²⁸⁶ *Ibid.*, 298.

²⁸⁷ 6:420.

freedom of will.”²⁸⁸ Transcendental idealism is not necessary to the account of moral worth outlined here; however, it is worth considering it briefly insofar as its rejection has implications for Guyer’s understanding of sympathy and moral motivation. Guyer is right that strictly speaking we cannot know an agent’s motivation on a noumenal level. However, the general hope he expresses toward the end of this paper, to “leave Kant’s transcendental idealist theory of the freedom of the will in the dustbin of history” and retrieve from Kant an independently standing empirical theory of moral motivation is more significantly out of keeping with Kant’s project than he suggests, even if we bracket the metaphysics of transcendental idealism.²⁸⁹ For Kant generally seems to allow that we can reasonably infer an agent’s maxims or fundamental motivation in many cases. His discussion of the grounding of character in the *AN*, and of character more generally in the *REL*, suggest this, as does his discussion of moral worth in the *GR*. In the *GR*’s familiar thought experiment involving the philanthropist, Kant imagines him with all positive inclinations removed in order to demonstrate that he does the right action “without any inclination, simply from duty” so that “genuine moral worth” can be attributed to this action.²⁹⁰ To accept Guyer’s account, which removes the noumenal from consideration altogether, would undermine the purpose of this thought experiment and the ability to infer anything about moral worth altogether, since moral worth is attributed to noumenal character alone. This is to move much farther from Kant than we should, if the goal is to achieve an interpretation of Kant that lies near to what

²⁸⁸ Paul Guyer, “Moral Feelings in the *Metaphysics of Morals*” in *Kant’s Metaphysics of Morals: A Critical Guide* ed. Lara Denis (Cambridge: Cambridge University Press, 2010), 132.

²⁸⁹ *Ibid.*, 151.

²⁹⁰ 4:398.

he actually (probably) thought. This may not be Guyer's goal here, but it is worth making that more clear.²⁹¹

Something similar happens with Guyer's interpretation of Kantian sympathy in the same paper. He says it is "only by the means" of sympathetic feelings that "we are ever capable of performing the beneficent actions that are called for by the maxim of beneficence in appropriate circumstances."²⁹² This claim is tantamount to a rejection of noumenal character and so is significantly out of keeping with Kant's project, once again, even after the details of transcendental idealism are put to one side. Further, it is the express purpose of the philanthropist of *GR I* to show first, that sympathetic feeling alone is not sufficient to motivate morally worthy action, and second, that morally worthy action can be performed without any trace of such feeling. So, while morally worthy action may include any degree of sympathetic feeling, contrary to Guyer it does not play a "causal role" in a phenomenal process of morally worthy action so sharply severed from the noumenal will.²⁹³ Sympathy may aid morally worthy action in the ways outlined in chapter two, but it may not ultimately cause it. Instead, the more apt picture concerning the role of sympathetic feeling in moral motivation is of the chain of embedded maxims described above, whereby action from an inclination like this may count as morally worthy if it is ultimately grounded in and constrained by the moral law as fundamental maxim.

²⁹¹ It is also why I set to one side more detailed questions concerning transcendental idealism in this thesis. It is perhaps likely that Kant himself did not quite feel the force of these questions, since he does not treat them in much detail and more or less proceeds with his practical philosophy as though unhindered by them. If these questions ultimately cause problems for Kant, they would be problems of coherence on the part of his broader system and insofar as it depends on transcendental idealism. However, if we are willing to set these more detailed questions aside, I suggest we make the sensible assumption that noumenal character obtains and, while it may be difficult to discern, it is not impossible to discern (this seems to be the assumption made by many contemporary Kantians with respect to the practical philosophy). This is different from Guyer's wholesale rejection of the noumenal will. This enables us to see that Kant provides a very interesting and workable account of moral agency and moral character, which I have taken as being the burden of this thesis to demonstrate.

²⁹² Paul Guyer, "Moral Feelings in the *Metaphysics of Morals*," 148.

²⁹³ *Ibid.*

Friendship

It is helpful to look at friendship as an example of morally worthy action. I would like to argue that there is a Kantian account of friendship that shows how the moral law can motivate and structure an agent's action in the presence of accompanying inclinations and without entering explicitly into consciousness. Friendship is standardly taken to be a feature of a good human life. In the *NE*, Aristotle identifies three types of friendship: friendships of utility, pleasure, and virtue.

Those who love each other for utility love the other not in his own right, but insofar as they gain some good for themselves from him. The same is true of those who love for pleasure; for they like a witty person not because of his character, but because he is pleasant to them. Those who love for utility or pleasure, then, are fond of a friend because of what is good or pleasant for themselves, not insofar as the beloved [friend] is who he is, but insofar as he is useful or pleasant.²⁹⁴

In Kantian terms, we might say that friendship based on utility or pleasure does not treat the friend as an end in herself, but rather as a means to an end. By contrast, Aristotle describes friendship based on virtue as complete.

But complete friendship is the friendship of good people similar in virtue; for they wish goods in the same way to each other insofar as they are good, and they are good in their own right. Now those who wish goods to their friend for the friend's own sake are friends most of all; for they have this attitude because of the friend himself, not coincidentally. Hence these people's friendship lasts as long as they are good and virtue is enduring. Each of them is both good without qualification and good for his friend, since good people are both good without qualification and advantageous to each other. They are pleasant in the same ways too, since good people are pleasant both without qualification and for each other.²⁹⁵

Here we see that friendship based on virtue only obtains between two people who are able to "wish goods to their friend for the friend's own sake". In being able to do this, they are

²⁹⁴ *NE*, 1156a9-18.

²⁹⁵ *NE*, 1156b7-17.

valuing “the friend herself”. In Kantian terms, we might say they treat the friend as an end in herself. This does not preclude these friends from obtaining advantage or pleasure from each other and from their friendship. Rather, it means that any advantage or pleasure they do receive is conditioned by virtue, or consistent with the moral law and treating the friend as an end in herself.

We can further imagine the motivational structure of each agent (friend) in the virtuous friendship in Kantian terms. For the sake of argument, let us assume the actions of the virtuous friend are always morally worthy. (As per the definition above, this would mean that her actions could proceed from inclination, provided the moral law restrains her if these inclinations are about to lead her astray; it would also mean that if an action is morally required, she would perform that action even if motives other than the moral law had not been present.) The virtuous friend would have the moral law as a second-order motivation and so the actions and intentions that go into any relationship she enters into, including the virtuous friendship, would be grounded in the moral law. Yet this would not preclude her from acting on inclination that is consistent with the moral law and respect for her friend. In this way, she might go for a walk or have a long conversation with her friend because, on the face of it, she enjoys doing these things. If her friend became ill, she might visit her in the hospital out of concern for her as well as a desire to share her company and know her condition firsthand. Are these desires moral in the Kantian sense? *Prima facie*, the maxim they are based on is one of pleasure: it does not enter into the agent’s mind that she has an obligation to visit her friend in the hospital, nor does she engage in conversation out of the explicit motive of respect for her friend’s dignity. If she were asked why she does these things, she might respond by saying she does them because she cares for her friend. But on

the Kantian model, this maxim of care can be incorporated into a deeper maxim that is itself incorporated into the moral law as a fundamental maxim.

For example, the desire to share a friend's company and inquire after her condition in a time of need could be incorporated into the greater goal of building and maintaining good personal relationships. That these personal relationships be good in this way – beneficial for both parties and their pursuit of the good – can mean this goal is grounded in the moral law. The agent does not need to be conscious of the moral law in every action, since the moral law does not need to be the direct or exclusive motivation in every action (though it does need to be sufficient to motivate an agent to action). Yet every such action must, as Kant says, have the moral law as a “first ground” (*ersten Grund*), or what we might call a second-order motivation.²⁹⁶ In other words, the moral law must be the grounding motivation. This means that every action remains accountable to the moral law, so the desires or inclinations the agent acts on at a higher level must be constrained and to some extent directed by it. It is not until either constraint or direction is necessary that the moral law becomes salient as this fundamental ground. For example, if the moral agent's friend were to ask her to participate in a fraudulent scheme or (for a more mundane example) malicious gossip, the moral agent would say ‘no’ on the basis of this moral constraint. This might be uncomfortable and may even threaten the relationship; however, it is often in situations like these that personal character is revealed. Also, it is often in situations like these that the character of a friendship is revealed (for example, whether it is a friendship of utility or virtue).

Consider another example: an agent may feel inclined to do something other than visit her friend in the hospital. Perhaps she has had a long day and is inclined to stay at home and watch television rather than visit her friend, who she knows is lonely and would

²⁹⁶ 6:21.

appreciate the company. In this case, her commitment to duty may direct her to overcome this inclination and visit her friend. It is true that if she too often had to be compelled by duty in this way, the friendship would seem less like friendship and therefore less desirable. It is worth noting that a Kantian understanding of moral worth does not necessitate this barren picture. Nevertheless, it is also worth noting that from time to time we do face cases like this in which we must serve our friends from duty and to some extent without immediate desire, if we are to maintain our friendship or serve our friend at all. Sometimes good friendship requires coming through for a friend when it is least convenient or most difficult. In these moments it is obvious that, in Aristotle's words, one is wishing "goods to their friend for the friend's own sake", or in Kantian terms, one is treating the friend as an end in herself.²⁹⁷ More generally, from time to time we face moments of decision in which we must evaluate and in some sense justify our life projects, including our personal relationships. In moments like these, the moral law provides the agent with a standard that aids in this justification by eliminating moral impossibilities and motivating moral action. Contrary to Williams, the moral motive does not provide the agent with one thought too many, since in most cases it will not be explicit.²⁹⁸ However, in some contexts the moral motive will become explicit, and in all contexts it is necessary as a ground for moral action in the way explained above. Kant takes for granted that people will pursue the kind of projects, relationships, and pleasure that

²⁹⁷ *NE*, 1156b10-11. For Kant, this will be the fulfillment of a wide duty. Interestingly, there is some ambiguity here, since a wide duty (say, of beneficence) gives a degree of latitude to an agent as to how she will fulfill it. So she might not strictly speaking be duty-bound to, for example, act beneficently toward her friend by visiting her in the hospital. Yet she is duty-bound to be beneficent and, given that she is perhaps in the best position to meet a close friend's need and in virtue of their friendship that friend might be counting on her to do so, it *almost* seems she is duty bound to visit. We might think she has a duty to visit and, in any case, this does not mean that she will at some point need the moral motive to act in this kind of way. Yet it seems Kantian ethics cannot strictly speaking say she *must* act beneficently in any particular case, just that she must act beneficently.

²⁹⁸ In contrast to Williams' assessment (cf. Bernard Williams, *Moral Luck*, 18).

make life meaningful.²⁹⁹ His concern is to articulate this moral motive and a system based on it, in order to guide and ground these life projects.

Degrees of Moral Worth

The above suggests a generous definition of moral worth, since it allows that seemingly standard actions consistent with an agent's inclinations can have moral worth, provided they are ultimately grounded in the moral maxim. This might lead us to think that all of an agent's actions subsequent to the change of heart would have moral worth: since her fundamental maxim would have changed to the moral law, all of her actions would be traceable to that maxim and so must have moral worth. This is not as odd as it sounds. What it does is shift the focus of moral worth from the relative difficulty of specific actions and the consequent merit accruing to an agent in virtue of performing those actions (what Richard Henson would refer to as a citation for gallantry in a battle against evil), to the fundamental ground and content of an agent's character (her character in the deep sense).³⁰⁰ It is true it would allow moral worth to be ascribed to an agent even for an action as mundane as conversing with a friend out of sheer delight, provided this surface maxim of delight is grounded in the moral law as described above. However, this is just to make moral worth more about agents than actions, since such an agent is estimable for the moral quality of her character across all actions, so long as she is acting out of that character.³⁰¹ This looks much like Henson's second definition of moral worth as moral fitness, which allows for the

²⁹⁹ He says, for example, "it is unavoidable for human nature to wish for and seek happiness" (6:387).

³⁰⁰ Regarding Henson's remark, see: Richard Henson, "What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action," 42.

³⁰¹ Kant speaks in the *GR* of actions having moral worth, which is why I have made actions the locus of description in my definition of moral worth above. However, since all actions proceeding from fundamentally moral character have moral worth, it still makes sense to say the focus of moral worth has shifted from character in specific actions to character in a deeper sense.

overdetermination of dutiful acts by inclination, provided the agent is in a fit moral condition at the time of the act.³⁰²

However, this definition appears to leave open the possibility of ascribing moral worth to agents who make honest mistakes. Prima facie it might seem wrong to ascribe moral worth to such agents, as it does not align with an idea of moral fitness in the standard sense. For the idea of moral fitness would seem to presuppose that an agent can exercise her practical reason effectively across a variety of situations, including those that present special difficulty for moral judgment. That is, a concept of moral fitness seems to entail a certain standard of cognitive fitness. So, despite my emphasis so far in this chapter on the permissibility of participating inclinations in morally worthy action, there must also be a substantive cognitive aspect to this definition of moral worth. And yet there will be some agents who are fundamentally committed to the moral law, are actively pursuing a course of moral fitness, and nevertheless have not yet acquired the sort of good judgment in applying moral principles we would expect of an agent with full moral worth. Such agents might make honest (though not gross) mistakes concerning moral action. This suggests two things: first, that we need to go beyond Kant's explicit discussion of moral worth to arrive at a satisfying definition of full moral worth and second, that even those agents who do not have good judgment concerning the application of moral principles (and so do not merit full moral worth) do merit moral worth in some mitigated sense of the term. Though Kant does not seem to recognize this distinction between what I have called full and mitigated moral worth, the theory of character I have outlined so far provides a way of thinking about moral agents

³⁰² Richard Henson, "What Kant Might Have Said: Moral Worth and the Overdetermination of Dutiful Action", 42. Though Henson suggests that Kant's "notion of moral worth is really two notions, [albeit] complementary and each perfectly coherent", whereas I am arguing it is one unified notion, roughly similar to his idea of moral fitness (*ibid.*, 53-54).

that corresponds well to this distinction. The most salient feature of this theory for this purpose is the disposition of sympathy from Kant's later work.

Kant does not engage in as focused a discussion of moral worth in his later work as he does in the *GR*. Nevertheless, it is reasonable to hold that a Kantian theory of moral worth can go beyond his explicit comments on the matter and draw from other important concepts in his moral philosophy. With this in mind, I suggest that Kant's notion of moral worth be understood against the backdrop of his complete conception of moral character. As argued in chapter two, moral character is a kind of dual concept for Kant, since it is divided into an intelligible and an empirical part. Since according to this dual concept, full moral character is constituted by a good will and well-developed sympathy, it makes sense to hold that full moral worth be constituted by the same, with the qualification that that well-developed sympathy and the above mentioned empirical inquiry associated with it be appropriately exercised. Since sympathy is a matter of having right beliefs concerning the well-being of others, to be eligible for an ascription of moral worth, an agent will have to have these same right beliefs. Full moral worth therefore has a crucial cognitive component, as stated above in point (c): *for an agent to receive an unequivocal attribution of moral worth, i.e. full moral worth, she must both have and exercise a well-developed disposition of sympathy, that is, she must possess and exercise good judgment in matters concerning morality and its application.* As I have mentioned, this agent will also have had to undertake an empirical inquiry of sorts to discover the morally relevant facts about others and the circumstances of her action, to guide her understanding of what constitutes morally right action in those circumstances. Further to this, I add that: (d) *an agent may receive an attribution of mitigated moral worth*

when she exercises the best moral judgment that can reasonably be expected of her, even though it may not meet the standard for full moral worth as stated in (c).

Somewhat controversially perhaps, this definition disqualifies certain types of agent from ascriptions of full moral worth. Consider the example of Huck Finn, who decides not to report his friend Jim, who is an escaped slave, on the basis of what appears to be a vague sense of fellow feeling and despite believing he is doing the wrong thing by not reporting him.³⁰³ Nomy Arpaly asserts that Finn constantly perceives but does not act upon the fact that Jim is a person just like him. She believes that although he cannot do the cognitive work to bring this fact to consciousness, he is acting from it as a reason and so is not acting on mere fellow feeling alone, but on a genuine concern for morality.³⁰⁴ On Arpaly's conception, knowledge is not necessary to moral worth because "moral worth is fundamentally about acting for moral reasons, not about acting for reasons believed or known to be such".³⁰⁵ She believes that if we required knowledge of right reasons for moral worth, we would have to view people like Finn as fundamentally bad, yet having some fortunate inclinations in their make-up. However, she suggests we more commonly treat these people as fundamentally good people who are responsive to moral reasons but happen to be incompetent abstract thinkers, and a definition of moral worth should track this common view.³⁰⁶ According to Arpaly's definition, agents like Finn qualify for ascriptions of moral worth in cases like these. However, Arpaly presents us with a false dichotomy: we do not have to view Finn as fundamentally bad if we decide that he does not qualify for an ascription of unequivocal or full moral worth. I suggest that in the case that Arpaly and Bennett refer to, Finn does not

³⁰³ This example is first discussed in Jonathan Bennett, "The conscience of Huckleberry Finn" *Philosophy* 49 (1974): 123-134.

³⁰⁴ Nomy Arpaly, *Unprincipled Virtue* (Cambridge: Cambridge University Press, 2003), 75-77.

³⁰⁵ *Ibid.*, 73.

³⁰⁶ *Ibid.*, 78.

qualify for an ascription of moral worth in either the full or mitigated sense of the term. Finn does not qualify for an ascription of moral worth in the full sense because he does not have right beliefs and appropriately developed sympathy, or good moral judgment. He does not qualify for an ascription of moral worth in the mitigated sense because he is not in this case fundamentally motivated by what is right, which in this case implies acting for the sake of Jim's well-being. However, Twain's story provides us with another example in which Finn is presented with a choice about whether he will report Jim or not. I argue that in this later example, Finn may qualify for an ascription of moral worth in the mitigated sense, because he makes a firm resolution to do what is right and help Jim to attain freedom – something analogous to the Kantian change of heart – and his moral judgment (sympathy) is perhaps as good as can reasonably be expected of him.

A closer look at how Finn handles the situation discussed by Arpaly casts doubt on the positive moral status of his reasons for not reporting Jim in the first instance. We know that his conscience errs when he prioritizes Miss Watson's supposed right to ownership of Jim over Jim's freedom, and it errs again when Jim is lowered in his eyes for wanting to 'steal' his children out of slavery in the case that their owner is not willing to sell them to Jim.³⁰⁷ (Of course, there is a deep irony here. Jim is so concerned to do right by prevailing moral standards that his first choice is to work until he can afford to buy his children's freedom; he is only willing to steal them as a last resort, if their owner refuses this possibility. This communicates the depth to which the contours of the prevailing moral system are internalized by characters in Twain's novel and by implication, people of that time, much to their detriment.) Yet, of course, Finn just could not give Jim up. When the slave hunters inquire after the other man on his raft, Finn fumbles his words and eventually

³⁰⁷ Mark Twain, *Adventures of Huckleberry Finn* (Oxford: Oxford University Press, 2008), 82.

lies, recounting this event in terms of a weakness in will.³⁰⁸ This much is already familiar. However, Finn's backsliding into right action does not speak as well of his character as we might want to believe. After further amiable interaction with these men, Finn reports feeling very bad and even depressed, because he 'knew' he had done wrong. Then his mind goes to work on the situation:

Then I thought a minute, and says to myself, hold on, - s'pose you'd a done right and give Jim up; would you felt better than what you do now? No, says I, I'd feel bad – I'd feel just the same way I do now. Well, then, says I, what's the use you learning to do right, when it's troublesome to do right and ain't no trouble to do wrong, and the wages is just the same? I was stuck. I couldn't answer that. So I reckoned I wouldn't bother no more about it, but after this always do whichever come handiest at the time.³⁰⁹

Here Finn appears less concerned with and responsive to Jim's humanity than he is with pacifying his own troublesome feelings.³¹⁰ He formulates a life plan on the basis of this, to always do what is most expedient, or "handiest" at the time. A resolution to do what is most expedient in the moment hardly merits an attribution of moral worth. Looking back on the moments immediately preceding Finn's encounter with the men, it is Jim's expression of friendship and admiration of Finn that ultimately tips him in favor of doing what he (incorrectly) believes to be wrong.³¹¹ Finn is much more powerfully steered by self-interested feelings and a casual liking of Jim than he is by Jim's humanity as such. He has neither correct beliefs (in the explicit sense), nor does he act on correct principles (in the implicit, even unconscious sense): his de re right action is primarily a function of self-interest. Since

³⁰⁸ He says: "I didn't answer up prompt. I tried to, but the words wouldn't come. I tried, for a second or two, to brace up and out with it, but I warn't man enough – hadn't the spunk of a rabbit. I see I was weakening; so I just give up trying, and up and says – 'He's white.'" (Ibid., 83.)

³⁰⁹ Ibid., 85.

³¹⁰ For further discussion of the role such feelings may play in matters of conscience and moral worth, see Thomas Hill, "Punishment, Conscience, and Moral Worth" in *Kant's Metaphysics of Morals: Interpretive Essays* ed. Mark Timmons (Oxford: Oxford University Press, 2002).

³¹¹ Jim says: "'Dah you goes, de ole true Huck; de on'y white genlman dat ever kep' his promise to ole Jim.'" Finn reports: "Well, I just felt sick. But I says, I got to do it – I can't get out of it" (ibid., 83).

full moral worth requires full moral character, in this case Finn does not qualify for an attribution of full moral worth. Since mitigated moral worth allows for incorrect beliefs yet still requires a fundamental commitment to the moral law, given the evidence at hand we cannot say that Finn qualifies for this either.

This does set the bar relatively high for ascriptions of moral worth in the full sense, but it is not unreasonable to do so. An agent's beliefs about morality are important because they often provide the determining ingredient in moral action. Finn's vacillation in this instance and later (unsuccessful) resolve to report Jim to Jim's 'owner' Miss Watson suggests the likelihood of him sticking by his decision act in Jim's best interest is at this point unstable, even if somewhere deep down he is aware that Jim's interest is deserving of moral attention. It is in this sense that it arises from self-interest rather than moral character: in this particular instance the right moral action won out, but we cannot be confident it will win next time, because Finn is not fundamentally concerned to perform the right action as such. In a way, it is fortunate that in this encounter with the slave hunters Finn fails to do what he (mistakenly) believes is right and does not report Jim, because he lacks the requisite strength of will. But it is not unreasonable to think that his mistaken conceptions of what is right could win out and things could easily go differently next time. Finn's moral position on the matter is hardly estimable and leaves the reader with an impression of precariousness.

I am suggesting that to some extent empirical knowledge and right beliefs about morality (practical anthropology and sympathy, in the cognitive sense outlined in chapter two) help to stabilize an agent and enable reliable, good judgment in counterfactual situations.³¹² It is also reasonable to hold that this stability of character is part of what makes

³¹² There is an objection against this claim from the situationist camp in contemporary moral psychology, but I will not deal with it in this thesis.

a morally worthy agent exceptional and so deserving of praise. So, even if Finn were acting on the moral law as a fundamental maxim, which his resolution to expediency rules out, his false beliefs concerning its application ought to disqualify him from full moral worth.³¹³ This does not present a serious challenge to standard intuition, since we can believe Finn performed the right action in a way that is not morally worthy. After all, it is evident that his moral framework is deeply confused.

We can imagine variations on Finn's case that bring about an outcome other than the moral one. Perhaps an alternate but similarly constituted Huck Finn did not meet Jim, who had expressed positive feelings toward him, who Finn judged to be a kind and likeable person, and with whom Finn had until that point already shared a number of meaningful personal experiences. Perhaps he instead met Jack, and Jack was much less pleasant than Jim. Due to the hardship and brutality Jack experienced at the hands of white slave owners, Jack strongly disliked all white people and spoke to them with disdain when he spoke to them at all. Perhaps Jack was prudent enough not to openly reveal his dislike of Finn, but Finn could not help but get the sense that Jack begrudged him, and he did not like Jack in return. Perhaps Jack angered Finn rather than expressed admiration of him immediately before Finn came upon the men looking for runaway slaves. Would Finn in this case respond to the men in the same way? It is difficult to say. At the point in the story when he encounters these men we cannot say that Finn is truly committed to doing what is right, or respecting others for the sake of their dignity rather than his own self-interest. Since his stated goal is expediency, it may be that if Jack were not helpful enough to Finn, it would be most

³¹³ Of course, to say that Finn is acting on the moral law in telling this lie runs contrary to Kant's beliefs about lying as evident, for example, in his essay, *SRL*. But we are not obligated to share Kant's beliefs on this point and can exclude them for the sake of this example.

expedient for Finn to turn him in for the reward on offer.³¹⁴ We can easily imagine a situation in which it is genuinely possible, if not probable, that he would do this. It may even have been that at this point in the story, Finn acted in Jim's best interest simply because he had grown attached to Jim. In any case, the integration of a commitment to the moral law and right beliefs concerning its application – namely that Jack deserves equal respect to Jim or himself – is necessary to ensure that Finn acts rightly in this case.

Later in the story, Finn considers handing Jim over a second time, this time by writing to Jim's 'owner' Miss Watson and disclosing Jim's whereabouts as well as confessing to the supposed wrong Finn had committed in helping Jim to escape.³¹⁵ He does this primarily to relieve the deeply troubling feelings his conscience causes him, though the draft of his letter to Miss Watson does not neglect to inquire after the reward for catching Jim. At first, Finn feels very relieved after having drawn up the letter, but he then begins to remember his time with Jim affectionately and thinks of all the ways Jim has been kind to him. He concludes that he cannot do this to Jim, so he tears up the letter and resolves to do everything he can to help Jim attain freedom.³¹⁶ In Finn's moral universe this is categorized as a decision for the bad and he takes it to be such. Nevertheless, it is *de re* a commitment to the good, or to Jim's humanity and so to the moral law.

As Kant would caution us, it is difficult to know Finn's true motivation. Given the risks Finn is now willing to take on Jim's behalf, this commitment appears to trump his earlier commitment to expediency; though it remains possible that his action is based on mere fellow feeling and not the moral law. If we were concerned about this, we could

³¹⁴ The men promise Finn money if he turns in any runaway slave (Mark Twain, *Adventures of Huckleberry Finn*, 85).

³¹⁵ *Ibid.*, 191-192.

³¹⁶ *Ibid.*, 192-193.

imagine Finn in an alternate situation that tests his commitment to the moral law, much in the same way Kant imagines the philanthropist of *GR I* in an alternate situation in order to test his will and reveal his motivation for action. For example, we might imagine Finn discovering around this time that Jim actually hates him and is merely putting on the guise of friendship for the sake of self-interest, much like the hypothetical character of Jack. If Finn nevertheless were to pursue the same course of action, not revealing Jim's whereabouts and instead, seeking his freedom, we could be confident Finn is acting on the moral law. (The fact that in the story Finn believes he is going to hell for Jim's sake suggests he is willing to endure a great amount of suffering in order to do what is right, making him out to be rather Kantian in the austere sense.³¹⁷) However, if Finn were at this point to betray Jim, this would suggest otherwise.

Finn is obviously an imperfect figure and far from an ideal type of the morally worthy agent. Of course, that is part of his charm and his purpose as a literary figure. Nevertheless, at this point we might say, provided we were confident he is appropriately motivated, that he acts with moral worth, albeit in a mitigated sense. Whereas in the earlier instance his stated resolution (of expedience) was self-interested and so immoral on a Kantian analysis, in this instance his commitment to do what is right is clear, even if he does not understand it as such. Finn is far from possessing ideal beliefs about morality, since his moral universe is deeply distorted. Nevertheless, this may be the intellectual best that can reasonably be expected of him, given not only his broader social setting but the particular network of relationships he finds himself in (including a violent and dysfunctional family setting). If it is

³¹⁷ Ibid. He conceives of this as a choice between heaven and hell, since he believes aiding Jim is tantamount to denying Miss Watson what is rightfully hers and this is punishable by damnation.

the best that can reasonably be expected of him, then given his de re commitment to the moral law we may reasonably attribute to him moral worth, albeit in a mitigated sense.

Full moral worth is reserved for those agents who are committed to principles that are grounded in the moral law and who, in addition to this, exercise a well-developed disposition of sympathy. To exercise a well-developed disposition of sympathy is to exercise good judgment in matters concerning morality and how it applies in the empirical world. This will require empirical knowledge of others and the circumstances of one's action, and so it will be necessary for the agent to pursue some form of empirical enquiry. It will also to some extent require right beliefs concerning morality. The idea of mitigated moral worth enables us to say that some agents merit praise for their commitment to the moral law and consequent good action, even when their beliefs and judgment are mistaken, provided these agents have good judgment to a standard that can reasonably be expected of them. (As mentioned, Eichmann and Shylock are not under consideration for this sort of ascription of moral worth, since we can grant that agents like them can reasonably be expected to believe other than they allege they did.) Though Kant did not recognize this distinction in his direct discussion of moral worth, it is consistent with the Kantian system because it is grounded in his theory of moral character. This theory of moral character is itself a systematic reconstruction of what is found in Kant's texts, but one that is faithful to the texts as far as that is possible. This theory of moral character is therefore attributable to Kant, as something he reasonably could have thought. It is also one that is interesting in its own right.

This complex notion of moral worth makes it possible for the moral appraisal of particular agents to come in degrees. It encourages us to take into consideration an individual's cultural, historical, and otherwise empirical environment when determining

whether she merits praise (or otherwise) as a moral agent. More importantly, it does not require us to abandon a fixed moral norm for appraisal, nor a strict definition of moral worth in the full sense: it enables us to retain a clear definition of the ideal moral agent we are to strive to become. Provided we understand the moral law correctly ourselves, it also enables us to be clearheaded about when a given agent is wrong, while recognizing that she may fundamentally still be a good person seeking to do what is right. This should encourage us to be humble, as we remain mindful of what we do not yet know about ourselves. Further, it encourages us to seek knowledge of ourselves, others, and that which is good, so that we may root out that which is bad in our own character as we further the good of others amidst the circumstances we find ourselves in.

4. The Good

Unless you strive after the virtues and practice them, you will never grow. For, as you know, anyone who fails to go forward begins to go back, and love, I believe, can never be content to stay for long where it is.

~ St Teresa of Avila³¹⁸

Closely related to the charge that Kant has a crude or incomplete account of education, emotion, and character more generally is the charge that he has a distorted or incomplete account of the good life. Chapters one through three provided a sustained reply to the former charge, arguing that Kant's moral psychology and theory of character in particular are much richer and contain a more persuasive account of these features than has otherwise been thought. The present chapter will respond to the latter charge. An account of the good life integrates factors of human life such as pleasure, personal relationships, motivation, and virtue into an ideal whole. These are all crucial concepts for moral character, so it is natural that a theory of moral character should be integrated with some account of the good life, even if only implicitly. If a theory of moral character fails to integrate with some standardly acceptable account of the good life, it is typically taken to have failed as a plausible theory of character, or at least to be that much less

³¹⁸ Theresa of Avila, *Interior Castle* trans. and ed. E. Allison Peers (New York: Dover, 1946), 164.

persuasive as a result. Kant has faced a number of objections along these lines: I will focus on several of these objections and respond to them in what follows.

One of the most prominent and, if correct, perhaps one of the most persuasive of these objections (which applies to Kantian ethics more generally) is the objection from moral demandingness. Broadly construed, it contends that in Kantian (along with utilitarian) ethics the standard for moral action puts too great a demand on the individual agent to be compatible with the enjoyment of a good life for that agent. Since it is reasonable to want to enjoy a good life, even if for non-moral reasons, we have good reason to reject Kantian ethics – so the argument goes. Michael Stocker and Susan Wolf each put forward poignant forms of this objection. Stocker is concerned that modern ethics, and Kantian ethics by implication, induces a fundamental split in an agent's personal psychology and prevents an agent from securing meaningful personal relationships, which are a deep source of value. Wolf is concerned with what she sees as the all-consuming nature of Kantian ethics and its inability to recognize non-moral value more generally as a part of the good life. I will look at each of these variations of the demandingness objection. I will argue that they both stem from a misunderstanding of Kant and that Kantian moral psychology, understood against the backdrop of his theory of moral character, has the resources to formulate a satisfying response to each. I will then look at Kant's concept of the highest good, which indicates a kind of ideal relation of happiness to morality, and respond to concerns that it is rigidly (hence undesirably) moralistic and that it has no practical import.

Moral Demandingness: Psychological Integration and Personal Relationships

Stocker's main concern is that living the theory of modern ethics – a label undoubtedly meant to include Kantian ethics – directly would make our lives fragmented and incoherent, because modern ethics is too narrowly focused on the notion of obligation.³¹⁹ He argues that modern ethics understands obligation in such a way that it makes action on the basis of standard human love impossible, since he appears to hold that action on the basis of obligation excludes the possibility that an agent also desire to do what she is under obligation to do. According to this picture, such an ethic would prevent an agent from engaging in healthy interpersonal relationships, because it would require her to do good out of a sense of obligation to the exclusion of love and desire. Stocker asks, suggestively, “[w]hat sort of life would people have who did their duties but never or rarely wanted to?”³²⁰ Clearly, it would not be a good life. He goes on to say that such an ethic of obligation would induce a split between what we would think of as an agent's healthy desires, for things such as friendship, and the moral reasons for which she must act. Stocker uses the term “moral schizophrenia” to name this split. On this picture, if an agent were to then choose to act morally, she would be prevented from securing healthy interpersonal relationships and psychological harmony or integrity. This would prevent her from securing a good life, since we understand these things to be necessary for a good life.

³¹⁹ Stocker directs this objection at modern ethical theory generally, by which I take him to mean both Kantian and utilitarian ethics, as these are the two dominant traditions of modern ethical theory and are usually the target of objections like his (Michael Stocker, “The Schizophrenia of Modern Ethical Theories,” 454-456).

³²⁰ *Ibid.*, 455.

This problem of moral schizophrenia has two parts, one internal and one external. The internal part concerns the agent's fractured personal psychology – the purported split between her healthy desires and justified moral reasons. The external part concerns the moral agent's purported inability to secure loving relationships with others. The internal part, which is at times referred to as alienation (from oneself and others), is easily dealt with for the Kantian: it does not apply to her. As I have argued throughout chapters one to three, commitment to doing one's duty – that is, commitment to the moral law – does not preclude the participation of accompanying emotions in dutiful action. More specifically, the nature of maxims and the Kantian account of moral character detailed in chapter three acknowledge that an individual will have a variety of goals and activities in life aside from the direct goal of undertaking to act on the basis of the moral law. What this account requires of an agent (on moral grounds) is to integrate these many goals and activities with the moral law. Successful integration would mean these goals and activities are neither themselves wrong, nor are being pursued in a way that is wrong. Plotting to murder someone would be an example of a wrong goal; stealing a violin so that one can give it to a friend as a gift would be an example of pursuing a right goal (of giving a gift) in a way that is wrong. Successful integration would also mean that self-perfection and the performance of benevolent acts would be among these goals, since they are required of the moral agent (though they need not consume an agent's life, a point about which I will say more below). But in summary – and importantly – this imperative of integration does not require that an individual either expunge feelings of love and affection, or necessarily overcome some negative feeling (for example, extreme apathy or antipathy) toward another person in order to act morally toward them.

The problem that Stocker believes he is seeing here would appear to stem from an incorrect (albeit to some extent forgivable) reading of Kant's account of moral worth – one that has acquired the mantle of a tradition through the past couple hundred years of philosophy. To use the example from chapter three, one may visit a friend in the hospital out of concern for this friend and desire to share her company, and this action will have moral worth, so long as whatever maxim it involves (say one of pleasure, in Kantian terms) is incorporated in a deeper maxim of respect for the moral law and the dignity of this friend.³²¹ In this way, there is no inconsistency between loving one's friend and doing one's duty (and in a way that merits moral worth). Further, as I have suggested above, to fulfill one's duty or act with moral worth does not in every instance require an act of self-overcoming, or a "citation for gallantry" in winning a difficult battle in the "war against evil", as Henson puts it.³²² To be sure, the moral agent will at times be required to undertake difficult acts of self-overcoming, given how deep the motive of self-love is rooted in the human constitution and how easily it can eclipse one's awareness of the needs of others (this is why Kant emphasizes the cultivation of sympathy, as described in chapter two). But such difficult acts do not have to be the norm, especially in relationships guided by human love at its best, since these sorts of relationships may require less active constraint and redirection of desire. As I have argued, Kant discusses these difficult sorts of acts in the *GR* primarily to make a more

³²¹ This was originally an example discussed by Stocker, of which mine is a variation. For Stocker's original discussion, see: *ibid.*, 462.

³²² Richard Henson, "What Kant Might Have Said: Moral Worth and the Over-determination of Dutiful Action," 42.

general point about the structure of moral action and to show how the moral motive operates in demanding cases.³²³

In other words, Kant's moral agent may be moved by what we might call the *de re* content of morality, that is, by particular people themselves, and is not required to in every case be moved by morality *de dicto* in some unhealthy or fetishistic way. Here I am following a fairly common use of the *de re/de dicto* distinction in contemporary moral philosophy, to refer to the concrete content of an obligation and the abstract (or token) concept of an obligation, respectively. For example, in the second case concerning Huck Finn (in which he resolves to help Jim attain freedom and so merits an attribution of mitigated moral worth) we could say he is moved by morality *de re* and not morality *de dicto*; that is, he is moved by the correct moral content of Jim's humanity, even though he is conceptually confused about what morality requires and acts against what he believes (the token concept of) morality requires. Of course, for an agent to merit an attribution of full moral worth she will be required to understand the contents of morality and how these contents apply to a particular situation (and she may still merit an attribution of mitigated moral worth if she fails to do this but acts on the right reasons *de re*). There is nothing unhealthy in this: to the contrary, it exhibits a greater degree of intellectual and personal integration.³²⁴ Morality on the Kantian picture does not necessitate a split in the moral agent's psychology, as Stocker suggests.

³²³ Allen Wood reminds us that Kant says in 4:397 that actions done from duty are difficult to distinguish from actions done from inclination, so presumably the standard moral agent wants to do both. For this reason (and others), it is a mistake to take the examples from *GR I* for a general account of what the motive of duty looks like. Kant does not require the moral agent to act on the thought of duty, rather than act directly on care for another person as an end in themselves. (Allen Wood, "The final form of Kant's practical philosophy" in *Kant's Metaphysics of Morals: Interpretive Essays* ed. Mark Timmons (Oxford: Oxford University Press, 2002), 16.)

³²⁴ Vanessa Carbonell makes a related point well when she says: "[t]here is a benign, perhaps even admirable, way of caring about the rightness of your actions *de dicto* – it manifests itself in reflective

The Kantian has a ready answer for the external part of Stocker's moral schizophrenia problem, too. This is the idea that the Kantian moral agent is unable to secure loving relationships with others, since this requires that one "care for the beloved and act for that person's sake as a final goal" and an ethic grounded in obligation does not allow for this.³²⁵ Most of the burden of this charge is dealt with in the above response concerning healthy desires and moral motivation: if doing one's duty is consistent with concern for another person or pleasure in one's relationship with them, one can enter into loving relationships without preventing the possibility of moral action or having to set aside one's moral commitments. Yet the worry might persist that there is still too great an emphasis on the concept of duty in the Kantian system and too little emphasis on personal relationships, which might still have the effect of eroding personal relationships, which is one of Stocker's main concerns. He states the point strongly when he says the "problem with these theories is [...] simply – or not so simply – with *people-as-valuable* [...] they are devoid of all people."³²⁶

Contrary to what Stocker's appraisal would have us think, Kantian ethics brings concern for other persons to the center of moral deliberation and action. As argued in chapter two, the two formulations of Kant's moral law most salient to moral psychology are the formula of universal law and the formula of humanity; together they form the basis of moral value for Kantian ethics and ground the notion of moral obligation. The formula of humanity holds that all rational beings have dignity and one therefore ought

people who are concerned about how their actions can be classified in the abstract, while being primarily motivated by their moral projects *themselves*" (Vanessa Carbonell, "What Moral Saints Look Like" in *Canadian Journal of Philosophy* 39:3 (2009): 395).

³²⁵ Michael Stocker, "The Schizophrenia of Modern Ethical Theories," 456.

³²⁶ *Ibid.*, 459-460.

always to treat each one as an end in themselves and never as a mere means.³²⁷ In the *MM*, Kant argues that this duty entails the pursuit of both self-perfection and the happiness of others, clarifying its demands.³²⁸ This straightforwardly affirms the value (and well-being) of persons: it is difficult to see how this would permit viewing and treating others “as externally replaceable, as mere instruments or repositories of general and non-specific value”.³²⁹ Perhaps Stocker would reply that, though this shows that persons have value in the Kantian system, the system requires that an agent not be motivated by the value of any person in particular, but by the thought that she has a duty toward them. This is not the case, however. The sorrowful philanthropist of *GR I* may be motivated directly by the humanity (dignity) of those whom he helps, rather than the naked thought that he has a duty to do so, even if in this limit case his apprehension of their value is predominantly intellectual, rather than emotionally felt.³³⁰ The account of maxims and of friendship I presented in chapter three provides a picture of how motivation can work in more standard cases, without requiring that one be motivated by the thought of duty.

It is worth clarifying that Kantian ethics does not, however, permit an agent to favor friends and loved ones by granting them a privileged *moral* status. Kant does allow that without violating the moral law one may privilege those closest to oneself by demonstrating a greater degree of benevolence toward them than to others, since one

³²⁷ 4:429.

³²⁸ 6:382-386.

³²⁹ Michael Stocker, “The Schizophrenia of Modern Ethical Theories,” 460.

³³⁰ See Allen Wood: “Kant identifies the ‘motive’ (*Bewegungsgrund*) proper to morality with the dignity of humanity as an end in itself. This means that, according to Kant’s theory, the sorrowful man who acts from duty is not moved merely by the stony thought ‘it is my duty to help’. He acts instead out of a recognition that those in need of his help are ends in themselves. Their *dignity* gives him a reason to care about them and gives them a claim on his help, whether or not he feels like helping them” (Allen Wood, “The final form of Kant’s practical philosophy,” 16).

cannot reasonably be expected to contribute equally to the well-being of all (if one is going to make a contribution of any significance). In this sense, one may have obligations toward those near to oneself that one does not have to others.³³¹ If this were all that Stocker's argument implied, Kant could agree. However, if Stocker's argument implies that one ought to grant one's loved ones a privileged moral status that would elevate their interests above the interests of others in a way that conflicts with the moral law, then the Kantian must disagree. For the moral law applies equally and unconditionally to all rational beings: it acts as a constraint on all actions and relationships an agent might engage in and it admits of no exceptions. Stocker might reply that such an unconditional commitment to moral principle is unhealthy or moralistic, but this reply would be unfair and incorrect, for such a moral constraint is neither unreasonable nor exceptional to standard intuition. On the contrary, it is necessary in order to guard against, for example, easily imaginable cases of nepotism and otherwise inappropriate favoritism that might be anticipated when personal and professional relationships overlap. In these and other cases, a moral constraint that requires people to be treated equally with respect, as ends in themselves and not mere means, seems a reasonable principle with which to regulate conduct.

That such a moral constraint on personal relationships is sensible and consistent with standard intuition is illustrated well in the fictional example of the young doctor Tertius Lydgate's changed love for the actress Laure in George Eliot's *Middlemarch*. Lydgate was deeply in love in Laure when she stabbed her husband, a fellow actor,

³³¹ "For in wishing I can be *equally* benevolent to everyone, whereas in acting I can, without violating the universality of the maxim, vary the degree greatly in accordance with the different objects of my love (one of whom concerns me more closely than another). (6:452.) It seems the closeness he has in mind in this passage could lie in either personal relation or geographical proximity. Further, Kant will of course recognize obligations to one's children and one's spouse (through contract, cf. 6:277-282).

apparently by accident during the performance of a murder scene in a theatrical play. Lydgate defends her innocence arduously, given his deep love for her and the lack of conclusive evidence to indicate her guilt. However, his love for her changes when she confesses to him that she had in fact meant to kill her husband and so had committed murder. It is worth reading their exchange to observe the way this moral requirement presents itself in the context of what would otherwise be a standard human relationship.

‘I know, I know,’ said Lydgate, deprecatingly. ‘It was a fatal accident – a dreadful stroke of calamity that bound me to you the more.’

Again Laure paused a little and then said, slowly, ‘*I meant to do it.*’ Lydgate, strong man as he was, turned pale and trembled: moments seemed to pass before he rose and stood at a distance from her. ‘There was a secret, then,’ he said at last, even vehemently. ‘He was brutal to you: you hated him.’

‘No! He wearied me; he was too fond: he would live in Paris, and not in my country; that was not agreeable to me.’

‘Great God!’ said Lydgate, in a groan of horror. ‘And you planned to murder him?’

‘I did not plan: it came to me in the play – *I meant to do it.*’

Lydgate stood mute, and unconsciously pressed his hat on while he looked at her. He saw this woman – the first to whom he had given his young adoration – amid the throng of stupid criminals.³³²

Evidently Lydgate’s commitment to Laure is not total. The evaporation of his love and the withdrawal of his commitment evince that both are conditioned by morality. I suggest this would strike most readers as the appropriate reaction for him to have. In a way similar to this, the moral law is meant to act as a condition on relationships, even personal relationships of deep love and affection. This is not a point against Kantian ethics; rather, it is a point in its favor.

In short, Kantian ethics is not dehumanizing, nor is it inimical to human relationships and psychological health. Once these objections have been answered, it remains difficult to anticipate what it is in a moral theory that Stocker might be satisfied

³³² George Eliot, *Middlemarch*, 143-144.

with. He suggests at one point that a “full scale philosophical anthropology” showing how personal relationships are possible, relate to larger structures of human life, and are constitutive of a good life is required to avoid dehumanization.”³³³ This would be a good and helpful thing, though it would not stand in for a justificatory framework for action. One gets the idea that Stocker is asking moral philosophy to do work nowadays more effectively done by disciplines like anthropology and psychology, or that he’s at least contesting the definition and purpose of moral philosophy. He certainly intends to do some of the latter, as it appears in its modern guise, but it is worth considering how useful an exercise moral philosophy would be if it did not in the first instance seek to understand and establish the grounds, objects, and limits of morally justified action. In the end it may be that moral philosophy is, to some extent, limited by this task – after all, it cannot be expected to do the work that a devoted anthropologist or behavioral psychologist could do better – however, its successful performance of this task need not entail the sort of maladies Stocker believes it does, as I have argued above.³³⁴ In general, it would seem sensible to not try to live any moral theory directly; and one might expect this point to generalize to any intellectual theory.³³⁵

³³³ Michael Stocker, “The Schizophrenia of Modern Ethical Theories,” 460.

³³⁴ With respect to a justificatory framework for action, Kant may be taken to have provided a complete theory, or one that purports to be so. As argued throughout this thesis, Kant provides a much richer account of moral psychology and anthropology than he is often credited with. However, of course, he does not say everything there is to say on the matter. In this respect, he provides what Stocker refers to as a “partial” rather than a “whole” theory and I imagine he would have seen it this way himself. In this respect, it does not make sense to say he considers duty “the whole, or at least the only important part, of ethics” (ibid., 464).

³³⁵ Stocker: “I have not argued that if you have a successful love relationship, friendship, . . . , then you will be unable to achieve the justifications, goals, goods posited by those theories. You can achieve them, but not by trying to live the theory directly” (ibid., 461).

Moral Demandingness: Happiness, Saints, and Moral Constraints

Susan Wolf's worry is similar to Stocker's, though she is more concerned with what she takes to be the all-consuming nature of Kantian duty and its tendency to obscure the value of non-moral goods in the broader sense, specifically with respect to the duty of benevolence (or beneficence). She suggests we have unlimited opportunities to do moral good and that Kantian ethics requires we act on all, or in any case, too many of these opportunities; so if we follow Kantian ethics, the need to act morally will dominate our lives. She suggests this state of affairs is undesirable and so we are warranted in limiting our commitment to morality on non-moral grounds.³³⁶

Indeed, some of what Kant says about duties to others in the *GR* lends itself to misinterpretation. For example, he says that there cannot be a "positive agreement with humanity as an end in itself unless everyone also tries, *as far as he can*, to further the ends of others"; he says also, "the ends of a subject who is an end in itself must *as far as possible* be also my ends".³³⁷ However, it is important to not read too far into comments like this from the *GR*. Wolf wants to read passages like this very determinately, in a way they are not meant and do not need to be read. In the Kantian system, benevolence is generally represented as a duty that affords the agent a significant degree of latitude in acting to fulfill it. In fact, it would seem there is no specific limit placed on benevolence for the reason that the agent is free to determine this limit as she sees and as she chooses she is reasonably able to, provided she put in a good effort. Kant's expectations are still serious – a life of indulgent luxury and only minimal concern for others would not do –

³³⁶ Susan Wolf, "Moral Saints," 435-436.

³³⁷ 4:430, emphases mine.

however, they are not fanatical. In his first mention of this duty in the *GR*, Kant only emphasizes that it is impossible to will that one *never* help those in need:

for whom things are going well while he sees that others (whom he could very well help) have to contend with great hardships, thinks: what is it to me? let each be as happy as heaven wills or as he can make himself; I shall take nothing from him nor even envy him; only I do not care to contribute anything to his welfare or to his assistance in need! [...] But although it is possible that a universal law of nature could very well subsist in accordance with such a maxim, it is still impossible to **will** that such a principle hold everywhere as a law of nature. For, a will that decided this would conflict with itself, since many cases could occur in which one would need the love and sympathy of others and in which, by such a law of nature arisen from his own will, he would rob himself of all hope of the assistance he wishes for himself.³³⁸

The contradiction in the will which grounds the justification for this duty is produced only when the agent wills this negative maxim (to never help those in need). So, the logic only entails that she not adopt this maxim; it does not entail that she adopt the maxim to always help every person in need. Of course, the duty of benevolence cannot be fulfilled by merely refraining from resolving to never help – to say as much would be confused, since benevolence implies action – and Kant is keen that one put a good degree of effort into acting for the benefit of others. Nevertheless, the logic that he employs warrants scaling back the purported demandingness of this duty. With this in mind, we might reread the sentence cited further above by placing the emphasis on the broader human community and the need for every agent to make some contribution: “[there is] not a positive agreement with humanity as an end in itself unless *everyone* also tries, as far as he can, to further the ends of others.”³³⁹ The idea Kant might have in mind generally is that, were everyone to do their part and contribute to the needs of others as they are reasonably able to, these needs would be met. So we might adjust this sentence by adding

³³⁸ 4:423.

³³⁹ 4:430, emphasis mine.

a ‘reasonable expectations’ clause to what Kant originally said, saying: there is not a positive agreement with humanity as an end in itself unless *everyone* also tries, as far as can *reasonably be expected*, to further the ends of others. Granted, this is slightly creative; but it seems to better capture Kant’s broad position concerning this duty.

Beyond this passage, his general presentation of this duty does not suggest that one ought maximally to fulfill one’s capacity to act beneficently. In his discussion of the duty of beneficence (*Wohltätigkeit*) in the *MM*, for example, he emphasizes that this duty has limits.

Providing for oneself to the extent necessary just to find satisfaction in living [...] belongs among duties to oneself. The contrary of this is depriving oneself (slavishly) of what is essential to the cheerful enjoyment of life, by *avarice*, or depriving oneself (fanatically) of enjoyment of the pleasures of life by exaggerated *discipline* of one’s natural inclinations. Both of these are opposed to a human being’s duty to himself.³⁴⁰

Here, the pursuit of beneficence may be limited by the conditions an agent needs to secure in order to enjoy the pleasures of life and to lead a satisfying life overall. This is grounded in the more general point that the moral law recognizes a duty to oneself in the formula of ends in themselves, meaning that one cannot make of oneself a mere means, including when this is a means to the happiness of others.³⁴¹ This is strikingly resonant with what Wolf refers to as the “not-so-ethical fact that we have sound, compelling, and not particularly selfish reasons to choose not to devote ourselves univocally to realizing this potential or to taking up this opportunity” (to be beneficent).³⁴² Kant presents this limit in terms of a duty to the self, so, unlike Wolf, he is appealing to another moral fact (duty) and not a non-moral fact in order to secure this limit. Though Kant appeals to one

³⁴⁰ 6:452.

³⁴¹ “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, bold emphasis mine).

³⁴² Susan Wolf, “Moral Saints,” 435.

duty in order to limit another, his doing so demonstrates that he grasps that we have strong and indeed justifying reasons to pursue projects and pleasures that will, because of the time they take, limit our performance of beneficent acts.³⁴³ Given that we are each free to determine what our particular projects and pleasures will be, the position Kant leaves us in is similar to the position Wolf desires – that is, one that allows ample room for non-moral projects in addition to reasonable latitude in choosing what those projects will be.³⁴⁴

This touches on a similar objection raised by Wolf, against the idea of morality as “an ever present filter through which all other values must pass”; to quote her in full: “[t]he role morality plays in the development of our characters and the shape of our

³⁴³ Bernard Williams, too, worries that morality can come to dominate a life to the point where either all non-moral goals and activities are squeezed out, or one is made constantly to feel guilty for pursuing them. With this concern in mind, he claims that duties to oneself are fraudulent and a manifestation of an illicit obsession with morality. (Bernard Williams, *Ethics and the Limits of Philosophy* (London: Routledge, 2006), 181-182.) However, for Kant, the purpose of duties to oneself is largely to protect the self from being harmfully dominated by another’s interests; it is not as though they have to constitute the primary content of day-to-day actions alongside duties to others and to the exclusion of other projects and activities. Instead, pursuing one’s own happiness as one defines it for oneself is perfectly consistent with being a moral agent: it is something we are expected to do. The difference is that there is no duty to do this, since Kant assumes we all do it anyway (cf. 6:387). So long as they do not conflict with the moral law and do include appropriate moral items among the greater set of which they are a part, we are free to choose a wide manner of activities and goals we wish to pursue. This does not evince the sort of obsession with morality or domination of it over one’s life that Williams is concerned to avoid. On this picture, a duty to oneself would then be activated, as it were, when an agent comes under threat of being instrumentalized by another agent or by oneself. Of course, Kantian duties are grounded in the moral law, which holds that one ought always to treat “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, italics redacted). This seems a sensible application of morality, or use of the concept of obligation, and not an obsessive one, insofar as one is obligated to not instrumentalize others or oneself. And as mentioned above, the moral law requires that one never instrumentalize oneself for the sake of performing beneficent acts toward others. In this way, Kantian ethics builds appropriate limits into morality itself, strengthening the case that its conception of morality is not excessively demanding, particularly in the way that Wolf claims; further, it mitigates against the alleged need for specifically non-moral constraints to be placed on it. This could be a way of filling in what Kant means when he says that only the concept of duty can entail constraint: “[w]hat everyone wants unavoidably, of his own accord, does not come under the concept of *duty*, which is *constraint* to an end adopted reluctantly” (6:386). (When he refers here to adopting an end “reluctantly” (*ungern*), I take him to be emphasizing only that this end is not one adopted out of inclination; and that it remains an end the agent wills, or is wholly motivated to do.)

³⁴⁴ Wolf further discusses the ‘one thought too many’ objection originally raised by Williams (Susan Wolf, “Moral Saints,” 431), to which I responded in my discussion of friendship in chapter three.

practical deliberations need be neither that of a universal medium into which all other values must be translated nor that of an ever-present filter through which all other values must pass.”³⁴⁵ The story of moral formation outlined in chapter one encourages the use of inclinations (for example, the “inclinations to be honored and loved” which are “aids to morality”) to form a child prior to her grasp of morality and prior to a decision to ground her character in morality along the lines Kant describes.³⁴⁶ As I argued, these inclinations may continue to be of use to her throughout her adult life. Further, action that proceeds from inclination may have moral worth, provided the moral law would restrain the agent if that inclination were to lead her to act contrary to the moral law. So, on the Kantian picture, morality does not function in the oppressive way Wolf suggests it does, in either the development of character or standard practical deliberation. Kant does not require that all other values be translated into moral value; nor does he require that one act on the basis of moral values recognized as such in every case, even in cases that merit attributions of moral worth. It is standard on the Kantian picture that a moral agent may initially be moved by inclinations, rather than the thought of morality.

Kant does require that morality act as a constant constraint against action that violates the moral law and in that way act as an ever-present filter. However, since this does not require constant, explicit attention to the moral law in practical deliberation and moments of action, it is difficult to see what is objectionable in this requirement. Further, I suggest it would not be acceptable for an agent to pursue even an otherwise admirable life-project at the expense of morality, namely at the expense of treating other people as means and not ends in themselves. Avoiding cases in which this would happen is the

³⁴⁵ Ibid., 438.

³⁴⁶ 9:482.

purpose of the moral law as constraint, or filter. Negative examples of this could include stealing a violin to give as a gift to a friend who has resolved to excel at the violin (and shows promise at it, etc.); it might also include, on a greater social scale, producing great works of art while depending on institutions of conquest and slavery to provide the conditions under which the creation of this art is possible. (Of course, ethical analysis of cases following the latter example will be complex, but the idea is straightforward enough.)

This brings me to disagree with Williams' analysis of his fictionally elaborated Gauguin, who neglects the claims of others in order to pursue a life devoted to realizing his gifts as a painter.³⁴⁷ Williams uses this example to argue that the standard of assessment for such a project can only be retrospective (that is, its justification can only be determined at a later point in time than when the decision to commence the project is made) and to therefore highlight the role that luck plays in the assessment of one's life projects. The idea is that if Gauguin acted according to the demands of morality in the first instance, he would not have gone on to realize his gifts as a painter and the world would be worse off for not containing his paintings; therefore, we should be glad he did not act according to the demands of morality. More generally, Williams argues that this should lead us to accept that moral values are not unquestionably supreme and that the limitation of moral values is necessary even for morality itself (insofar as morality is concerned with the sort of retrospective justifications Williams points to in the case of Gauguin).³⁴⁸ I wish only to say that while one may be free to admire Gauguin's work, an agent in Gauguin's position ought not to neglect certain claims of others, namely those

³⁴⁷ Bernard Williams, *Moral Luck*, 23.

³⁴⁸ *Ibid.*, 37-38.

claims that are grounded in the moral law, even if that would make the production of his (arguably) excellent work impossible. Williams does not say much about what his fictional Gauguin's neglect entails, but he allows the consequence of his neglect to be "grim".³⁴⁹ It is common knowledge that the real Gauguin treated his wife terribly and had severely mistreated his family before his wife pushed him out of the home. It is also known that while in Tahiti, the real Gauguin took a thirteen- and two fourteen-year-old wives to whom he knowingly gave syphilis. If these are to be necessary consequences of a pursuit of artistic excellence, I suggest such a pursuit is not worth the cost. In the case of Gauguin, if this meant the world were deprived of his artistic contributions, I suggest we should accept that deprivation for the sake of those whose lives were affected by his actions (of course, the significance of the loss in artistic value is a matter of taste).

Williams and others do not like the idea of an ever-present moral filter for action. But, for my part, it is difficult to see how removing a filter like this would be desirable, or at the risk of circularity, permissible – it seems that we here run into a kind of justifying intuition. To return to Wolf, her objection misses the mark on this point. Requiring that one's projects, goals, and activities etc. obtain a kind of moral permission is just to ensure that they do not violate morality; beyond that one is free to pursue those projects and activities that one decides will bring happiness to one's life, provided one also contributes to the happiness of others and pursues self-perfection, as noted above. In this way, because happiness is a simple concept for Kant, it is a liberal one.

³⁴⁹ Ibid., 23.

The Highest Good: Moralism and Practical Import

Discussion of Kant's understanding of happiness is incomplete without reference to his notion of the highest good, first mentioned in the *CPR* and laid out in more detail in the subsequent *CPrR*, *REL*, and *CJ*.³⁵⁰ In the *CPrR*, he explains that the highest good consists in a world where "happiness [is] distributed in exact proportion to morality".³⁵¹ In what follows, I will address two objections concerning his account of the highest good; first, that it is rigidly (hence undesirably) moralistic and second, that it has little or no practical import, since he ultimately appeals to a divine ruler and an afterlife in order to bring it about. I will conclude by examining its effect on moral motivation as discussed in the *CJ* and will argue it may perform a positive function for moral character consistent with the account of moral worth presented throughout this thesis.

Kant maintains that this highest good is the "whole and complete good as the object of the faculty of desire of rational finite beings".³⁵² At this point, he has argued extensively that the moral law applies unconditionally to all action (as a categorical imperative). So it makes sense to think of morally worthy action (virtue) as a "supreme" good and a condition under which any further good must be realized, if that further good is to count as moral. In this way, virtue is the first constitutive part of the highest good.

In chapter two, we observed that in the *MM* there are two ends that are also duties, both of which follow from the moral law; those duties are one's own perfection and the happiness of others.³⁵³ Similarly, in the *CPrR*, Kant argues that practical reason requires

³⁵⁰ For discussion of the highest good in the *CPR*, see A805/B833-A819/B847

³⁵¹ 5:110.

³⁵² 5:110.

³⁵³ 6:385-388.

the attainment of happiness as a part of the highest good, though he is careful to qualify that the attainment of happiness in the highest good is conditional on virtue. He explains that morality (here, virtue or moral worth) is to be the ground and condition of happiness, so that happiness is distributed in proportion to moral worth.³⁵⁴ In this way, happiness is the second constitutive part of the highest good: even though it is conditional on virtue and in that way is not the supreme good, it is still necessary to obtain a complete account of the good.³⁵⁵ So the moral law requires the satisfaction of happiness in the attainment of the highest good, as well as in the fulfillment of duties to others as we observed earlier in the *MM*. Hence, the satisfaction of sensible nature is incorporated in the broader satisfaction of rational nature, so that the general “faculty of desire of rational finite beings” may be satisfied.³⁵⁶ Put another way, in the highest good Kant provides a higher order concept under which all aspects of human nature are coordinated and their respective desires properly satisfied. This is what practical reason and the moral law direct us toward, and it is what Kant has in mind when he describes the highest good as the “unconditioned totality of the object of pure practical reason”.³⁵⁷ In this way, the highest good consists in happiness distributed in proportion to moral worth. Further, because (i) we cannot expect happiness necessarily to follow from virtuous action, since

³⁵⁴ As Kant says, the highest good “means the whole, the complete good, in which, however, virtue as the condition is always the supreme good, since it has no further condition above it, whereas happiness is something that, though always pleasant to the possessor of it, is not of itself absolutely and in all respects good but always presupposes morally lawful conduct as its condition” (5:110-111).

³⁵⁵ “The highest can mean either the supreme (*supremum*) or the complete (*consummatum*). The first is that condition which is itself unconditioned, that is, not subordinate to any other (*originarium*); the second is that whole which is not part of a still greater whole of the same kind (*perfectissimum*). That *virtue* (as worthiness to be happy) is the *supreme condition* of whatever can even seem to us desirable and hence of all our pursuit of happiness and that it is therefore the *supreme* good has been proved in the *Analytic*. But it is not yet, on that account, the whole and complete good as the object of the faculty of desire of rational finite beings; for this, *happiness* is also required” (5:110).

³⁵⁶ *Ibid.*

³⁵⁷ 5:108. In the *REL*, he eases the abstraction in his language to describe the highest good as meeting a “natural need”, going on to say that “the human being evinces the need, effected in him by morality, of adding to the thought of his duties an ultimate end as well, as their consequence” (6:5-6).

it manifestly does not in this world, (ii) the highest good not only fulfills a rational desire but represents a duty we ought necessarily to fulfill, and (iii) Kant has argued (in the *CPR*) that we are already warranted in thinking of ourselves as belonging to a noumenal world of pure understanding, he argues that (iv) we are therefore also warranted in postulating the existence of a being that can bring about the highest good in the noumenal world, that is, God.³⁵⁸

Prima facie, some of what Kant says here does have a rather moralistic (in the pejorative sense) ring to it. The notion of “happiness distributed in *exact* proportion to morality” might be taken to suggest that happiness only be given to those who earn it with good behavior, like candy given to an obedient child.³⁵⁹ Moreover, insofar as we have a duty to promote the highest good, it might seem as though we have a duty to decrease the happiness of those who are less virtuous, as well as to increase the happiness of those who are more virtuous. Though this is not an entirely unreasonable thought, it is not an entirely attractive one: decreasing the happiness of others is likely not a task most of us would be keen to undertake.³⁶⁰ However, the justification Kant provides for the immortality of the soul makes clear that pursuit of the highest good does not commit one to such a negative distribution of happiness. He points out that an agent whose will is determined by the moral law will necessarily have as a goal the attainment of holiness (the complete conformity of her dispositions with the moral law). Relying on the principle that ought implies can, Kant argues that the attainment of holiness must

³⁵⁸ 5:125.

³⁵⁹ 5:110, emphasis mine.

³⁶⁰ It is not entirely unreasonable because it generally does not seem right when someone immensely evil is allowed to live a happy life. This would seem to acknowledge as valid a certain retributivist sentiment. It should also be noted that this is something Kant in fact is willing to undertake in the case of criminals. In defending the right to punish he states quite simply: “[t]he *right to punish* is the right a ruler has against a subject to inflict pain upon him because of his having committed a crime” (6:331).

therefore be possible for her and since holiness “can only be found in an *endless progress* toward that complete conformity”, it requires an endless duration of time, hence the immortality of the soul.³⁶¹ The highest good would entail the highest degree of virtue and the proportionate degree of happiness – the highest degree of happiness – for all rational beings. So the idea is that every agent would eventually attain perfect virtue (holiness, or the complete conformity of her dispositions with the moral law) and maximal happiness. On this picture, there are no vicious agents and so there is no negative distribution of happiness in the highest good: everyone is meant to attain happiness.

It remains somewhat puzzling what is to happen to the apparently vicious agents we observe in this world, if they were to choose not to undertake the journey of self-perfection. Kant wrestles with this in the *EA* and seems to settle on the idea that we cannot know, so we must leave it in the hands of divine wisdom.³⁶² He explains that we are not in a position to know either our own character or that of others “through and through”, so any personal judgment or dogmatic stance on the matter should be met with skepticism.³⁶³ It seems the hope is that everyone will choose to pursue self-perfection and, in any case, that is what we each have a duty to do. Further, with regard to the matter of pronouncing judgment on others, toward the end of the *CPrR*’s Dialectic of Pure Practical Reason he reminds us that a person’s moral worth may only be discerned in the “eyes of supreme wisdom”.³⁶⁴ He maintains this point throughout his practical philosophy, pointing out in the *REL* that

through experience we can indeed notice unlawful actions, and also notice (at least within ourselves) that they are consciously contrary to law. But

³⁶¹ 5:122.

³⁶² 8:337.

³⁶³ 8:829-830.

³⁶⁴ 5:147.

we cannot observe maxims, we cannot do so unproblematically even within ourselves; hence the judgment that an agent is an evil human being cannot reliably be based on experience.³⁶⁵

It makes sense to hold that we have a good enough grip on our own maxims to be able to think and act rationally, as well as to engage in the project of self-improvement (self-perfection, in Kantian terms), attending to and rooting out those maxims that are ill-formed or evil. This is not inconsistent with the idea that our insight into our own maxims and the maxims of others is limited and prone to some degree of error. In light of this point, too, it would be unwise to distribute happiness negatively on moral grounds, that is, to undertake to reduce the happiness of the vicious or less virtuous, except perhaps in the case of punishment (a topic the discussion of which is beyond the scope of this thesis). It is therefore reasonable to interpret this aspect of the duty to strive for the highest good primarily positively, as a requirement to increase the happiness of the virtuous.

Is this a reasonable thing to want to do? Insofar as we find it difficult to see a genuinely good person suffer and wish for them to enjoy a correspondingly good life, this does seem reasonable. At times, we find ourselves saying things along the lines: ‘so and so is a good person, she really *deserves* to be happy.’ Allen Wood notes: “[t]he happiness of such a deserving person is a good for morality, something which morality commands us to pursue. On the other hand, the happiness of a ruthless embezzler who, ignoring his obligations to his family and community, steals a large sum of money and lives comfortably in Rio de Janeiro is not a moral good at all.”³⁶⁶ Discussion of punishment and right aside, we do not need to agree to seek the unhappiness of the embezzler to see

³⁶⁵ 6:20.

³⁶⁶ Allen Wood, *Kant's Moral Religion* (Ithaca: Cornell University Press, 1970), 82-83.

that his happiness is not a good for morality. The idea seems to be that if we wish to bring morality to bear on human happiness as an object for moral pursuit, the two concepts would relate as they do in the highest good. This is not the only way happiness may be pursued within the scope of the moral life; beyond our duty to strive for the highest good, we retain the sort of latitude I have described above to pursue our own happiness and contribute to the happiness of others as we choose. Happiness apart from the highest good is still a form of well-being and a natural good, though not an object of moral striving.³⁶⁷ Understood this way, the idea of the highest good and our duty to pursue it sits closer to standard intuitions and is not excessively or otherwise undesirably (pejoratively) moralistic.

This does leave open the question of whether Kant's conception of the afterlife in particular, as a place in which the dominant activity seems to be the purification of one's moral disposition and the making of oneself into a more virtuous being still remains undesirably moralistic. It seems this could be the case – Kant's description of the afterlife is rather solemn and bare, and the God of practical reason comes off as a petty albeit powerful taskmaster. However, if such a picture is true to what he is suggesting, we are not obliged to follow Kant in accepting it. For it is generally not clear why it should be necessary to attain the highest good rather than simply to pursue it; so it is not very clear why it is necessary to postulate the existence of God, since we could still have a duty to pursue the highest good in our earthly life in the hope that future generations will further approximate its attainment. In other words, it could be that we are meant to approximate the highest good as an earthly species rather than as individuals in an afterlife, along the lines Kant gestures toward in other works, such as *TPP*. There are big questions here

³⁶⁷ Ibid., 83-84.

concerning the role and viability of the concept of God in Kant's practical philosophy, the answers to which are not clear; however, for the purposes of this thesis it is sufficient to point out that there is an obligation to pursue the highest good, and then to explain what that pursuit looks like and the fact that all any individual can do is set out to do just that.

Notably, this does not imply that the concept of the highest good has no practical import, as Lewis White Beck has suggested.³⁶⁸ In the *REL*, Kant tells us the following:

Assume a human being who honors the moral law, and who allows himself to think (as he can hardly avoid doing) what sort of world he would *create*, were this in his power, under the guidance of practical reason – a world within which, moreover, he would place himself as a member. Now, not only would he choose a world precisely as the moral idea of the highest good requires, if the choice were entrusted to him alone, but he would also will the very existence of [such] a world, since the moral law wills that the highest good possible through us be actualized[.]³⁶⁹

The duty to pursue the highest good entails the duties of self-perfection and beneficence, since the duty to pursue the highest good proceeds from morality itself and morality entails these duties.³⁷⁰ So in this way, the duty to pursue the highest good just is the duty to contribute to these practical, moral ends. Further, a world in which everyone acted on the moral law and lived thoroughly happy lives would be satisfying from the point of view of morality, as I have indicated above, and Kant is suggesting that we have a duty to do what we can to contribute to a distribution like this. In the *CPrR*, Kant uses the language of striving to describe this duty: “in the practical task of pure reason, that is, in the necessary pursuit of the highest good, such a connection [of happiness proportioned

³⁶⁸ “The truth of the matter is that the concept of the highest good is not a practical concept at all, but a dialectical Ideal of reason. It is not important in Kant's philosophy for any practical consequences it might have, for it has none except those drawn from the concept of *bonum supremum*.” Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason* (Chicago: University of Chicago Press, 1960), 245.

³⁶⁹ 6:5-6.

³⁷⁰ “Yet an end proceeds from morality just the same; for it cannot possibly be a matter of indifference to reason to answer the question, *What is then the result of this right conduct of ours?* nor to what we are to direct our doings and nondoings” (6:5).

to virtue] is postulated as necessary: we *ought* to strive to promote the highest good.”³⁷¹

The language of striving indicates he means this to be a wide duty, since the notion of striving makes less sense with respect to avoiding violations of narrow duty.³⁷² Further, the highest good is an end that is also a duty and Kant describes the two ends that are also duties in the *MM* as wide duties; so it is consistent with his broader position that the duty to promote the highest good is a wide duty.³⁷³ Understanding the highest good as a wide duty helps to ease its demandingness as an ideal. Though Kant is keen that the agent employ a good degree of effort in fulfilling her wide duties, we are nevertheless warranted in scaling back their demand to a reasonable level, as I have argued above.

It is somewhat more difficult to think of ways an agent might make progress in fulfilling this duty than, say, the duty of beneficence. But it is not impossible. For instance, it could involve a general reorientation in how we value work and reward it with remuneration or social status. An example could be providing better remuneration for those who work in certain caregiving professions that require virtues like sympathy and patience and might otherwise not pay well enough to provide for a comfortable life.

³⁷¹ 5:125.

³⁷² Unless that meant, for example, that one ought to strive after a wide duty like self-perfection or the cultivation of sympathy in order to increase the chance that one not violate a narrow duty in future instances. In that case, fulfillment of narrow duty would be, as it were, embedded in fulfillment of wide duty.

³⁷³ “This proposition follows from the preceding one; for if the law can prescribe only the maxim of actions, not actions themselves, this is a sign that it leaves a playroom (*latitude*) for free choice in following (complying with) the law, that is, that the law cannot specify precisely in what way one is to act and how much one is to do by the action for an end that is also a duty” (6:390). Of course, the two ends that are also duties (from the *MM*) are one’s own perfection and the happiness of others. To take this position on the highest good as wide duty is to affirm Wood’s stance against John Silber. Silber claims: the “moral law must pose the *attainment* of the highest good as the necessary object of the will in order to carry out its function as the *ratio cognoscendi* of freedom.” (John Silber, “Kant’s Conception of the Highest Good as Immanent and Transcendent” in *The Philosophical Review* 68:4 (1959): 491.) Wood responds along the lines I have above: “[t]here is a distinction for Kant between what the law commands us to *do* (to accomplish), and what it commands us to *seek* (to make our end). If the law commanded us to *attain* the highest good, then the highest good would be a ‘narrow’ or ‘perfect’ duty, something we were obligated to *accomplish*. But Kant points out that both components of the highest good involve ‘imperfect’ or ‘wide’ duties, duties to act from a certain maxim, to *adopt* a certain end, but not necessarily to achieve or attain that end.” (Allen Wood, *Kant’s Moral Religion*, 94.)

Though this may sound radical, it need not be and could take place in the smaller communities one finds oneself in, such as a college or an office. In the context of personal relationships, one might out of gratitude direct more of one's time and attention to the people who have given the most to them in life (perhaps one's parents, grandparents, or other loved ones). Perhaps, interestingly, the duty to promote the highest good could provide a way of focusing our more general duty to contribute to the happiness of others, prioritizing how we do that in line with the above suggestions. In any case, Kant recognizes a wide duty to promote the highest good and this duty appears to have realistic practical import. So, while Beck is correct that the moral law does not require the highest good to explain its binding nature for rational beings, he is incorrect to say it "is not a practical concept at all".³⁷⁴

Finally, the *CJ* indicates that the idea of the highest good might also provide a kind of motivational supplement for the moral agent, at least if understood in a certain way. Here, Kant discusses an example of a righteous atheist like Spinoza who strives "unselfishly to establish the good" required of him by the moral law.³⁷⁵ He suggests that over time such a person will inevitably be worn down by the "purposeless chaos of matter". In other words, they will be worn down by a world – namely, the one we live in – which provides no reward for moral action and allows good people to suffer greatly; they will experience a kind of depression and pointlessness in pursuing a seemingly Sisyphean task. In saying this, Kant might seem to be suggesting that an agent in this intellectual position would suffer a weakened ability to be influenced by the moral law and hence to be motivated by it: this agent's lack of belief in an afterlife in which the

³⁷⁴ Lewis White Beck, *A Commentary on Kant's Critique of Practical Reason*, 245.

³⁷⁵ 5:452.

highest good is established would “weaken the respect, by which the moral law immediately influences him to obedience”.³⁷⁶ It might seem Kant is saying that an agent to whom the hope of the highest good is not available lacks access to a helpful motivational supplement that removes or diminishes inclinations, such as despair, that could undermine his capacity for rational reflection and so undermine the influence of the moral law on him.³⁷⁷

However, it should be emphasized again that if Kant is going to be consistent, the moral law alone should be sufficient to motivate an agent to perform her duty, as Kant seems to forget it was in the case of Spinoza himself. Further, if any agent does what is right strictly for the sake of some reward they understand the highest good to offer – namely, one’s own happiness – and so acts on a maxim of self-love, then that agent will not be acting in a way that merits moral worth. So, to speak of this as a motivational supplement is somewhat misleading, since the good will needs no supplement to act morally. Nevertheless, the idea of a motivational supplement might work on the Kantian picture in a different sense. It could be that the idea of the highest good and an author that can bring it about is meant to make the prospect of a morally better world more vivid to the agent and to stir up in her the desire to see this better world come about as a result of her actions. In a case in which the desire accompanying the hope of the highest good brings about the correct action such that the agent is motivated by inclination and not the moral law, this desire may still be useful for morality in a way analogous to external moral habit. This is because it would lead her to perform the correct, moral action and

³⁷⁶ 5:452.

³⁷⁷ This is roughly Lawrence Pasternack’s position, outlined in Lawrence Pasternack, *Kant on Religion Within the Boundaries of Mere Reason* (London: Routledge, 2014), 51-54. With regard to the loss in capacity for rational reflection, this might be in the same way that Kant believes too much drink can have this effect, as when he says that a “human being who is drunk is like a mere animal” (6:427).

would produce better consequences (contributing to the happiness of others, for example), though not in a way that would garner moral worth strictly speaking.

A similar and perhaps more concrete, though no less difficult ideal to achieve than the highest good is a state of perpetual peace between nations. In closing his treatise on this topic (*TPP*), Kant says:

[i]f it is a duty to realize the condition of public right, even if only in approximation by unending progress, and if there is also a well-founded hope of this, then the *perpetual peace* that follows upon what have till now been falsely called peace treaties (strictly speaking, truces) is no empty idea but a task that, gradually solved, comes steadily closer to its goal (since the times during which equal progress takes place will, we hope, become always shorter).³⁷⁸

Kant's apparent optimism about the attainment of this ideal might strike us as naïve from the point of view of our current moment in history. In any case, the belief that progress is being made toward perpetual peace and a better world more generally may serve as a kind of motivational supplement in the way outlined above. There may be moments when people are encouraged along in the journey of moral perfection by the sincere and seemingly well-founded belief that the world is becoming a better place as a result of their efforts. However, there may also be moments when the story of progress is more difficult to believe. Though we may ultimately be unable to know the end toward which history is moving, we can nevertheless know that we have a duty to work toward this better world. So, motivational supplements aside, the duty to strive after the highest good remains.

According to Kant, our ideals may be motivationally efficacious, both because we can act on them directly (as when we act directly on the highest good, which just is the moral law under a different description) and because these ideals can further interact with

³⁷⁸ 8:386.

and structure our inclinations. An effect this can have on us is to enable us to feel the force of the moral law more firmly and to see what it requires of us more clearly.³⁷⁹ In the above way, the highest good acts as a kind of fixed point that focuses moral effort. Ideals like this are helpful in enabling us to see beyond the vicissitudes of circumstance, and they are necessary in enabling us to think strategically about the kind of world our actions are meant to create.³⁸⁰ In short, they enable us to situate our actions in a greater project that transcends the moral limitations of our particular time and place, in the project Kant would refer to as the bringing about of a kingdom of ends.

Conclusion

In this thesis, I have argued that Kant's work contains a coherent and compelling account of moral character. This account is consistent with his broader practical philosophy, in particular because it connects with his account of moral worth. Moral worth is of deep concern to Kant because it tells us about the sort of person an agent fundamentally is – whether she is good and can be counted on to do the right thing. (In

³⁷⁹ With regard to enabling us to feel the force of the moral law more firmly, this could be something analogous to external moral habit, as mentioned, or something analogous to the moral feeling Kant speaks of in the *CPrR*, which by “removal of the counterweight” of inclinations that would lead one to act contrary to the moral law, enables the agent to feel the “relative weightiness of the law” and hence, act on the moral law (5:76). With regard to seeing what the moral law requires of us more clearly, I have in mind the duty of sympathy from the *MM*, the immediately relevant features of which I discussed at length in chapter two.

³⁸⁰ See Rae Langton's idea of “strategy, for the Kingdom's sake”, which she adapts from Korsgaard, who argues that “when circumstances are far from the ideal, the ideal becomes a goal, something to strive towards, not something to live up to.” (Rae Langton, “Duty and Desolation,” in *Philosophy* 67 (1992): 502-504.) Kant, of course, would protest the kind of strategy that allowed an agent to, for example, lie to preserve the dignity (or life) of another person, as is clear in *SRL*. But we may depart from Kant on this point if we like.

this way, what has moral worth is worthy of praise independently of whether it is in fact praised by others or not.) Perhaps more importantly, it tells us about the sort of agents we should aspire to become – people who are committed to furthering and not violating the dignity and well-being of humanity. For Kant, an account of moral character will entail an account of moral worth, because moral character is meant to encourage and enable morally worthy agency and full moral character is required for full moral worth. For him, character without moral worth is not moral.

Chapter one outlined a theory of moral education in which the development of character in the standard sense (of training one's inclinations so that they direct one toward right action) is meant to encourage the grounding of character. The grounding of character, or change of heart, is Kant's term for a decisive commitment to the moral law as a guiding principle for one's life and action. This grounding of character is necessary for moral worth, because moral worth requires that an agent's fundamental maxim, which is precisely a commitment of this kind, be the moral law. This chapter further outlined four types of habit in Kant's system, including external moral habit and moral aptitude, among others. External moral habit encourages moral worth because it encourages self-control and the development of mental faculties that can grasp correct moral principles. But it cannot guarantee these things and does not qualify an agent for an ascription of moral worth. Moral aptitude, on the other hand, does qualify an agent for ascriptions of moral worth because it is a cultivated proficiency for thinking and acting on the basis of moral maxims.

Chapter two developed Kant's account of sympathy, arguing it is a kind of moral aptitude of special importance on the Kantian conception of moral character. It is of

special importance because it provides knowledge of how the moral law applies, or what one's duty is in particular circumstances. It does this primarily by telling an agent both about what furthers and what violates the well-being of others; it helps one to fulfill one's wide duties as well as avoid violations of narrow duty. In this way, Kantian sympathy is primarily a cognitive disposition; it is about acquiring correct beliefs to enable moral action and, while it may make use of feelings, it need not necessarily do so. This makes it analogous to a kind of practical wisdom, or Aristotelian *phronēsis*, and I argued it is necessary for full moral character. It also underscores that Kant's account of character is an intellectualist one, because it emphasizes the primacy of intellectual capacities like belief, cognition, and judgment in the constitution of good moral character.

Chapter three then developed an account of moral worth in light of what Kant says about character and some of the more difficult things he says about moral worth. I explained how morally worthy action may involve participating inclinations, so long as it is grounded in the moral law. This means that if the action were morally required, it would be performed even in the absence of these inclinations and, if these (or other) inclinations were to encourage wrong action, the moral law would prevent the agent from performing that wrong action. I also explained how full moral worth requires sympathy, or full moral character, and introduced the idea of mitigated moral worth for agents who act rightly but without proper sympathy (for example, agents who act for confused reasons as in the case of Huck Finn) yet are doing the best that can reasonably be expected of them. Chapter four responded to several objections concerning the nature and role of the good in Kant's moral psychology, to which I argued Kant has a more than satisfying response.

It has also been my aim to show that Kant's account of moral character is much closer to standard intuition than many have thought. However, this does not mean his account will be entirely satisfying on every point. In particular, his tendency to emphasize the instrumental value of feeling and not to appreciate that there may be value in experiencing certain feelings regardless of what one does with them, might be less than entirely satisfying. Perhaps this also overlooks the complexity and so the moral value of certain emotional structures that can support the disposition of sympathy, even as Kant conceives of it. Further, there can be something beautiful, for example, in a person who is moved by the suffering of others and chooses to suffer alongside them even when this might not be necessary, perhaps out of a desire to express love or solidarity.³⁸¹ This is not something Kant seems to have room for in his official theory.

In any case, he does have room for positive, participating inclinations in his account of moral worth and character: after all, human happiness is of central moral concern for him. In sum, Kant's account of moral character is coherent, much richer, and more interesting than has often been thought. It indicates that further investigation into and development of his broader moral psychology would be deeply worthwhile.

³⁸¹ It could be argued that a choice like this is of value because it can inspire moral action in others, but this is again to place the value elsewhere and to make the feeling instrumental.

Bibliography

- Allison, Henry. *Kant's Theory of Freedom*. Cambridge: Cambridge University Press, 1990.
- Annas, Julia. *Intelligent Virtue*. Oxford: Oxford University Press, 2011.
- Anscombe, G.E.M. "Modern Moral Philosophy." *Philosophy* 33:124 (1958): 1-19.
- Aristotle. *Nicomachean Ethics* 2nd Edition. Translated by Terence Irwin. Indianapolis: Hackett, 1999.
- Arpaly, Nomy. *Unprincipled Virtue*. Cambridge: Cambridge University Press, 2003.
- Baron, Marcia. *Kantian Ethics Almost Without Apology*. Ithaca: Cornell University Press, 1995.
- Baxley, Anne Margaret. *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press, 2010.
- Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. Chicago: University of Chicago Press, 1960.
- Bennett, Jonathan. "The Conscience of Huckleberry Finn." *Philosophy* 49 (1974): 123-134.
- Bostock, David. *Aristotle's Ethics*. Oxford: Oxford University Press, 2000.
- Carbonell, Vanessa. "What Moral Saints Look Like." *Canadian Journal of Philosophy* 39:3 (2009): 371-398.
- Charles, David. *Aristotle's Philosophy of Action*. London: Duckworth, 1984.
- Cicero. *On Obligations*. Translated by P.G. Walsh. Oxford: Oxford University Press, 2008.
- Cooper, John. "The Unity of Virtue." *Social Philosophy and Policy* 15:1 (1998): 233-274.
- Duden Deutsches Universalwörterbuch*. Mannheim: Dudenverlag, 2003.
- Eliot, George. *Middlemarch*. Oxford: Oxford University Press, 2008.
- Engstrom, Stephen. "The Concept of the Highest Good in Kant's Moral Theory." *Philosophy and Phenomenological Research* 52:4 (December 1992): 747-780.
- Frierson, Patrick. "Affects and Passions." *Kant's Lectures on Anthropology: A Critical Guide*. Edited by Alix Cohen. Cambridge: Cambridge University Press, 2014.
- Goethe, J.W.V.. *Iphigenia auf Tauris*. Project Gutenberg: <http://www.gutenberg.org/cache/epub/2054/pg2054-images.html>.
- Greene, Joshua. "The Secret Joke of Kant's Soul." *Moral Psychology Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders and Development*. Edited by Walter Sinnott-Armstrong. Boston: MIT Press, 2007.
- Guyer, Paul. *Kant on Freedom, Law, and Happiness*. Cambridge: Cambridge University Press, 2000.

- _____. "Moral Feelings in the *Metaphysics of Morals*." *Kant's Metaphysics of Morals: A Critical Guide*. Edited by Lara Denis. Cambridge: Cambridge University Press, 2010.
- _____. Personal Conversation with Carl Hildebrand. Istanbul, June 2016.
- Hare, R.M.. *Freedom and Reason*. Oxford: OUP, 1963.
- Hegel, G.W.F.. *Outlines of the Philosophy of Right*. Translated by T.M. Knox. Oxford: Oxford University Press, 2008.
- Henson, Richard. "What Kant Might Have Said: Moral Worth and the Over-determination of Dutiful Action." *Philosophy Review* 88:1 (1979), 39-54.
- Herman, Barbara. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press, 1993.
- Hill, Thomas. "Punishment, Conscience, and Moral Worth." In *Kant's Metaphysics of Morals: Interpretive Essays*. Edited by Mark Timmons. Oxford: Oxford University Press, 2002.
- Horace. *The Complete Odes and Epodes*. Translated by David West. Oxford: Oxford University Press, 2008.
- Kant, Immanuel. *Anthropology From A Pragmatic Point of View*. Translated by Robert Louden. Cambridge: Cambridge University Press, 2006.
- _____. *Groundwork of the Metaphysics of Morals*. Translated by Mary Gregor. In *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- _____. *Lectures on Pedagogy*. Translated by Robert Louden. In *Anthropology, History, and Education*. Edited by Robert Louden and Günter Zöllner. Cambridge: Cambridge University Press, 2007.
- _____. *On a Supposed Right to Lie from Philanthropy*. Translated by Mary Gregor. In *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- _____. *Religion Within the Boundaries of Mere Reason*. Translated by Allen Wood and George di Giovanni. Cambridge: Cambridge University Press, 1998.
- _____. *The Critique of Judgment*. Translated by Paul Guyer and Eric Matthews. Cambridge: Cambridge University Press, 2000.
- _____. *The Critique of Practical Reason*. Translated by Mary Gregor. In *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- _____. *The Critique of Pure Reason*. Translated by Paul Guyer and Allen Wood. Cambridge: Cambridge University Press, 1998.
- _____. *The End of All Things*. Translated and edited by Allen Wood and George di Giovanni. Cambridge: Cambridge University Press, 1998.
- _____. *The Metaphysics of Morals*. Translated by Mary Gregor. In *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- _____. *Toward Perpetual Peace*. Translated by Mary Gregor. In *Practical Philosophy*. Edited by Mary Gregor. Cambridge: Cambridge University Press, 1996.
- Korsgaard, Christine. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- Kuehn, Manfred. "Kant's *Metaphysics of Morals*: The History and Significance of its Deferral." *Kant's Metaphysics of Morals: A Critical Guide*. Edited by Lara Denis. Cambridge: Cambridge University Press, 2010.
- Langton, Rae. "Duty and Desolation." *Philosophy* 67 (1992): 481-505.

- Louden, Robert. "Introduction to Kant's *Lectures on Pedagogy*". *Anthropology, History, and Education*. Edited by Robert Louden and Günter Zöller. Cambridge: CUP, 2007.
- _____. *Kant's Impure Ethics*. Oxford: Oxford University Press, 2000.
- MacIntyre, Alasdair. *A Short History of Ethics 2nd Edition*. London: Routledge, 1998.
- Mandela, Nelson. *The Long Walk to Freedom*. London: Abacus, 1995.
- Moore, Adrian. "Maxims as Thick Ethical Concepts." *Ratio* 11:6 (2006): 129-147.
- Nagel, Thomas. *Mortal Questions*. Cambridge: Cambridge University Press, 1979.
- Nietzsche, Friedrich. *The Gay Science*. Translated by Josefine Nauckhoff. Cambridge: Cambridge University Press, 2001.
- Nussbaum, Martha. *Poetic Justice: The Literary Imagination and Public Life*. Boston: Beacon Press, 1995.
- _____. *Love's Knowledge: Essays on Philosophy and Literature*. Oxford: Oxford University Press, 1990.
- O'Neill, Onora. *Acting on Principle*. New York: Cambridge University Press, 1975.
- _____. *Constructions of Reason*. Cambridge: Cambridge University Press, 1989.
- Oxford English Dictionary*. Oxford: Oxford University Press.
- Pasternack, Lawrence. *Kant on Religion Within the Boundaries of Mere Reason*. London: Routledge, 2014.
- Plato. *Phaedrus*. Translated by R. Hackforth. Cambridge: Cambridge University Press, 1997.
- _____. *Phaedrus*. Translated by Stephen Scully. Newburyport: Focus Publishing, 2003.
- Schiller, Friedrich. *Über die Grundlage der Moral*. Translated by A.B. Bullock and quoted in H.J. Paton, *The Categorical Imperative*. London: University of Pennsylvania Press, 1971.
- Schopenhauer, Arthur. *On the Basis of Morality*. Translated by E.F.J. Payne. Indianapolis: Hackett, 1995.
- Seneca. "On Tranquility of Mind." *Seneca: Moral Essays Volume II*. Translated by John Basore. Cambridge, MA: Harvard University Press, 2006.
- Shakespeare, William. "Merchant of Venice." *The Oxford Shakespeare: The Complete Works 2nd Edition*. Edited by John Jowett et. al. Oxford: Oxford University Press, 2005.
- Sherman, Nancy. *Making a Necessity of Virtue*. Cambridge: Cambridge University Press, 1997.
- Silber, John. "Kant's Conception of the Highest Good as Immanent and Transcendent." *The Philosophical Review* 68:4 (1959): 469-492.
- Simmons, Keith. "Kant on Moral Worth". *History of Philosophy Quarterly* 6:1 (1989): 85-100.
- Smit, Houston and Mark Timmons. "The Moral Significance of Gratitude in Kant's Ethics." *The Southern Journal of Philosophy* 39:4 (2011): 295-320.
- Smith, Adam. *The Theory of Moral Sentiments*. London: Penguin Books, 2009.
- Stocker, Michael. "The Schizophrenia of Modern Ethical Theories." *Journal of Philosophy* 73 (1976): 453-466.
- Theresa of Avila. *Interior Castle*. Translated and edited by E. Allison Peers. New York: Dover, 1946.
- Timmermann, Jens. "Kant's Puzzling Ethics of Maxims." *The Harvard Review of Philosophy* 8 (2000): 39-52.

- Twain, Mark. *Adventures of Huckleberry Finn*. Oxford: Oxford University Press, 2008.
- Walker, Ralph. "The Rational Imperative: Kant Against Hume." *Proceedings of the British Academy* LXXIV (1988): 113-133.
- Williams, Bernard. *Ethics and the Limits of Philosophy*. London: Routledge, 2006.
- _____. *Moral Luck*. Cambridge: Cambridge University Press, 1981.
- _____. *Problems of the Self*. Cambridge: Cambridge University Press, 1973.
- Wolf, Susan. "Moral Saints." *The Journal of Philosophy* 79:8 (August 1982): 419-439.
- Wood, Allen. *Kant's Moral Religion*. Ithaca: Cornell University Press, 1970.
- _____. *Kantian Ethics*. Cambridge: Cambridge University Press, 2008.
- _____. "The Final Form of Kant's Practical Philosophy." *Kant's Metaphysics of Morals: Interpretive Essays*. Edited by Mark Timmons. Oxford: Oxford University Press, 2002.

Word Count

64,753 with notes and with bibliography etc.
61,994 with notes and without bibliography etc.
50,783 without notes and without bibliography etc.
299 on abstract