

KANTIAN ETHICS AND THE FORMULA OF HUMANITY: TOWARDS VIRTUES AND ENDS

Omar Bachour

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
in partial fulfillment of the requirements
for the Master of Arts degree in Philosophy

Department of Philosophy
Faculty of Arts
University of Ottawa

ABSTRACT

The aim of this work is to show that criticisms of Kantian ethics from the field of virtue ethics misfire because they rely on a widespread reading of Kant which centers on the *Groundwork* and the Formula of Universal Law as the key elements in his moral philosophy. This reading, I argue, is susceptible both to charges of “empty formalism” and moral “rigorism” as well as the complaint voiced by virtue ethicists that Kantian ethics lacks a full-blooded account of the virtues, along with the attendant desiderata of sociality, character and the emotions. In response, I defend the proposal that the Formula of Humanity and the Doctrine of Virtue in the *Metaphysics of Morals* represent the final form of Kant’s ethical thought. If this is accurate, a rich and novel ethical theory emerges, and many of the criticisms from the field of virtue ethics are subsequently disarmed.

TABLE OF CONTENTS

Introduction.....	1
I. The ‘Aretaic Turn’	
II. A Short Conspectus of Criticism	
III. What is Virtue Ethics?	
IV. General Overview	
Chapter 1: The <i>Groundwork</i>, Duty and the Formula of Universal Law.....	9
I. The Good Will	
II. Acting from Duty: Three Appeals to Common Rational Moral Cognition	
III. Kant’s Rejection of the Inclinations as Moral Motives	
IV. Unsociable Sociability	
V. The Imperatives and Happiness	
VI. The Formula of Universal Law	
VII. Applying FUL/FLN: The Four Examples	
VIII. Beyond FUL/FLN	
Chapter 2: Rational Nature and the Formula of Humanity.....	37
I. The Formula of Humanity as an End in Itself	
II. The Radical Egalitarianism of FH	
III. The Derivation of FH	
• <i>Part I: The Argument by Elimination</i>	
• <i>Part II: The Regressive Argument</i>	
• <i>The Crucial Premises Explained</i>	
IV. Objections to FH	
• <i>Objection 1: ‘Rational nature is ambiguous.’</i>	
• <i>Objection 2: ‘FH is formalistic.’</i>	
• <i>Objection 3: ‘FH is not egalitarian.’</i>	
• <i>Objection 4: ‘End-setting does not determine objective goodness.’</i>	
• <i>Objection 5: ‘The Regressive Argument relies on a faulty inference.’</i>	
V. Applying FH: The Four Examples Revisited	
Chapter 3: The Doctrine of Virtue.....	66
I. Expressive Ethics	
II. The Metaphysics of Morals	
III. Virtue	
IV. Obligatory Ends: ‘The Ends that are also Duties’	
• <i>One’s Own Perfection</i>	
• <i>The Happiness of Others</i>	
V. The Taxonomy of Ethical Duties	

- VI. The Duties of Virtue
- *Perfect Duties to Oneself as an Animal Being (Against):*
 1. *Suicide*
 2. *Lust*
 3. *Drunkenness and Gluttony*
 - *Perfect Duties to Oneself as a Moral Being (Against):*
 4. *Lying*
 5. *Avarice*
 6. *Servility*
 - *Imperfect Duties to Oneself:*
 7. *Natural Perfection*
 8. *Moral Perfection*
 - *Duties to Others: Respect (Against)*
 9. *Arrogance*
 10. *Defamation*
 11. *Ridicule*
 - *Duties to Others: Love*
 12. *Beneficence*
 13. *Gratitude*
 14. *Active Sympathetic Participation (Teilnehmung)*
 - *The Malevolent Dispositions: 'Vices of Hatred for Human Beings'*
- VII. How Much Latitude Do Wide Duties Permit?

Chapter 4: From the Inclinations to the Kingdom of Ends.....117

- I. Desires, Affects, Passions, Delusions
- II. Moral Feelings
- III. Two Conceptions of Virtue
 - *Aristotelian Virtue*
 - *Kantian Virtue*
- IV. Are the Inclinations the Enemy of Virtue?
- V. The Kingdom of Ends
 - *Friendship*
 - *Ground Projects*
 - *Desolation*

Conclusion.....157

Bibliography.....161

ABBREVIATIONS

Works by Immanuel Kant (all citations refer to the following English translations which include the volume and page number of the Prussian Academy edition of *Kant's gesammelte Schriften*)

- G* *Groundwork of the Metaphysics of Morals*. Revised Edition. Ed. and trans. Mary Gregor and Jens Timmermann. Introduction. Christine M. Korsgaard. New York: Cambridge University Press, 2012.
- I* *Idea of a Universal History with a Cosmopolitan Aim*. Trans. Allen W. Wood. *Anthropology, History, and Education*. Ed. Günter Zöller and Robert B. Loudon. New York: Cambridge University Press, 2007.
- KpV* *Critique of Practical Reason*. Ed. and trans. Mary Gregor. Introduction. Andrews Reath. New York: Cambridge University Press, 1997.
- KrV* *Critique of Pure Reason*. Ed. and trans. Paul Guyer and Allen W. Wood. New York: Cambridge University Press, 1998. Cited by A/B pagination.
- KU* *Critique of the Power of Judgement*. Ed. Paul Guyer. Trans. Paul Guyer and Eric Mathews. New York: Cambridge University Press, 2000.
- MS* *The Metaphysics of Morals*. Ed. and trans. Mary Gregor. Introduction. Roger J. Sullivan. New York: Cambridge University Press, 1996.
- R* *Religion within the Boundaries of Mere Reason*. Ed. and trans. Allen Wood and George di Giovanni. Introduction. Robert Merrihew Adams. New York: Cambridge University Press, 1998.
- VA* *Anthropology from a Pragmatic Point of View*. Ed. and trans. Robert B. Loudon. Introduction. Manfred Kuehn. New York: Cambridge University Press, 2006.
- VE* *Lectures on Ethics*. Trans. Peter Heath. Ed. Peter Heath and J. B. Schneewind. New York: Cambridge University Press, 1997.

Other Works Cited Frequently

- CI* Paton, H. J. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. Philadelphia, Pennsylvania: University of Pennsylvania Press, 1947.
- CKE* Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- DL* Seymour, Melissa M. *Duties of Love and Kant's Doctrine of Obligatory Ends*. Diss. Indiana University, 2007. Ann Arbor: UMI 3277970.

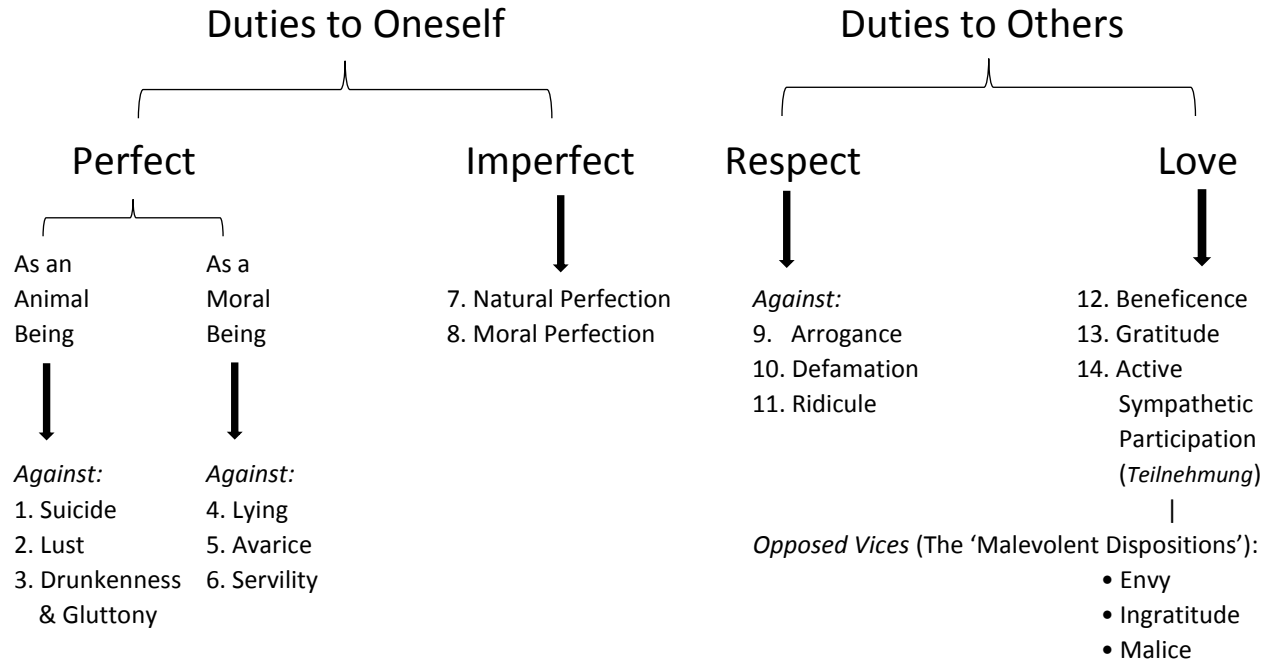
- HW* Hill Jr., Thomas E. *Human Welfare and Moral Worth: Kantian Perspectives*. New York: Clarendon Press, 2002.
- KE* Wood, Allen W. *Kantian Ethics*. New York: Cambridge University Press, 2008.
- KET* Wood, Allen W. *Kant's Ethical Thought*. New York: Cambridge University Press, 1999.
- KEWA* Baron, Marcia W. *Kantian Ethics (Almost) Without Apology*. New York: Cornell University Press, 1995.
- ML* Williams, Bernard. *Moral Luck: Philosophical Papers 1973-1980*. New York: Cambridge University Press, 1981.
- NE* Aristotle. *Nicomachean Ethics*. Trans. Christopher Rowe. Introduction and Commentary. Sarah Broadie. New York: Oxford University Press, 2002.
- PMJ* Barbara, Herman. *The Practice of Moral Judgement*. Cambridge: Harvard University Press, 1993.
- SS* Langton, Rae. *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*. New York: Oxford University Press, 2009.

Formulas

- FUL* The Formula of Universal Law: “Act only according to that maxim through which you can at the same time will that it become a universal law” (G 4:421).
- FLN* The Formula of the Law of Nature: “So act as if the maxim of your action were to become by your will a **universal law of nature**” (G 4:421).
- FH* The Formula of Humanity as an End in Itself: “So act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means” (G 4:429).
- FKE* The Formula of the Kingdom of Ends: ““Act according to the maxims of a member universally legislating for a merely possible kingdom of ends” (G 4:439).

KANT'S COMPLETE TAXONOMY OF THE DUTIES OF VIRTUE

DUTIES OF VIRTUE



Introduction

I. The 'Aretaic Turn'

"Modern Moral Philosophy," Elizabeth Anscombe's 1958 essay which ushered in the "aretaic turn" in moral philosophy, involved a scathing attack on both the Kantian and consequentialist traditions that had dominated the field from the eighteenth century onwards. In their place, Anscombe called for a return to the Aristotelian language of virtue understood in terms of human flourishing. Until the notion of virtue is illumined by a robust moral psychology, we should, she argued, stop doing moral philosophy.

The contemporary revival of virtue ethics that followed in the wake of Anscombe's essay took up her call in different ways, often in directions she had not anticipated. Despite its various stripes, one of the shared features of this literature was a pointed dissatisfaction with Kantian ethics. The criticisms, many of which had been around since Kant's time (via Herder and Hegel, for example), centered on Kant's moral "rigorism" and "empty formalism."

The criticisms that fell under the heading of "rigorism" took issue with the notion, derived almost entirely from the *Groundwork*, that the moral law "remains in full force" (G 4:439), as a source of absolute, exceptionless prohibitions, regardless of all human circumstances, consequences and desires. As a result, Kantian ethics was charged with failing to appreciate the moral complexities of situations, being unable to accommodate moral change, and advocating a rigid system of deontological duties in the face of immoral outcomes.

The problem of "empty formalism" was tied closely to the former: if Kant's theory failed to take into account contingent circumstances (for whatever reason) and abstracted from *all* ends, as its critics alleged, then it left us with nothing more than a husk of unalloyed rationalism, bereft of any determinate content. This naturally led to a host of criticisms from the virtue ethicists relating to virtue, sociality, character and the emotions which they saw as sacrificed to the *form* of Kant's ethics.

Relying predominantly on a reading of the *Groundwork*, virtue ethicists were quick to point to Kant's infamous example in which a compassionate man's beneficent action "has no true moral worth"—while the same man's action, when his temperament has been clouded by life's sorrows, and he acts beneficently from duty, is for the first time imbued with "genuine moral worth" (*G* 4:398).

This ill-fated example garnered Kant more hostility than any other part of his moral doctrine, making appearances in the works of almost all his detractors, and resulting in the now familiar image of Kantian ethics as "representative of moralistic strictness and sternness, downright hostile to human happiness, mercilessly unsympathetic to human weakness, [and] allowing no place in the moral life for natural human feelings and desires" (Wood, *KE* 2). In the following section, we briefly review some of these criticisms.

II. A Short Conspectus of Criticism

In 1797, Friedrich Schiller wrote the following satirical verse in *Xenien*, mocking Kant's ostensible emphasis on duty over feeling:

Gladly I serve my friends, but alas I do it with pleasure.

Hence I am plagued by doubt that I am not a virtuous person.

To this the answer is given:

Sure, your only resource is to try to despise them entirely,

And then with aversion to do what your duty enjoins you.

(qtd. in Paton, *CI* 48)

The Romanian thinker, E. M. Cioran, points to Kant's ethical works as the catalyst which precipitated his break with philosophy. In bristling Nietzschean prose, he writes:

I turned away from philosophy when it became impossible to discover in Kant any human weakness, any authentic accent of melancholy....[Kantian] philosophical activity proceeds from a suspect depth, prestigious only for the timid and the tepid....impersonal anxiety, refuge among anemic ideas—[it] is the recourse of all those who would elude the *corrupting* exuberance of life. (Cioran, *A Short History of Decay* 47)

Theodor Adorno, in his lectures on moral philosophy, argues that Kant “banishes sympathy, compassion and the direct expression of pity from his ethics because all impulses of this sort are merely natural impulses” (Adorno, *Problems of Moral Philosophy* 120). In the same passage, Adorno issues the following caveat: “Because this extreme view...of freedom is based on its absolute independence from all existing beings, [and] from nature as such, it threatens to become transformed into unfreedom” (Adorno, *Problems of Moral Philosophy* 120).

In *A Short History of Ethics*, in his section on Kant, MacIntyre concludes:

The logical emptiness of the test of the categorical imperative is itself of social importance. Because the Kantian notion of duty is so formal that it can be given almost any content, it becomes available to provide a sanction and a motive for the specific duties which any particular social and moral tradition may propose. Because it detaches the notion of duty from the notions of ends, purposes, wants and needs, it suggests that, given a proposed course of action, I may only ask whether, in doing it, I can consistently will that it shall be universally done, and not ask what ends or purposes it serves. Anyone educated into the Kantian notion of duty will, so far, have been educated into easy conformism with authority. (MacIntyre, *A Short History of Ethics* 191)

The charge of “empty formalism” is not only of social importance, but also holds grave implications for character. Thus, in “Persons, Character and Morality,” Bernard Williams writes that “once one thinks about what is involved in having a character, one can see that the Kantians’ omission of character is a condition of their ultimate insistence on the demands of impartial morality, just as it is a reason for finding inadequate their account of the individual” (Williams, *ML* 14).

At the end of his article, Williams openly pits a “life of substance” against one that is informed by Kant’s ethics, which he believes is antipodal to the former: “Life has to have substance if anything is to have sense, including adherence to the impartial system; but if life has substance, then it cannot grant supreme importance to the impartial system...” (Williams, *ML* 18).

Addressing Kant's case which compares the person who acts beneficently from sympathy to one who does so strictly from duty, Michael Stocker identifies a *moral schizophrenia* in any ethical theory which countenances a split between one's motives and one's reasons. Before discussing a now famous example in which, asked by your friend why you have come to visit her at the hospital, you reply unsatisfactorily that it was your duty that prompted the visit, Stoker asks:

What sort of life would people have who did their duties but never or rarely wanted to? Second, duty, obligation, and rightness are only one part—indeed, only a small part, a dry and minimal part—of ethics. There is the whole other area of the values of personal and interpersonal relations and activities; and also the area of moral goodness, merit, virtue...in both motive and reason must be in harmony for the values to be realized. (Stocker, "The Schizophrenia of Modern Ethical Theories" 67)

Considering a scenario in which a man, acting from duty, offers a woman with a crying child his bus seat, N. J. H. Dent describes him as "schooled in the duties of a gentleman" acting in "a somewhat wooden or truculent fashion" (Dent, *The Moral Psychology of the Virtues* 15). "Either way," Dent tells us, "it is no kindly impulse that prompts his action" (Dent, *The Moral Psychology of the Virtues* 15-6). Both Philippa Foot and Rosalind Hursthouse return, time and time again, to Kant's "curious doctrine" with its "strange conclusion" (Foot, "Virtues and Vices" 173).

In her well-known essay, Susan Wolf decries moral saintliness as an unhealthy and undesirable ideal, including what she refers to as the "Rational Saint":

I suspect that the range of activities acceptable to the Kantian saint will remain objectionably restrictive....the Kantian would have to value his activities and character traits in so far as they are manifestations of respect for the moral law....This is a good and noble motivation, to be sure. But it is hardly what once expects to be dominantly behind a person's aspirations to dance as well as Fred Astaire, to paint as well as Picasso, or to solve some outstanding problem in abstract algebra, and it is hardly what one hopes to find lying dominantly behind a father's action on behalf of his son or a lover's on behalf of her beloved. (Wolf, "Moral Saints" 90)

Instead of being mitigated or dispelled by recent shifts in Kantian scholarship, the caricature of Kantian ethics as one of such moral rigor and self-mastery that it borders on inhumanity has persisted.

In 1970, Iris Murdoch compared the Kantian man, in a poetic turn of phrase, to Milton's Lucifer:

[Kant's] man is with us still, free, independent, lonely, powerful, rational, responsible, brave, the hero of so many novels and books of moral philosophy....increasingly aware of his alienation from the material universe his discoveries reveal....his alienation is without cure....It is not such a very long step from Kant to Nietzsche, and from Nietzsche to existentialism and the Anglo-Saxon ethical doctrines which in some ways closely resemble it. In fact Kant's man had already received glorious incarnation nearly a century earlier in the work of Milton: his proper name is Lucifer. (Murdoch, *The Sovereignty of Good* 78)

Here is Simon Blackburn's vivid depiction of the "Kantian Captain" 28 years later:

[T]he Captain, the will, yourself as an embodiment of pure practical reason, detached from all desires...always stands ready to stop things going wrong with the [Humean] crew's handling of the boat. Sometimes, it seems, that happiest ship will have no crew at all, but only a Captain....Thus the Kantian Captain. He is a peculiar figure, a dream—or nightmare—of pure, authentic self-control....Context-free, non-natural, and a complete stickler for duty, perhaps the Kantian self is nothing but the sublimation of a patriarchal, authoritarian fantasy. (Blackburn, *Ruling Passions* 246-8)

In the opening paragraph of his essay, "Partiality and the Virtues", John Cottingham notes that "despite the scholarship and eloquence of its defenders, Kant's insistence that moral worth is reserved for the austere motivated act of pure duty...seems to bleach out the moral worth from much of our lives—conditioned as they are by the ties of partiality, the 'sensible warm motions' of the human heart" (Cottingham, "Partiality and the Virtues" 57).

The startling and vituperative culmination of this line of criticism is captured well by the philosopher Richard Taylor, who writes:

With every year of my life I become increasingly overwhelmed with the basic offensiveness of Kant's system of metaphysical morals, and I am very sure that this is not because I have somehow failed to

understand it. On the contrary, I understand it very well, and that kindles my dislike....[and] if I were ever to find...a man who assured me that he really *believed* Kant's metaphysical morals, and that he faithfully modeled his conduct and his relations with others after those principles, then my incredulity and distrust of him as a human being could not be greater than if he told me he regularly drowned children just to see them squirm. (Taylor, *Good and Evil* xii)

III. What is Virtue Ethics?

Without knowing what virtue ethics is, or what distinguishes it as a rival theory, we cannot fully appreciate the charges it levels at Kantian ethics or adequately defend the latter from them. Further bolstering the need for a definition is Julia Driver's distinction between virtue theory and virtue ethics, since both Kantian ethics and utilitarianism can offer a *theory* of virtue. In posing the question of what virtue ethics is, however, we run into three difficulties.

The first concerns the various forms of virtue ethics: one need only compare MacIntyre's approach with that of Hursthouse or Michael Slote, for example, to see that different philosophers understand virtue ethics very differently and put forward diverse conceptions of character and virtue.

The second difficulty stems from the fact that modern virtue ethics is still at an early stage of development (compared to its traditional counterparts) and therefore "resists precise definition" (Swanton, *Virtue Ethics* 5). Christine Swanton, whose distinction between virtue ethics as a genus and virtue ethics as a species will help us overcome these difficulties, argues that "virtue ethics in its modern development is still in its infancy. It should not therefore be shackled by preconceived ideas about its progeniture and nature" (Swanton, *Virtue Ethics* 5).

The third rejects the question as unduly burdensome from the outset. In *On Virtue Ethics*, Hursthouse complains:

No one, as far as I know, is bothered by the fact that there are no longer satisfactory short answers to the questions 'What is deontology?' and 'What is utilitarianism?', but currently, at least some philosophers seem bothered by the fact that we virtue ethicists cannot come up with one to answer 'What is virtue

ethics?'. The demand that virtue ethics, unlike the other two approaches, should be able to state its position succinctly, in terms both sufficiently broad (or disjunctive?) to get all virtue ethicists in and sufficiently keep all deontologists and utilitarians out, seems a bit excessive. Why should anyone expect us, uniquely, to be able to do it? (Hursthouse, *On Virtue Ethics* 4)

Hursthouse resists the demand for such a definition because she locates it in the belief that virtue ethics is not an a novel and challenging rival to deontological and utilitarian approaches,¹ but our interest is precisely in *what* makes it this sort of (challenging) rival.

To bypass these three difficulties, we can, following Marcia Baron, adopt Swanton's definition of virtue as a genus:

In virtue ethics, the notion of virtue is central in the sense the sense that conceptions of rightness, conceptions of the good life, conceptions of the 'moral point of view' and the appropriate demandingness of morality, cannot be understood without a conception of the relevant virtues. (Swanton, *Virtue Ethics* 5)

This definition is (a) general enough to include various species of virtue ethics, along with their diverse conception of character and virtue, (b) does not limit virtue ethics to a narrow, inchoate definition which will hamper its development and (c) because it aims at identifying what all the species of virtue ethics have in common, cannot be used as a short list of theses to discredit the theoretical field as a whole.

The question now becomes, does Kantian ethics lack a robust conception of the virtues and, with it, a place for the desiderata of sociality, character and the emotions, as its critics allege?

IV. General Overview

Chapter 1 offers a deflationary account of the good will and acting from duty, while arguing that the Formula of Universal Law is plagued by structural difficulties stemming from its formal, provisional and negative nature. Here we are in agreement with Kant's critics: the formula, insofar as it is viewed in

¹ "A deeper reason for the demand that we should come up with a crisp answer to 'What is virtue ethics?', I suspect, is the persistence of the belief that virtue ethics is not...a rival to the deontological and utilitarian approaches, as interesting and challengingly different from either as they are from each other." (Hursthouse, *On Virtue Ethics* 7).

isolation, fails to establish the conclusions that Kant's four examples are meant to deliver. Fortunately, the form of the categorical imperative is merely the first step in Kant's presentation of the moral law.

Chapter 2 turns from the *form* of the moral law to its *matter* (rational nature as an end in itself), dispelling charges of "empty formalism." After advocating a broad, egalitarian conception of rational nature as the capacity to set ends, and establishing the Formula of Humanity as the primary formula in Kant's ethics, we defend its derivation in the form of a "regress on conditions." Finally, we survey five possible objections to the idea of humanity as an end in itself, and revisit Kant's four examples through the lens of the second formula.

Chapter 3 is concerned with the ethical duties furnished by Kant's Formula of Humanity. The two obligatory ends which are also duties are (a) our own perfection and (b) the happiness of others. These, in turn, give rise to fourteen duties of virtues, which we examine in detail. Having established the central role of the virtues in Kant's ethics, we address the question of latitude, arguing *against* any "rigorist" interpretation of Kant's wide or imperfect duties.

Chapter 4 begins with an outline of Kant's account of desire, the affects, passions and delusions. Here we find the anthropological underpinnings of the idea of "unsociable sociability" or "self-conceit" running throughout the previous chapters. Responding to specific charges, we defend the place of *rational* desires as well as empirical feelings in Kant's ethics, and draw parallels between certain aspects of Kantian and Aristotelian virtue. A discussion of the social nature of self-conceit leads us to the kingdom of ends, and its corresponding formula, as the moral answer to our unsociable sociability. Having finally arrived at the *complete determination* of the moral law, we expound Kant's account of friendship, character and sentiment by responding to his critics, and championing the kingdom of ends.

Chapter 1: The *Groundwork*, Duty and the Formula of Universal Law

I. The Good Will

Kant begins the first section of the *Groundwork* with the following sentence: “It is impossible to think of anything at all in the world, on indeed even beyond it, that could be taken to be good without limitation, except a **good will**” (G 4:393). A few paragraphs later, Kant draws a direct link between the good will and the concept of duty, “which contains that of a good will though under certain subjective limitations and hindrances, which, however, far from concealing it and making it unrecognizable, rather bring it out by contrast and make it shine forth all the more brightly” (G 4:397).

Without understanding Kant’s aim in first section of the *Groundwork*, we are easily misled by these early passages into holding two erroneous beliefs about his practical philosophy: “(1) The basis of all Kantian ethics is the unlimited goodness of the good will. [And] (2) For Kantian ethics, the good will is only the will that acts from duty” (Wood, *KE* 39). Combining these two ideas, the reader then fallaciously infers that Kant is recommending that we *maximize* the good will by acting from duty whenever possible which, in turn, translates into a blanket suppression of our natural desires. Nothing, however, could be more alien to Kant’s ethics. Let us begin with the second point.

Kant maintains that a will which acts from duty is a good will, but as H. J. Paton reminds us, “it must not be supposed that a good will is necessarily one which acts for the sake of duty” (Paton, *CI* 46). In other words, acting from duty is a *special case* of the good will. It is therefore a mistake to conclude, as MacIntyre does, that “the good will’s motive is to do its duty for the sake of doing its duty. Whatever it intends to do, it intends because it is its duty” (MacIntyre, *A Short History of Ethics* 185).

We can see that this characterization is false when turn to Kant’s notion of a holy will, whose maxims necessarily harmonize with the laws of morality without the need for self-constraint (G 4:414;

439; *MS* 6:379-80). Duty, in the form of necessitation, for Kant, applies only to imperfect rational beings “who are unholy enough that pleasure can induce them to break the moral law” (*MS* 6:379). This is why there can be no imperatives for the holy will: “here the *ought* is out of place, because *willing* already of itself necessarily agrees with the law” (*G* 4:414). The fact that a holy will is not subject to necessitation, and thus duty, however, does not divest it of its goodness (as MacIntyre’s statement suggests it would). The holy will, which sits beyond duty, is, according to Kant, an “absolutely good will” (*G* 4:439).

Under adverse conditions, however, in which we are forced to struggle against our inclinations when they threaten the moral law, a good will manifests itself in acting from duty. This explains Kant’s qualification that the concept of duty “contains that of a good will *though under certain subjective limitations and hindrances* which...bring it out by contrast and make it shine forth all the more brightly” (*G* 4:397, emphasis added). Heralding the configuration of Kant’s motivational examples, we are now able to see that the good will *is* present in cases where a person assists those in need out of sympathy, for example, but that it “shines forth more brightly” when, in the absence of such inclinations, it must tear itself from this deadly insensibility in order to continue offering aid.

Kant defines the will as “nothing other than practical reason” (*G* 4:412). A *good* will is, therefore, one that adopts good principles (e.g. succour those in need), “whether it thereby acts from self-interest, immediate inclination, or duty” (Wood, *KE* 32). This means that not all acts of the good will have genuine moral worth, nor should we expect them to always involve duties, since there are many cases of action in which no duties apply. More strikingly, even in cases involving duty, Kant believes that we only have a wide or imperfect duty to make duty alone a sufficient motive for our actions (*MS* 6:393; 446-7). Therefore, we incur no blame if we fail to act from duty as long as we are *striving* to make it a sufficient motive for our actions, and (b) “we can not only have a good will but even can achieve moral merit and deserve esteem in cases where that striving is not perfectly successful” (Wood, *KET* 27).

Finally, Kant distinguishes between actions done “*in conformity with duty*” [*pflichtmäßig*] and those done “*from duty*” [*aus Pflicht*] (G 4:390, 398; KpV 5:81). If Kant really believed that actions done in conformity with (but not from) duty reflected an *immoral* or *morally worthless* will, as his critics seem to imply, this would be self-contradictory, since it would suggest that actions done in conformity with duty are immoral and thus *not* in conformity with duty after all (Wood, KE 33). A careful reading of the *Groundwork* militates against such an ascription. We will have much more to say about these conclusions, and others, in our examination of Kant’s examples in the following section.

This leaves us with the first point, which views the unqualified goodness of the good will as the cornerstone of all Kantian ethics. In fact, outside the early pages of the *Groundwork*, the notion appears only rarely in Kant’s other ethical works. In both the Second Section of the *Groundwork* and the *Metaphysics of Morals*, not only are there only incidental references to the unqualified goodness of the good will, if any at all, but the basic value of Kantian ethics is located in the dignity of rational nature as an end in itself, and not in the former. In his essay “Is a Good Will Overrated”, Thomas E. Hill argues that Kant’s thesis (regarding the unqualified goodness of the good will) is “less distinctive and controversial” than both his critics and his sympathizers have typically maintained (Hill, HW 37). The thesis, according to Hill, “has a practical, choice-guiding function, but can be employed only in conjunction with Kant’s fuller account of the fundamental features of a moral attitude. In effect it is just *one of the many ways* of affirming that in deliberation moral considerations should be overriding” (Hill, HW 37, emphasis added). Interpreting Kant’s notion of the good will along the same lines, Nelson Potter writes, “we could also express proposition 1 [A good will only has absolute worth] as a statement about moral value: (1a) Moral value always outweighs any other kind of value” (Potter, “The Argument” 31). This intuitive reading, as we shall see, fits nicely with Kant’s appeals to common rational cognition in the next section.

So why does Kant, in Paton’s words, “begin his argument dramatically” with the unqualified goodness of the good will (Paton, CI 34)? Kant’s strategy in the *Groundwork* is concerned with grounding

the supreme principle of morality by charting its “progression”² scholastically, beginning with its “*form*” then its “*matter*” and finally its “*complete determination*” (G 4:436). In the First Section of the *Groundwork* Kant is attempting to derive the *form* (as opposed to the matter or content) of the supreme principle of morality from common rational moral cognition. This is why he is quick to “rob the will of all impulses that could arise for it from following some practical law” (G 4:402), including any effects of willing in order to arrive at the first formulation of the categorical imperative.

Kant begins by arguing that the “true” or “genuine” moral worth of a human action lies in it being done not from self-interest or immediate inclination, but from duty, which he defines as “*the necessity of an action from respect for the law*” (G 4:398-400). An action from duty, according to Kant, has its moral worth not in the purpose that is attained by it, but in the *principle of willing* according to which it is done. Kant then concludes that, having “robbed the will of all impulses that could arise for it from following some practical law, nothing remains but...the universal conformity of actions with law” (G 4:402). This yields the first formulation—or mere form (universality)—of the supreme principle of morality.³ Our main concern, however, lies only with Kant’s first premise: why does Kant insist that only actions done from duty have genuine moral worth?

II. Acting from Duty: Three Appeals to Common Rational Moral Cognition

Barbara Herman is surely correct in claiming that the best approach to understanding Kant’s strategy is to determine what question he is investigating when he returns the answer, “action[s] done from the motive of duty” (Herman, *PMJ* 2). “Both sympathetic and hostile critics of the doctrine”, Herman writes, “take the question to be obvious: What motive (or motives) distinguish the good moral agent from those of an agent whose actions are merely morally correct” (Herman, *PMJ* 2)? If this is the etiological question that Kant sets out to answer, then Herman finds the harsh reactions to his doctrine justified.

² “...the progression takes place as through the categories of *unity* of the form of the will (its universality), the *plurality* of the matter (of objects, i.e. of ends), and the *allness* or totality of the system of these” (G 4:436).

³ The first variant of FUL reads: “I ought never to proceed except in such a way *that I could also will that my maxim should become a universal law*” (G 4:402)

After introducing the good will, and immediately preceding the three examples, Kant tells us that his task is to “unravel [*entwickeln*]⁴ the concept of a will to be highly esteemed in itself and good apart from any other purpose” (*G* 4:397). In order to do this, he examines cases in which the good will “shines forth more brightly” by subjecting it to “limitations and hindrances” which allow us to study our reactions to it apart from other motives (*G* 4:397). Paton refers to this as “the method of isolation” since each example pits two distilled motives against each other to see which of the two elicits feelings of moral esteem [*Hochschätzung*] from our common rational moral cognition (Paton, *CI* 47-8). Bearing out this interpretation is Kant’s own description of this method in the second *Critique* in which he compares his role to that of a chemist:

A philosopher, however, has...the advantage that, almost like a chemist, 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, namely, by adding the moral law (as determining ground) to the empirically affected will (e.g. that of someone who would gladly lie because he can gain something by it). (*KpV* 5:93)

~

We have at hand examples of reason judging morally. We can analyze them into their elementary concepts and...adopt a procedure similar to that of *chemistry*—the *separation*, by repeated experiments on common human understanding, of the empirical from the rational that may be found in them... (*KpV* 5:163)

The method of isolation results in Kant dealing with “extreme cases in which the two things, duty and desire, work at cross-purposes” but to mistake this oppositional procedure for Kant’s ethics is, as Lewis White Beck reminds us, “based upon many readers’ failure to remember the polemical situation in which Kant found himself, one in which he had to separate and set apart, and seem to set in opposition those things that had been confused...by others” (*A Commentary* 120). This context is also

⁴ Paton translates *entwickeln* as “elucidate,” Lewis White Beck and Allen Wood as “develop” and Mary Gregor (in the first Cambridge edition) as “explicate.”

vital for understanding some of Kant's more strident passages denouncing the inclinations in the *Groundwork* and the second *Critique*.

Kant outlines four types of action: (1) actions contrary to duty, (2) actions in conformity with duty, but not from immediate inclination, (3) actions in conformity with duty from immediate inclination and (4) actions from duty (G 4:397-8). It must not be thought, as it commonly is, that Kant is merely concerned with the psychological causality of actions in the corresponding examples: the real distinction he is after is the one between our *reactions*, via common rational moral cognition, to each action.

The first case Kant "pass[es] over" since, being openly contrary to duty, there is no ambiguity as to whether the actions were done from duty or not (G 4:397). Once again, Kant's interest lies only in cases in which it is difficult to distinguish between the two motives *in order to determine* which motive properly warrants our moral esteem: actions which are decidedly contrary to duty (e.g. murder) clearly fail this criterion.

Case 2: The Prudential Shopkeeper. The second case involves a shopkeeper who charges all his customers, including inexperienced ones, a fixed price so that they are all "served *honestly*" (G 4:397). Thus, the shopkeeper acts in conformity with duty. He does not, however, really want to serve all his customers this way, but does so merely out of concern for his reputation. This is what Kant means by saying that the shopkeeper does not act from immediate inclination (e.g. love for his patrons), or duty, but rather from a self-interested, prudential motive. Like the first case, Kant moves through the example quickly since acting prudentially, even when in conformity with duty, "is not nearly enough for us to believe that the merchant proceeded in this way from...principles of honesty" (G 4:397). In other words, we are not likely to accord *moral* esteem to actions done merely from self-interest.

Case 3: Preserving One's Life. Although he would later come to disagree with them, Kant was influenced by the British moral sense theorists in his early published essays, and they continued to make

appearances in his lectures on ethics.⁵ Thus, in the third example, Kant turns to what he considers to be the strongest source of confusion plaguing our ordinary moral thought: everyone, he argues, has an immediate inclination to preserve their life, but only those who continue to preserve their life when “adversities and hopeless grief have entirely taken away the taste for life” have acted on maxims that have actual moral content (G 4:397-8). To further explore our reactions to actions done from duty compared to those done merely from immediate inclination, Kant introduces his notorious fourth case.

Case 4: The Friend of Humanity. Still in the sphere of immediate inclination, Kant considers the case of a friend of humanity “so attuned to compassion” that without any ground of prudence, vanity or self-interest, he finds “inner gratification in spreading joy” and “can relish the contentment of others” (G 4:398). Kant then asks us to imagine the *same* friend of humanity, beset by such crippling grief that it “extinguishes all compassion for the fate of others” in him: if he “tear[s] himself out of this deadly insensibility” and continues to benefit those in need out of duty, only then do his actions have genuine moral worth (G 4:398-9). Because this example is singlehandedly responsible for much of the hostility directed at Kantian ethics, we would do well to proceed carefully. Kant might have been overly (and, given the history of criticism, disastrously) optimistic in the endorsement he thought his example would secure from common rational moral cognition—but this failure of *reception* should not replace a judicious reading of his argument.

First, in contrasting actions done from duty (the beneficent friend of humanity) to those done solely from immediate inclination (the compassionate friend) Kant is *not* asking the following questions:

1. Which of the two actions has moral value?
2. Which of the two situations do we prefer to be in?
3. Which of the two situations should we educate people to be in?

⁵ Kant’s 1765 *Announcement of the Programme of His Lectures for the Winter Semester 1765–1766* included Shaftesbury, Hutcheson and Hume under the heading of ethics (Schneewind, Introduction xiv).

Rather, the question Kant is asking is the same one he has been concerned with throughout the two previous cases: which action elicits in us feelings of proper moral esteem (Wood, *KE* 27-9)?

III. Kant's Rejection of the Inclinations as Moral Motives

The traditional reading of Kant's fourth case is that helping others from duty is incompatible with being sympathetically inclined to help them. Since Kant attributes moral worth to the former, readers naturally infer that, for Kant, having an inclination to help prevents one's actions from having any moral value.

At first blush, it seems that Kant's distinction is precisely one regarding the moral value of the two actions. After all, he openly declares that the actions stemming from immediate inclination lack "moral content" or "genuine moral worth" (*G* 4:398). Kant's language here is informative: modifiers such as "true", "actual" and "genuine" are reserved for actions done from duty, which he views as having a fundamental moral value worthy of our esteem (*G* 4:398). Now if—in stark contrast to the former—Kant viewed the actions of the compassionate friend as morally worthless (i.e. of *no* moral value whatsoever), his language should reflect this putative disdain for the inclinations. This is, however, not what we find.

Instead, in the same passage Kant refers to actions stemming from immediate inclination, such as those of the compassionate friend of humanity, as "amiable", "in the general interest...and hence honorable" deserving "praise and encouragement, but not high esteem" (*G* 4:398). In the second *Critique*, Kant once again avers that "it is very beautiful to do good to human beings from love for them and from sympathetic benevolence, or to be just from love of order, but this is not *yet* the genuine moral maxim of our conduct (*KpV* 5:82, emphasis added). In other words, we clearly approve of such actions from a moral standpoint, but they do not yet contain what Kant believes is essential to morality. Why is this?

In “Morality and the Emotions”, Bernard Williams sums up what he takes to be Kant’s reasons for his rejection of the inclinations as moral motives as follows: “that the emotions⁶ are too capricious; that they are passively experienced; that a man’s proneness to experience them or not is the product of natural causation” (Williams, “Morality and the Emotions” 226).

Williams correctly argues that *these* reasons assume a crude view of the emotions in which “there is no way of adjusting one’s emotional response in the light of other considerations...without abandoning emotional motivation altogether” (Williams, “Morality and the Emotions” 226). The question then becomes are these the reasons Kant offers?

The ‘No Accident’ Principle. Herman locates Kant’s rejection of the inclinations as moral motives not in the argument that they are capricious or partial, but in the fact that they hit upon the right action *fortuitously*, without any real regard for morality (Herman, *PMJ* 4-5). She offers the example of a person of sympathetic temper who acts firmly, consistently and impartially from the motive of sympathy (meeting Williams’s demand for a sophisticated sentimentalism). During a late night walk, he spots someone struggling with “a heavy burden” at the back door of the Museum of Fine Arts, and immediately offers his aid out of sympathy (Herman, *PMJ* 4-5). Like the friend of humanity, who acts *solely* from immediate inclination, Herman argues that the agent is not concerned with the morality of his actions. The difference is that in Kant’s example, the friend is fortunate to hit on something beneficial and right while in Herman’s example (in which the agent aids an art thief), the same maxim is exposed as lacking moral content. In both cases, the agents are, according to Kant, *morally* indifferent. Herman concludes: “For while sympathy can give an interest in an action that is (as it happens) right, it cannot give an interest in its being right” (Herman, *PMJ* 5).

⁶ “Inclination” and “emotion” are often used synonymously in discussions of Kantian ethics, despite the fact that Kant reserves the term inclination [*Neigung*] more narrowly for a subspecies of desire, which he characterizes as habitual *empirical* desire or “*concupiscentia*” (*MS* 6:212, *R* 6:29). We will ignore this distinction for our purposes here, but as we shall see in Chapter 4, § II, it results in a gross (and invidious) misreading of Kant’s fundamental moral psychology.

At this point of the discussion, Kant's hedonistic theory of empirical desire lands him in some trouble, since it falls prey to Williams's crudeness charge. Kant argues that any desire which *precedes* a practical rule (by presenting an end for it to realize) must be (a) empirical, and (b) based in pleasure (KpV 5:21). He then classifies all such pleasures under the heading of self-love or happiness (KpV 5:22),⁷ and concludes that no inclination can act as a moral motive because "consciousness of virtue is immediately connected with satisfaction and pleasure, and consciousness of vice with mental unease and pain, so that everything is still reduced to the desire for one's happiness" (KpV 5:38). The problem is that a sophisticated sentimentalism need not be committed to the view that makes emotional motivation dependent on an agent's desire for pleasure.

In "Kant's Criticisms of Eudaemonism" T. H. Irwin recognizes the problematic nature of Kant's account, but argues that "it would be a mistake...to assume that Kant has no plausible arguments left if we reject his hedonist assumptions" (Irwin, "Kant's Criticisms of Eudaemonism" 68). If we can locate such an argument in Kant's ethics, then even if we expand the web of emotional motivation to incorporate a host of deliberations or a sophisticated sentimentalism that issues in the right action (so that Herman's agent, in the final analysis, refuses to aid the art thief), the 'no accident' problem persists. The argument can be found in Kant's second *Critique*:

What is essential to any moral worth of actions is *that the moral law determine the will immediately*. If the determination of the will takes place *conformably* with the moral law but only by means of a feeling, of whatever kind, that has to be presupposed in order for the law to become a sufficient determining ground of the will, so that action is not done *for the sake of the law*, then the action will contain *legality* indeed but not *morality*. (KpV 5:71)

⁷ "...all material principles which place the determining ground of choice in pleasure or displeasure...are wholly of the same kind insofar as they belong without exception to the principle of self-love or one's happiness" (KpV 5:22). See also G 4:442: "I class the principle of moral feeling [in moral sense theory] with that of happiness because every empirical interest promises to contribute to our well-being by the agreeableness that something has to offer, whether this happens immediately and without a view to advantages, or with regard for them."

Kant argues that all the objects of inclination, sought as a determining ground of the will, can be “cognized only empirically” (*KpV* 5:23). Because the moral sentimentalist seeks an *object* of the will, experientially, in order to make it into the material foundation of the moral law, “this relation, whether it rests on inclination, or on representations of reason, makes possible hypothetical imperatives only: I ought to do something *because I want something else*” (*G* 4:441). Whether the determining ground is sought materially in pleasure, happiness, a state of contentment or enlightened ataraxia in which virtue is its own reward—by presupposing the former as a *condition* for the moral law, practical reason is, according to Kant, subordinated to judgements of weal (*KpV* 6:63-4). This is why Kant believes that all such attempts lapse into heteronomy. In short, we are confronted not with *moral* judgments of Good and Evil [*Gut* and *Böse*], but rather ones of weal and woe [*Wohl* and *Weh*]. The judgements need not be aimed at pleasure at all. Whatever their aim, the *priority of the inclinations* (sought as a determining ground) to the moral law yields only hypothetical imperatives since it implies that “whether or not we act on moral principles depends on whether we have a sufficiently strong inclination toward the end that we would promote by acting on that principle” (Irwin, “Kant’s Criticisms of Eudaemonism” 68-72). Here even if the inclinations have legality (i.e. they *conform* to the moral law), they are not yet moral since they do not determine the will directly.

Universality / Necessity. This leads us to a second, more general reason for Kant’s rejection of the inclinations as moral motives. Because Kant is convinced by the Humean sceptical argument that no universal or necessary propositions can be inferred experientially (*KrV* B4-5)⁸—his search for the universally obligating, apodeictic principles that he believes morality requires cannot lie in experience, and thus in the inclinations:

⁸ “Experience never gives its judgements true or strict but only assumed and comparative universality (through induction)....Empirical universality is therefore only an arbitrary increase in validity from that which holds in most cases to that which holds in all”—thus, Kant concludes, “Necessity and strict universality are...secure indications of an *a priori* cognition” (*KrV* B4).

Everyone must admit that a law, if it is to hold morally, i.e. as the ground of an obligation, must carry with it absolute necessity...hence the ground of obligation must not be sought in the nature of the human being, or in the circumstances of the world in which he is placed, but a priori solely in concepts of pure reason, and that any other prescription that is founded on principles of mere experience—even a prescription that is in some certain respect universal...can indeed be called a practical rule, but never a moral law. (G 4:389)

In the same section, Kant identifies “universal” principles as those that are normative for all rational beings by virtue of their rationality, and “necessity” as the feature of moral ends to obligate or constrain us as rational beings, regardless of circumstance or desire (G 4:389). “Only the *law*,” according to Kant, “carries with it the concept of an *unconditional* and indeed objective and hence universally valid *necessity*” which morality requires (G 4:416).

In the face of mounting empirical evidence of “fundamental moral disagreement,”⁹ Wood convincingly argues that those who would reject Kant’s moral rationalism and ground morality in the inclinations, are forced to either: (a) claim “a greater empirical uniformity of human nature than experience shows to be there” or (b) embrace cultural relativism (Wood, *KE* 36-7). Since Kant grounds morality, not in theoretical, but in practical reason, he makes no appeals to human nature as such. The universal rationality he speaks of does *not* denote a state of affairs which all rational beings find themselves in, or one that translates into standardized acts across cultures. This explains why, in the *Anthropology*, Kant jettisons the traditional classification of a human being as a “*rational animal* (*animal rationale*)” in favor of “an animal endowed with the *capacity of reason* (*animal rationabile*)” (VA 7:321). We are indeed *capable* of acting rationally, but we fall woefully short of this ideal.

⁹ For a good overview of the empirical evidence of fundamental moral disagreement and the philosophical challenge it poses for moral realism, see John M. Doris and Alexander Plakias, “How to Argue About Disagreement: Evaluative Diversity and Moral Realism,” *Moral Psychology, Volume 2: The Cognitive Science of Morality: Intuition and Diversity*. Ed. Walter Sinnott-Armstrong. (Cambridge: MIT Press, 2008): 303-31.

Practical reason issues in *fundamental* normative ends whose rational authority is not contingent upon culture, desire or circumstance. Being fundamental, the ends conferred *a priori* by reason are not only subject to different interpretations (often fallible and partial) across cultures, but because the individual *in the social condition* is beset by *unsociable* inclinations (envy, jealousy, rivalry, joy in other's misfortunes, etc.), he is "indeed capable of the idea of a practical pure reason, but not so easily able to make it effective *in concreto* in the conduct of his life" (G 4:389), which takes us to Kant's final reason for his rejection of the inclinations as moral motives.

IV. Unsociable Sociability

Kant's debt to Rousseau is well known, but its impact on his practical philosophy remains largely unappreciated. Kant third, and most revealing, reason for the rejection of the inclinations as moral motives stems from his belief that the inclinations are unreliable insofar as the vices "grafted upon" them are "vices of *culture*" (R 6:27). This is important for two reasons.

First, it belies the commonly held view that ascribes to Kant the belief that the inclinations are inherently evil. In what may come as a surprise to many of his critics (and some defenders), Kant explicitly denies any such doctrine: "*Considered in themselves* natural inclinations are *good*, i.e. not reprehensible, and to want to extirpate them would not only be futile but harmful and blameworthy as well..." (R 6:58). We have been arguing that Kant's emphasis on the good will and his mistrust of the inclinations is merely an expression of the Kantian notion that moral considerations should be overriding. If the inclinations are naturally good, why should they fail to act as moral motives or jeopardize the latter?

Kant defines "*unsociable sociability*" as the "propensity [of human beings] to enter into society, which...is combined with a thoroughgoing resistance that constantly threatens to break up this society" (I 8:20). In other words, individuals are drawn to society by their need for others, while simultaneously being repelled by them as obstacles to the fulfillment of their own desires and personal projects (I 8:21).

For Kant, this results in a precarious marriage of *socializing* affinity and *individualizing* social antagonism: the human being seeks “to obtain for himself a rank among his fellows, whom he cannot *stand*, but also cannot *leave alone*” (I 8:21). Now, because the inclinations, as we encounter them in society, are never merely natural, they exhibit this same dichotomy: feelings of love and sympathy draw us closer to one another in weal and woe, while at the same time exhibiting our proclivity to seek a comparative advantage over others. Even feelings of kinship and bonhomie are easily subverted since “we take our stand on the relativities of things” and “even good-natured souls feel grudging” (VE 27:438-9):

We are more than willing to listen, in company, to the tale of another’s misfortunes...or to hear of the downfall of certain wealthy personages, and though we show no satisfaction, it is still privately pleasing to us. When we sit by a warm fire at the coffee-table, during foul and stormy weather, and talk of those who are travelling or at sea in such conditions, we enjoy our good fortune the more for it, and it heightens our sense of comfort. There is thus a grudging element in our nature... (VE 27:439)

This is why Kant maintains that love, like water, “is directed more downwards than upwards” (VE 27:670). According to Kant, we “love everything over which we have a decisive authority” because we...need something to love with which we don’t stand in rivalry. So we love birds, dogs, or a young, fickle and cheerful person. Even love between men and women is partly grounded on this, since no struggle for superiority occurs between people of different kinds. (qtd. in Wood, *KET* 272)

Therefore, the inclinations *in the social condition* are not only unreliable, but they are susceptible to the quotidian corruption of what Rousseau calls *amour propre* [vanity]: “Amour propre is only a relative sentiment, factitious, and born in society, which inclines every individual to set greater store by himself than by anyone else [and] inspires men with all the evils they do one another...” (Rousseau, *The Discourses* 218, n. XV). In light of the former, Kant’s notorious doctrine of radical evil is, in fact, strikingly ordinary. This explains why Kant’s moral examples and table of vices often revolve around ridicule, ingratitude, scandal and other social failings—and not “murder and mayhem” as the title of one of Herman’s essays goes.

If what we have been arguing so far is correct, and Kant's emphasis on the good will concerns the priority of moral considerations over prudential ones, then we should expect to locate Kant's conception of evil not in our natural inclinations, or happiness (which, as we shall see in the next section, he believes we have a rational imperative to pursue), but in our social tendency to *invert* this hierarchy. Kant states the point as follows:

This propensity to make oneself as having subjective determining grounds of choice into the objective determining grounds of the will in general can be called *self-love*; and if self-love makes itself a lawgiving and unconditional practical principle, it can be called *self-conceit* [*Eigendünkel*]. (*KpV* 5:74)

In the sphere of Kant's practical anthropology, this social antagonism born of our unsociable sociability serves to develop our reason through history; in the moral realm, however, it involves the *willful* (and often guileful) inversion of the rational order of incentives so that our inclinations for honour, power and wealth, as well as any superiority we may afford ourselves over others, become the determining ground of the will rather than the moral law. Therefore, as Wood points out, what Rousseau calls *amour propre*, Kant refers to as "unsociable sociability" (*I* 8:20), "self-conceit" (*KpV* 5:74), or the "radical propensity to evil" (*R* 6:28-32).

Whatever name we assign to it, what Kant's doctrine of radical evil amounts to "is simply the claim that morally good people put the moral law first, while the rest of us don't....even if it does not count as one of the most profound statements of moral psychology to come from a philosopher's pen" (Louden, "Evil Everywhere" 103). We will have much more to say about this when we discuss the inclinations in Chapter 4, § IV.

For now what is important to note is that Kant's opposition to self-love is not an opposition to the claims of happiness. In line with our earlier comments about the good will, Kant's chief and rather uncontroversial point is that moral law should not be subordinated to self-love.

V. The Imperatives and Happiness

In the introduction we saw that many of Kant's critics, especially from the field of virtue ethics, object to what they consider to be Kant's oppositional attitude towards human happiness, which they see as sacrificed to the empty form of the moral law. For example, in *From Morality to Virtue*, Michael Slote takes issue with a putative "self-other asymmetry" in Kantian ethics: "On Kant's view...we have an obligation to benefit or contribute to the happiness of other people, but no parallel obligation to seek our own well-being or happiness" (Slote, *From Morality to Virtue* 11).

An overemphasis on the *Groundwork* and FUL has undeniably contributed to these mischaracterizations of Kant's ethics, but even within the former, Kant not only "presupposes safely and a priori" that happiness is an end "in every human being, because it belongs to his essence" (G 4:415), but he also furnishes us with a rational imperative to pursue it.

All practical reason, for Kant, involves the setting of ends (G 4:437; MS 6:384-5). To set an end is to place oneself under a normative commitment or plan for achieving that end:

Everything in nature works according to laws. Only a rational being has the capacity to act *according to the representation* of laws, i.e. according to principles or a *will*. Since *reason* is required for deriving actions from laws, the will is nothing other than practical reason. (G 4:412)

It follows that the will can be determined by objective laws of reason to which our subjective interests are sometimes opposed. Kant refers to this determination in conformity with objective laws as "*necessitation*" (G 4:413). The representation of a necessitating objective principle is a command which he identifies as an imperative. All imperatives are normative in nature since they express an "*ought*": "They say that to do or to omit something would be good, but they say it to a will that does not always do something just because it is represented to it that it would be good to do it" (G 4:413).

This is noteworthy for two reasons: (a) Kant views all imperatives as *normative* in nature and, as such, (b) every imperative, including hypothetical ones, aims at something *good* (G 4:413-4)—that is,

something which we have a practical reason to pursue even in the face of contrary inclinations. For Kant, this is not contingent on any moral dimension: “to set an end is to judge it as *good*, which means: as *worth pursuing*” (Wood, *KET* 55).

If the end is good as *a means to something else*, the imperative is hypothetical. If the end is good *in itself*, the imperative is categorical. Kant identifies two classes of hypothetical imperatives: “*rules of skill*” and “*counsels of prudence*,” and a single categorical imperative, the “*commands (laws) of morality*” (G 4:416).

Rules of Skill. The first hypothetical imperative involves instrumental or means-ends reasoning: if I want X, I ought to do Y. Kant refers to these as “problematic” or “technical” imperatives (G 4:415-7). He sees these basic rules of skill as permeating our daily endeavours and “the use of means to all sorts of *discretionary ends*” (G 4:415). What is relevant for our purposes is that rules of skill are grounded on contingent ends. This means that it is *rationaly permissible* to refrain from a prescribed action, if we abandon the end in question.¹⁰ The same, however, cannot be said of the counsels of prudence.

Counsels of Prudence. Kant labels imperatives of prudence as “assertoric” or “pragmatic” because he believes they involve an end attributable to all rational beings: happiness (G 4:415-7). Happiness, for Kant, is “the satisfaction of all inclinations” (VE 27:276, 29:599; G 4:399; MS 6:387), an idea for which “an absolute whole is required, a maximum of well-being, in my present and every future condition” (G 4:418). Counsels of prudence, therefore, “prescribe universally to all rational beings the actions that are required to achieve the greatest total satisfaction with life” (Wood, *KET* 65).

The first thing to note is that this immediately establishes a hierarchy which forgoes any technical imperatives that might compromise our happiness. Kant gives the example of an individual suffering from gout who recognizes that he ought to refrain from certain dietary predilections that may exacerbate his ailment in the future (G 4:399). Because prescriptions of prudence entail the collective

¹⁰ This does not impugn the fact that technical imperatives are universally binding on all rational beings. See Chapter 2, § V for a detailed defense of this claim.

satisfaction of all our inclinations, they take precedence over technical imperatives or short-term pleasures, which may individually jeopardize the former. Unlike technical imperatives, however, the end of happiness is not one we are free to rescind,¹¹ since it belongs “safely and a priori” to our “essence” as rational beings (G 4:415). Kant is clear on this front:

The imperatives of prudence do not enjoin under a problematic condition, but under an assertoric universally necessary one, found in all men. I do not say: ‘Insofar as you wish to be happy, you must do this and that’, but ‘Because everyone wishes to be happy, which is presupposed for all, he must observe this’. (VE 27:246)

Readers are often puzzled by Kant’s claim (that happiness can be presupposed *a priori*, in the form of a practical necessity, for every rational being) because they interpret it in one of two ways: either (a) that individuals never prefer short-term pleasures to their overall happiness or (b) that prudential considerations win out on every occasion (i.e. there are no imprudent persons)—both of which are obviously false.

Taken normatively, Kant’s claim is one about *rationality*. Although Kant does not explicitly state the general “prescription of prudence” (G 4:416), Wood suggests the following formulation:

Form an idea for yourself of the greatest achievable sum of your empirical satisfaction (under the name ‘happiness’) and make happiness your end, always preferring it over any limited empirical satisfaction.
(Wood, *KET* 67)

Understood this way, the failure to adopt one’s overall happiness as an end, favouring momentary pleasures at its expense, is a failure of practical reason. In other words, I ought to be prudent because being imprudent is failing to will, according to reason, one’s overall happiness as an end. The prudential ‘ought’ is rational in character and, as such, *rationally obligating*, staking a middle ground between mere desire-satisfaction (technical imperatives) and morality (categorical imperatives).

¹¹ Assertoric imperatives are, of course, still hypothetical in that they prescribe actions as *a means to happiness*, but the end of happiness is not a contingent one.

Commands (Laws) of Morality. At the top of the hierarchy are moral, or categorical, imperatives. Much of the *Groundwork* is dedicated to showing how these *a priori* synthetic practical propositions are possible, but this is not what we are concerned with here. What is significant for us is that, just as prudential imperatives override instrumental ones, moral imperatives override both. This is merely an expression of the fact that the categorical imperative “declares an action to be of itself objectively *necessary...apart from any other end*” (G 4:414-5). This last qualification has been adduced by Kant’s critics as evidence of his “empty formalism” since it suggests that categorical imperatives abstract from *all* ends, leaving only a husk of unalloyed rationalism in its wake. But as we saw earlier, all action, for Kant, involves acting on an end (G 4:437; MS 6:385). So when Kant says that the categorical imperative represents an action as objectively necessary by itself, *without reference to another end*, he cannot mean that it lacks an end.

The key, which we encountered earlier in Kant’s rejection of the inclinations as moral motives, is the idea that moral imperatives command our will *immediately* “without presupposing as its condition any other purpose to be attained by a certain course of conduct” (G 4:416). Kant’s main point is that that laws of morality do not constrain the will relative to any *further* end or purpose. As we shall see in Chapter 2, § II, the end that motivates the categorical imperative is neither one to be produced, nor one which is grounded in the inclinations as a material end. Kant’s language is, once again, revealing: “Practical principles are formal if they abstract from all **subjective** ends; they are *material* if they have these” (G 4:427, emphasis added).¹² In both cases, an end is being acted on—the difference lies in the *nature* of the end.

If what we have been arguing so far concerning the intermediate place of happiness in Kant’s hierarchy of imperatives is correct, then criticisms of his treatment of happiness in the *Groundwork*

¹² See also *KpV* 5:20-8.

must be modified as follows: in Kantian ethics we have no direct *moral* obligation to seek our own happiness. Slote acknowledges this when he writes:

[A]ccording to Kant there are no categorical imperatives of self-interest, no...moral obligations to seek one's own well-being or happiness of the sort he believes he can demonstrate with people other than the agent....Even if Kantianism allows or were to allow for the conditional value and virtue status of certain self-regarding prudence, such value exists...at a less fundamental level than the moral value of pursuing the happiness of others. (Slote, *From Morality to Virtue* 28)

Notice that charges such as these are considerably weaker, and far less persuasive, than those that claim Kant has *no* place for human happiness: they take issue with Kant's belief that the rationally obligating claims of happiness are (even in our pursuit of another's happiness) subject to morality as a *limiting condition* and, as such, are of a lesser fundamental value. We will have more to say about this complaint when we turn to William's hypothetical Gauguin example in Chapter 4, § V.

VI. The Formula of Universal Law

Immediately after his discussion of the three imperatives, Kant arrives at the first formulation of the categorical imperative, the Formula of Universal Law:

FUL: *"Act only according to that maxim through which you can at the same time will that it become a universal law"* (G 4:421).

Endless literature and commentary on Kant's ethics is dedicated to discussions of FUL and its universalizability tests which are often shown to yield false positives and negatives by means of creative "*maxim-fiddling*" (Sneddon, "A New Kantian Response" 67).¹³ Instead of delving into this vast field, it is

¹³ Kant's critics in this regard are by no means limited to the field of virtue ethics but range from Hegel and Herder to Mill and Sidgwick, among others. For some well-known treatments of FUL and its universalizability tests among Kant's defenders see Onora O'Neill, "Consistency in Action," *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Ed. Paul Guyer. (New York: Rowman & Littlefield Publishers, 1998): 103-31; Christine M. Korsgaard, "Kant's Formula of Universal Law," *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996): 77-105; and Barbara Herman, "Moral Deliberation and the Derivation of Duties," *The Practice of Moral Judgement* (Cambridge: Harvard University Press, 1996): 132-58.

best to examine the problems that plague FUL through Kant's own examples, which we can then revisit once we have made our case for the primacy of the Formula of Humanity in Kant's ethical project.

However, before moving on to Kant's four examples, these are three aspects of FUL that have contributed to the recurring difficulties highlighted by Kant's critics.

Formalism. First, FUL is *formal* in nature. As we saw, Kant arrives at the formula after having "robbed the will of all impulses that could arise from following some particular law" (G 4:402)—leaving only the form of the law (i.e. universality) to serve as its principle. Kant's aim is to find a practical law valid in form to act as the "formal supreme determining ground of the will regardless of all subjective differences" (KpV 5:32). Being cognizant of the formal nature of FUL (understood *in abstracto*) and the difficulty in applying it *in concreto*, Kant immediately introduces the Formula of Universal Law of Nature as a variant of FUL to bring it closer to common understanding, and make it easier to apply:

FLN: "So act as if the maxim of your action were to become by your will a **universal law of nature**" (G 4:421).

Where Kant's critics err is in viewing the formalism of FUL as encapsulating his entire moral philosophy, which takes us to our second point.

Provisionality. FUL is provisional, marking Kant's first step in his systematic presentation of the moral law in the *Groundwork* as he proceeds scholastically from the *form* to the *matter* and, finally, the *complete determination* of the categorical imperative (G 4:436). Viewed independently of this progression, FUL remains silent about both the moral content of the law as well any positive duties it might entail. Thus, Wood writes, "No doubt Kant would have proceeded in a more reader-friendly manner (and provoked fewer charges of 'empty formalism') if he had dealt from the start directly with the content of moral value, rather than proceeding...from form through matter to...its systematic completeness" (Wood, KET 77). Wood also assigns Kant's critics their fair share of the blame:

These criticisms [of ‘empty formalism’] completely ignore the fact that Kant does not provide merely a formal principle, but proceeds immediately to specify the substantive value on which the principle rests. Hence such critics behave as though they had suddenly stopped reading the *Groundwork* in the middle of the Second Section. (Wood, *KET* 107)

Only midway in the Second Section of the *Groundwork* (G 4:428), do we arrive at the fundamental value or “matter” underpinning Kantian ethics. Thus Wood is right when he complains that “it is deplorably common to regard FUL and FLN...as the chief, if not the only, formulation of the moral law....That seems...almost as misleading as giving the name ‘Newtonian Physics’ to the law of inertia—as if there was nothing else to Newtonian mechanics besides the First Law of Motion” (Wood, *KE* 69).

Inability to Furnish Positive Duties. As a universalizability test, FUL can only establish the permissibility or impermissibility of acting on specific maxims—it cannot furnish a positive duty¹⁴ of any kind to follow such maxims. For example, from the fact that it is morally impermissible to make promises that we do not intend to keep, it does not follow that we have a positive duty to make promises we intend to keep. Nor does it follow that we must keep all the promises we make, since there may be universalizable maxims in which it is permissible to break our promises (this is possible because FUL tests individual *maxims*—not actions of a certain kind).

If FUL is strictly formal, provisional and negative in nature, then we should expect Kant’s critics, as well as his defenders, to encounter fatal difficulties and lacunas if they view his entire moral philosophy through FUL as the chief formulation of the moral law. This is exactly what we find when we turn to Kant’s first treatment of his four examples.

VII. Applying FUL/FLN: The Four Examples

Before applying FUL/FLN to his four examples, Kant’s alludes to a taxonomy of duties of virtue that he will present in the *Metaphysics of Morals*, which divides duties we have to ourselves and to others into

¹⁴ By “positive duties” we mean *both* perfect and imperfect duties. In other words, FUL and its variant, FLN, are “tests of maxims, not of action-kinds, still less of omission-kinds” (Wood, *KET* 100).

perfect and imperfect duties (G 4:421). The first two cases Kant discusses are meant to be examples of perfect duties, while the last two represent imperfect duties.

1. *Suicide*. Kant asks us to imagine someone “weary of life because of a series of ills that has grown to the point of hopelessness” who asks himself if whether it is “contrary to duty” to take his own life (G 4:422). Kant formulates his maxim as follows: “from self-love I make it my principle to shorten my life if, when protracted any longer, it threatens more ill than it promises agreeableness” (G 4:422). Applying FLN, Kant then asks if this principle could become a universal law of nature—to which he answers that it could not, because “a nature whose law it were to destroy life itself by means of the same sensation the function of which it is to impel towards the advancement of life, would contradict itself and thus not subsist in nature” (G 4:422).

It is immediately clear that to arrive at Kant’s contradiction in conception there is an enthymeme at play: why does a law of nature which destroys life from self-love contradict itself?

In the *Groundwork*, Kant argues that “in the natural predispositions of an organized being, i.e., one arranged purposively for life, we assume as a principle that no organ will be found in it for any end that is not also the most fitting for it and the most suitable” (G 4:395). The contradiction arises, for Kant, when self-love, *which has the function of furthering life* (in his natural teleology), is used to shorten life when it “threatens more ill than it promises agreeableness.” But why should we accept this assumption?

Kant’s argument is made weaker still when we see that even if we accept a teleological or purposive view of nature, it is far from evident that the “natural purpose of self-love is to extend the life of the individual organism as far as possible under all circumstances” (Wood, *KET* 86). Even Paton, who embraced Kant’s teleology, regarded this as “the weakest of Kant’s arguments” and concluded that “unless we have an exaggerated idea of the perfection of teleology in nature, unless indeed we commit ourselves to some theory of the working of divine Providence, this argument can carry little conviction except to those already convinced” (Paton, *CI* 154).

2. *False Promises*. This is Kant's most widely discussed example, and one subject to competing interpretations among his own defenders. By examining the two leading interpretations we can bring to light some of the systemic inadequacies of FUL and FLN.

An individual in need of money cannot obtain a loan unless he promises to repay it at a determinate time. He asks himself if it is permissible to make a false promise (of repayment) in order to secure the loan. His maxim reads: "when I believe myself in need of money I shall borrow money, and promise to repay it, even though I know it will never happen" (G 4:422). Kant argues that the maxim would contradict itself if universalized since "it would make the promise and the end one may pursue with it itself impossible, as no one would believe he was being promised anything, but would laugh about such an utterance, as a vain pretense" (G 4:422). The question, once again, is what is the nature of the contradiction that arises? Korsgaard identifies three different interpretations (Korsgaard, *CKE* 78). We will examine the main two.¹⁵

The "Logical Contradiction" interpretation, championed by Dietrichson and Wood, locates the contradiction in a logical impossibility in which the universalized maxim undercuts itself, making the proposed action impossible. In this case, promises themselves would be impossible since there would be no ground for ever believing them.

Under the "Practical Contradiction" interpretation, defended by O'Neill and Korsgaard, the contradiction arises because the universalized maxim frustrates our primary goal, here obtaining a loan, since we are unable to act on our original maxim in a world in which false promises are universalized. The practical interpretation involves a step beyond the possibility of promises and centers instead on their *efficacy*—which is defeated, thwarting our initial purpose, once the maxim is universalized.

Whichever of the two interpretations we adopt, if FUL and FLN are to provide a reliable and accurate method for the assessment of maxims, as their supporters contend, we should expect them to

¹⁵ The third alternative which we will not consider is Paton's "Teleological Contradiction" interpretation.

deliver correct moral outcomes for normal cases via their universalizability tests. Consider the following two maxims, the first from Herman, the second from O'Neill:

M1: In order to avoid crowded theatres at 8:00 p.m., I will go to the movies at 6:00. (Herman, *PMJ* 139)

M2: "I will buy a clockwork train, but never sell one." (qtd. in Wood, *KET* 105)

M1 fails the practical interpretation since, if everyone went to the movies at 6:00 in the evening, this would frustrate the purpose of the original maxim (to avoid crowded theatres). Herman refers to these as "coordination or timing maxims" (Herman, *PMJ* 139). Movie-going is not shown to be impossible in the universalized maxim, but the *efficacy* of going to the movies at 6:00 p.m. is defeated.

M2, on the other hand, fails the logical interpretation since if no one sold a clockwork train, no one could ever buy one: the action itself is made impossible. Both maxims are, however, harmless, reflecting normal cases that should pass FUL and FLN, yet yield false negatives. Notice that neither the logical nor the practical interpretations can deliver the correct moral outcome in both cases. As a result, we are left with the false conclusion that M1 and M2 are on a par with a maxim of deceitful promising: all three are *impermissible*. This is not only devastating for those who argue for a derivation-of-duties model based on FUL and FLN, but it also points to persistent and deep-seated flaws in the latter.

3. *Resting Talents*. A man harbours a certain talent by which he could make himself "a useful human being in all sorts of respects", but prefers to give himself up to gratification and idle amusement rather than cultivate his talent (*G* 4:423). He asks himself whether this "agrees with what one calls duty" (*G* 4:423). Kant doesn't formulate his maxim specifically but we can imagine it being as follows: I will ignore my natural gifts in order to pursue a life devoted to idleness, amusement and pleasure. Unlike the previous two examples which involved contradictions in conception, Kant believes that universalizing this maxim involves a contradiction in *willing* (in which we can conceive, but not will, that the latter become a universal law of nature).

Kant's offers us this pithy explanation in defence of his conclusion: "For as a rational being he [the agent of the maxim] necessarily wills that all his capacities in him be developed, because they serve him and are given to him for all sorts of purposes" (G 4:423).

If we focus solely on the last clause, the spectre of natural teleology rears its head, and we can, once again, question Kant's telic assumptions. But this is not the same argument that Kant appealed to in his suicide example. His main point is rather that a *rational* being cannot universally will the neglect of her rational capacities and talents in order to live a life of idle pleasure. The question is why?

If there is some rational end of substantive value or absolute worth that the agent would be neglecting (in ignoring her talents), Kant has yet to provide us with one. Within the strict parameters of FUL and FLN, he has no recourse to any such principle. This point is quite significant, since we will be arguing for the primacy of the Formula of Humanity, which Kant appeals to here *prematurely*. Outside such an appeal, however, FUL and FLN fail to deliver the contradiction Kant has in mind.

After expounding Kant's rusting talents example in her Cambridge Commentary on the *Groundwork*, Sally Sedgwick puts the point as follows: "Kant has not yet made explicit what our specifically rational ends are; it is therefore impossible to comprehend his line of reasoning here. He will begin clarifying the nature of our rational ends at [G 4:]428" (Sedgwick, *Kant's Groundwork* 120). This is the section in which Kant derives the Formula of Humanity from rational nature as an end in itself.

4. Refusing to Succour Those in Need. In Kant's fourth example, a person who is prospering while others struggle, thinks to himself, "What's it to me?" (G 4:423). His maxim reads: "May everyone be as happy as heaven wills, or as he can make himself, I shall take nothing from him, not even envy him; I just [will not contribute]...anything to his well-being, or his assistance in need!" (G 4:423). Like the previous example, Kant argues that universalizing the former results in a contradiction in willing, not conception (since we can conceive of a world in which no one helped another, yet not *will* to live in such a world). Kant explains: "[For] many cases can yet come to pass in which one needs the love and compassion of

others, and in which, by such a law of nature sprung from his own will, he would rob himself of all hope of the assistance he wishes for himself" (G 4:423).

Wood argues that Kant's fourth case is easily misunderstood. It is *not* a prudential argument (that some unabashedly individualistic agent may choose to opt out of) despite this reading being forced upon Kant, nor does it commit a "temporal fallacy" that argues speciously from the agent needing aid *someday* to her refusal to will the universalized maxim *today* (Wood, *KET* 93-4).¹⁶

Instead, Wood interprets the claim as a normative one regarding human interdependency (albeit drawn from empirical grounds). Under this reading, it would be irrational given our vulnerability and the precariousness of the human condition to live in a world in which no one succoured those in need. The Kantian idea, however, of shared kingdom of ends, in which "we make ourselves an end for others" (*MS* 6:393)—is, once again, not one that Kant can appeal to until he has discussed the *complete determination* of the moral law, and its corresponding formula.

Without this rational end (specified by the Formula of Humanity) and its relationship vis-à-vis other rational beings (specified by the Formula of the Kingdom of Ends), FUL and FLN have purchase only on the prudential argument from self-interest that fails to deliver Kant's desired conclusion, since a robust libertarian, for example, might easily reject it.

VIII. Beyond FUL/FLN

We have been arguing that the devastating problems that plague Kant's initial treatment of his famous four examples are endemic to FUL and FLN. Specifically, they stem from the formalism, provisional nature, and inability of the two formulas to establish any duties (to do or omit a general kind of action). Kant's critics are therefore justified in the charges of "empty formalism" and "rigorism" they level at FUL and FLN (and his defenders locked into a hopeless apologetics)—*if they continue to ignore the merely*

¹⁶ "The issue Kant means to pose is...whether it would be rational for me *at any time* to will to deprive myself *for all times* (by a universal law of nature) of every possibility of voluntary help from others. Kant's contention is that, human beings and human life being what they are, this could not be rational" (Wood, *KET* 93).

formal, provisional, and negative nature of the latter as well as *the final form of Kant's moral philosophy* which we turn to next.

Kant employs FUL and FLN in the *Groundwork* to highlight an important and recurring point about the moral law that we saw both in our treatment of the inclinations, as well as the various imperatives: that we should not “take the liberty of making an *exception* to it [the moral law] for ourselves, or (just for this once) to the advantage of our inclination” (G 4:424). In Kant’s doctrine of radical evil, the term used to denote this is “*self-conceit*” (KpV 5:74)—in Kant’s hierarchy of imperatives, it involves a willful inversion of the moral order, in which prudential imperatives (illicitly) assume categorical form.

Fortunately for Kantian ethics, Kant has a great deal more to say regarding the primacy of the Formula of Humanity, the virtues and practical anthropology before we gauge his ethics against the charges of his critics.

Chapter 2: Rational Nature and the Formula of Humanity

In Chapter 1, we looked at some of the most recognizable and oft-discussed aspects of Kant's ethics in the *Groundwork*. Specifically, we argued for the following deflationary hypotheses: (a) the good will, far from being central to Kant's ethics is only one expression of Kant's claim that moral considerations should be overriding; (b) Kant's emphasis on duty and his notorious mistrust of the inclinations, especially in light of his discussion of the *form* of the moral law, is merely a corroboration of this thesis; (c) what Kant is ultimately chary of are not the claims of happiness or self-love, but the willful *inversion* of the hierarchy of imperatives so that prudential considerations become overriding—that is, not subject to morality as a limiting condition. Kant's term for this reversal of imperatives is "self-conceit." And, finally, (d) the Formula of Universal Law (FUL) is the first step in Kant's presentation of the moral law. Being merely formal, provisional, and negative in nature, when viewed in isolation, it suffers from congenital defects that fail to secure the conclusions that Kant's four examples are meant to deliver.

In what follows, we present the foundation of Kant's Formula of Humanity and argue for its central role in providing the basic substantive value in Kantian ethics. After examining the arguments and implications of Kant's formula, we respond to the main objections to it in Section V. In Chapter 1, we examined the *form* of the moral law. Now, we turn to its *matter*.

I. The Formula of Humanity as an End in Itself

After his discussion of material principles, which are based on subjective ends, and his insistence that practical principles "abstract from all subjective ends" (G 4:427), Kant asks us to consider a different kind of end: "...suppose there were something *the existence of which in itself* has an absolute worth, that, as an *end in itself*, could be a ground of determinate laws, then the ground of a possible categorical imperative, i.e. of a practical law, would lie in it, and only in it" (G 4:428).

This key passage establishes two important points. First, the charge of empty formalism tells us that Kantian ethics is empty because “it detaches the notion of duty from the notion of ends” and thus “is so formal that it can be given almost any content” (MacIntyre, *A Short History of Ethics* 191), but after having introduced FUL, Kant immediately turns to the *content* or objective ground of the moral law. Far from being empty, it is this end which commands both our respect and adherence to the categorical imperative. Since Kant defines ethics as “the system of the *ends* of pure practical reason” Kantian ethics, *pace* MacIntyre, intimately links duties to ends: “that ethics contains duties that one cannot be constrained by others...to fulfill follows merely from its being a doctrine of *ends*” (*MS* 6:381).¹⁷ We are neither free to give the moral law any content of our choosing, nor to repeal the end in question since it is, unlike our subjective ends, not one effected by us.

Second, Kant ascribes three features or “value conceptions” to the end discussed in the passage, which Wood outlines as follows: (a) it is an “*end in itself*”, (b) it is an “*existent end*” and (c) it has “*absolute worth*” (Wood, *KET* 115).

An end in itself is one that is unconditional, representing “*objective ends*, i.e. entities whose existence in itself is an end, an end such that no other end can be put in its place” (*G* 4:428). As such, it applies to all rational beings regardless of their nature, desires or circumstances. Since Kant argues that this is not a hypothetical or subjective end worth having only for us, it “must here be thought [of] not as an end to be effected *but as an independently existing end*”—i.e., one to be “esteemed in every willing” (*G* 4:437). Drawing on Kant’s distinction between ends effected by us and an existent end, Wood describes the latter as “something to be esteemed, preserved and furthered” (Wood, *KET* 115). Finally, Kant ascribes absolute worth to this existent end-in-itself in the form of *dignity* [*Würde*]. That which has

¹⁷ See also *MS* 6:385: “[B]ecause this act which determines an end is a practical principle that prescribes the end in itself...it is a categorical imperative of pure practical reason, and therefore an imperative which connects a *concept of duty* with that of an end in general.” For more supporting texts, see Chapter 3, § IV.

dignity, for Kant, “is elevated above any price, and hence allows of no equivalence”: it cannot be measured against, or replaced with, something else “without...violating its sanctity” (G 4:434-5).

The question we are left is: what is this (a) existent (b) end-in-itself that has (c) absolute worth, or dignity? In short, what end satisfies the three value conceptions? Kant’s answer is found in the Second Section of the *Groundwork*:

“[R]ational nature exists as an end in itself.” (G 4:429)

Kant locates the fundamental value of Kantian ethics, which serves as the determining ground of the moral law, in humanity or rational nature as an end in itself. He asserts that every human being and “generally every rational being *exists* as an end in itself, *not merely as a means* for the discretionary use for this or that will, but must in all actions, whether directed towards itself or also to other rational beings, always be considered *at the same time as an end.*” (G 4:428) Tidying up the assertion in the form a practical imperative, the Formula of Humanity reads:

FH: *“So act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means.” (G 4:429)*

Before discussing his derivation of FH, we should briefly examine the implications of Kant’s formula.

II. The Radical Egalitarianism of FH

Today vague appeals to the “equality of all” have become commonplace, while our interest in the moral and economic implications of the notion has waned. Wood has rightly pointed out that, given this fact,

we should not fail to hear in the phrase [the ‘dignity of humanity’] Kant’s defiant and paradoxically egalitarian assertion that the highest possible worth any human being can have consists in the value that all human beings have equally—whether well born or ill born, rich or poor, intelligent or stupid, even good or evil. (Wood, *KE* 94)

One way of appreciating the radical egalitarianism of FH is by surveying Kant's three original "predispositions" [*Analgen*], or "fundamental global capacities" of human nature outlined in the *Religion* (Wood, *KET* 118). They are: (1) "*Animality*," (2) "*Humanity*" and (3) "*Personality*" (R 6:26).

Animality refers to the individual as a *living being* and includes self-preservation, the sexual drive (ensuring the propagation of the species) and the social drive responsible for bringing human beings together in communities.

Humanity refers to the individual as a *rational being*, to wit, a being capable of setting ends, selecting the means to them and acting on them (individually or with regard to her overall wellbeing), *without any specific reference to morality*. Kant describes humanity as "rooted in reason which is indeed practical, but only subservient to other incentives" (R 6:28).

Personality, on the other hand, refers to the individual as a *responsible being* and is defined by Kant as the "susceptibility to respect...the moral law as *of itself a sufficient incentive to the power of choice*" (R 6:28). In other words, personality is the predisposition "rooted in reason practical of itself" which characterizes individuals as moral beings (R 6:28).

In the *Groundwork* Kant equates dignity with personality: "Thus morality and humanity, *in so far as it is capable of morality*, is that which alone has dignity" (G 4:435, emphasis added). Many defenders of Kant as well as philosophers who appeal to his moral philosophy in their positions have taken this to mean that, for Kant, eligibility for moral consideration is contingent upon personality.¹⁸

This widespread idea often attributed to Kant (that personality or moral agency is a condition, or at least one of the necessary conditions, for moral consideration) is, tellingly, *not* his position. Despite equating dignity with personality, Kant identifies humanity, not personality, as an end in itself.

If we return to the three predispositions of human nature outlined by Kant, we can see precisely where the egalitarian nature of FH lies. The end that satisfies Kant's three value conceptions is identified

¹⁸ Richard Dean takes this position: "The aspect of rational nature that marks a being as an end in herself...is the aspect related to legislating and acting on moral principles" (Dean, "The Formula of Humanity" 86).

by him as “*rational nature...as an end in itself*” (G 4:429). This corresponds to the predisposition of humanity which refers to the individual as a “*rational being*” (R 6:26).¹⁹ The crucial point is that the intermediary position of humanity involves rational nature *in all its broad functions* without any specific reference to morality. Therefore the end that “must...be esteemed in every willing” is not moral agency (which belongs to personality), but the rational capacity to set ends, understood broadly (G 4:437).

Korsgaard sums up the relation as follows:

When Kant says that the characteristic of humanity is the power to set an end, then, he is not merely referring to personality, which would encompass the power to adopt an end for moral or sufficient reasons. Rather he is referring to a more general capacity for choosing, desiring, and valuing ends....the distinctive feature of humanity, *as such*, is simply the capacity to take a rational interest in something: to decide, under the influence of reason, that something is desirable, that it is worthy of pursuit or realization....It is this capacity that the Formula of Humanity commands us never to treat as a mere means, but always as an end in itself. (Korsgaard, CKE 114)

If there is any textual ambiguity regarding Kant’s own characterization of rational nature in the *Groundwork*, there is none in the *Metaphysics of Morals*:

The capacity to set oneself an end—***any end whatsoever***—is what characterizes humanity (as distinguished from animality). Hence there is also bound up with the end of humanity in our own person the rational will, and so the duty, to make ourselves worthy of humanity by culture in general, by procuring or promoting the *capacity* to realize ***all sorts of possible ends***, so far as this is to be found in the human being himself. (MS 6:392, emphasis added)

What follows from Kant’s emphasis on the fundamental value of humanity over personality may shock even his defenders: because Kant characterizes humanity as a “*pragmatic predisposition*” (VA 7:323-4), that is, as one involving the counsels of prudence (see Ch. 1, § V), “the absolute worth that grounds morality is the predisposition towards prudence (rational self-love and the end of our own

¹⁹ This is why Kant uses “humanity” and “rational nature” interchangeably. See, for example, G 4:428-31.

happiness), rational social interaction, and the cultivation of ourselves and all our faculties through society” (Wood, *KE* 88). This corresponds, as we shall see in the next chapter, to the two obligatory ends that follow from FH: our own perfection and the happiness of others (*MS* 6:392-5).

The question now arises: given the fact that Kant identifies the end in itself with humanity why does he reserve dignity for personality? The key lies in the relationship of the two predispositions vis-à-vis the idea of freedom. Personality, which is “*lawgiving of its own*” represents the positive freedom to legislate or give oneself the moral law, whereas humanity, as “*independence*” from alien causes represents the negative freedom to freely set ends (*KpV* 5:33). Kant argues that positive freedom “flows from” negative freedom (*G* 4:446). Thus the foundation for personality is rational nature in all its broad functions, which finds its highest expression in the former. Korsgaard puts the point as follows: “Humanity, completed and perfected, becomes personality, so that in treating the first as an end in itself we will inevitably be lead to realize the second” (Korsgaard, *CKE* 114). Korsgaard believes this will happen because the ends of humanity can only be “perfected” if they are fully determined by reason, and “this occurs only when we respond to moral incentives” (Korsgaard, *CKE* 114). This explains why both Korsgaard and Wood regard humanity and personality as necessarily coextensive. But personality considered *in isolation*, as merely one aspect of rational nature, cannot ground moral imperatives.

In the previous chapter, we saw that Kant’s aim was to find an objective and universally valid determining ground of the will. Kant defines *objective* as “holding for the will of every rational being” (*KpV* 5:19). He also maintains that such a ground can be *universally valid* “only when it holds without the contingent, subjective conditions that distinguish one rational being from another” (*KpV* 5:21).

Now if personality were the sole end in itself, it could not ground the categorical imperative since: (a) it would not hold for the will of *every* rational being, but only for responsible or moral beings (i.e. it would not be objective) and (b) it would be contingent on the moral conduct of individuals, assigning worth only to the virtuous (i.e. it would not be universally valid). This is what Richard Dean

alludes to when he says that such a move “appears to have unpalatably moralistic consequences—it appears to demand judgement of others’ moral character, and differential treatment based on those judgements” (Dean, “The Formula of Humanity” 87).

The radical egalitarianism of FH lies, first, in respecting rational nature *in all its functions* which extends equality to all human beings regardless of their rank, class, race, sex, intelligence or rectitude. Kant is clear about this when he discusses what is owed to the “vicious man” by virtue of his humanity: “I cannot deny all respect to even a vicious man as a human being; I cannot withdraw at least the respect that belongs to him in his quality as a human being, even though by his deeds he makes himself unworthy of it” (MS 6:463).

In an attempt to defuse this reading of the passage, Richard Dean offers three “psychological claims” for treating the vicious man as an end in himself, even if he has not “earned this treatment” (Dean, “The Formula of Humanity” 100). In the same article, Dean maintains that “the most common way to give up one’s rational nature is to choose to act contrary to moral requirements” (Dean, “The Formula of Humanity” 92). But, if the worth of rational nature must be “earned” and can be “given up” as he contends, then it would be neither absolute, nor binding on all rational beings. The fact that the vicious man makes himself *unworthy* of his humanity through his deeds is only possible if he already possesses the humanity in question. Therefore, Dean’s strong sense of rational nature fails to satisfy Kant’s value conceptions across the board.

The second egalitarian dimension of FH is found in the respect it extends to *all rational nature* wherever it may be found. Thus, as Wood points out, FH does not fall prey to Singer’s speciesism charge, since it does not assign worth to humanity on the basis of membership to the human species (Wood,

KET 199-20). Kant states the point as follows: “It [the principle of morality] is...not limited to human beings only but applies to all finite beings that have reason and will” (*KpV* 5:32).²⁰

III. The Derivation of FH

Kant offers two arguments in support of FH in the *Groundwork*: an argument by elimination, followed by a precipitate derivation of FH which has been described by his own admirers as “terse and obscure” (Wood, *KE* 88), “too brief to be intelligible” (Paton, *CI* 176), “remarkably cryptic” and “so compressed as to be largely mysterious” (Dean, “The Formula of Humanity” 90-1).

To aid in Korsgaard’s main reconstruction of the proof, which involves a “regress on conditions,” we will, following Wood, examine Kant’s argument in standard form:

The ground of this principle is: *a rational nature exists as an end in itself*.

- P1:** That is how a human being by necessity represents his own existence; to that extent it is thus a *subjective* principle of human actions.
- P2:** But every other rational being also represents its existence this way, as a consequence of just the same rational ground that also holds for me;
- P3:** Thus it is at the same time an *objective* principle from which, as a supreme practical ground, it must be possible to derive all laws of the will.
- C:** The practical imperative will thus be the following: *So act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means.* (*G* 4:429)

²⁰ Interestingly, Kant thought that it was likely that such finite rational beings existed on other planets. A passage at the end of the *Anthropology*, tinged with Kant’s usual pessimism, reads: “If one now asks whether the human species (which, when one thinks of it as a species of rational *beings on earth* in comparison with rational beings on other planets, as a multitude of creatures arising from one demiurge, can also be called a *race*)—whether, I say, it is to be regarded as a good or bad race, then I must confess that there is not much to boast about in it” (*VA* 7:331). In the next paragraph, he ponders the existence of rational extraterrestrials openly: “It could well be that on some other planet there might be rational beings who could not think in any other way but aloud...” (*VA* 7:332)

P3 acts as an intermediate conclusion drawn from P1 and P2, which appear to be doing the heavy lifting. Both Wood and Korsgaard acknowledge that the argument hinges precariously on P1, which they refer to as the “crucial” premise or step (Wood, *KET* 125; Korsgaard, *CKE* 122). Accordingly, Korsgaard’s “regress on conditions” serves as a clarification and defence of the latter.

Before turning to her argument, however, we can rule out two natural interpretations of P1: under the first interpretation, the premise serves as an empirical observation that people subjectively ascribe value to their existence. This cannot be what Kant means, since there are many cases in which people deem their existence to be worthless, that is, of no value whatsoever. Citing a despairing Macbeth, Alan Donagan reminds us that “a man may not care about rational nature—his own or anybody else’s” (Donagan, *The Theory of Morality* 229-30).

The second natural interpretation we can rule out takes the premise to mean that, when we act, we empirically affirm our existence as an end in itself. Once again, this is obviously false: in the typical course of events, we hardly dwell on the ground of our actions, let alone the idea of an “end in itself.”

The best way to interpret P1, according to Wood, is take Kant as meaning that “there is something in the way people act, and think about their action, which *necessitates* (or commits them to) representing their own existence as an end in itself. The question is: What could that be?” (Wood, *KE* 91). This is where we turn to Korsgaard.

Part I: The Argument by Elimination

Korsgaard’s reconstruction involves a “regress on conditions” or regressive argument: her main contention is that Kant reasons from the value we place on the ends we set, to the ground of ultimate value in the rational nature that set them. The first part of her exposition, however, follows Kant as he eliminates the other possible candidates which might satisfy the three value conceptions examined in Section II. Korsgaard couches Kant’s search for absolute worth, or the end in itself, in terms of the good. She begins with the following question: “Suppose that you make a choice, and you believe what you

have opted for is a good thing. How can you justify it or account for its goodness?" (Korsgaard, *CKE* 121)? Kant considers three possible candidates.

1. *Objects of Inclinations.* We already know that in order to ground the categorical imperative, the end in itself, or source of value, which Kant is seeking must be objectively and universally binding. This end, Kant argues, cannot lie in the objects of inclinations themselves, since "all objects of inclinations have conditional worth only; for if the inclinations, and the needs founded on them did not exist, their object would be without worth" (*G* 4:428). In other words, the worth of objects of inclination is contingent upon our desire for them. Absent that desire (or need), they are without worth. Therefore, they cannot be objectively and universally binding and, as such, the source of goodness we seek.

2. *Inclinations Themselves.* Having eliminated the objects of inclination, we are now tempted to ascribe the source of goodness to the inclinations themselves (since, as we saw above, without them, the objects have no worth)—an idea Kant is quick to reject for "the inclinations themselves, as sources of need, are so far from having absolute worth...that to be free of them must rather be the wish of every rational being" (*G* 4:428). As Korsgaard reminds us, we need not endorse Kant's view to realize that there are some inclinations which we would prefer not to have: "namely those whose existence is disruptive to our happiness" (Korsgaard, *CKE* 121). Consider, as an example, any deleterious habit, such as dipsomania, which involves strong addictions or cravings. This means that there are some inclinations we would be better off without. Given this fact, Korsgaard argues, "it would not be right to say that the object craved was good simply because of the existence of the craving when the craving itself is one that you would rather be rid of" (Korsgaard, *CKE* 121). In short, the goodness of the object cannot be ascribed to *any* inclination for that object, which prompts Kant to conclude that "the worth of any object *to be acquired* by our action is always conditional" (*G* 4:428). The point is captured well by Wood: "even when something is judged good because it satisfies an inclination, it need not follow that the inclination, as the *source* of this value, is something we judge to be good" (Wood, *KET* 123).

3. “*Beings whose existence rests not...on our will, but on nature*” (G 4:428). Having eliminated the first two candidates, Korsgaard arrives at P1. Kant, however, considers a third and final alternative: “non-rational beings” (G 4:428). According to Kant, non-rational beings “have only a relative worth, as means, and are therefore called *things*” while rational beings have worth as ends in themselves are called “*persons*” (G 4:428). And “things” as mere *means*, cannot serve as the ultimate ground of value.

This is not only the weakest of Kant’s claims, but it also seems to be a case of circular reasoning since Kant assumes in his premises what he is seeking to prove, mainly, that it is *rational nature* which exists as an end in itself. The distinction between persons (as ends in themselves) and things (as means), which is crucial to the argument, hinges on the truth of this conclusion—precisely the one Kant is seeking to establish.

For this reason among others, Dean dismisses the argument by elimination as “not compelling” (Dean, “The Formula of Humanity” 91), while Samuel Kerstein says that “on its face, the defense seems inadequate” (Kerstein, “Deriving the Formula” 204).²¹ In light of the former, what Kant needs is a positive argument for FH.

Part II: The Regressive Argument

The question we have been investigating so far is “What makes the objects of our choice good?” Having arrived at P1, or the “crucial step,” Korsgaard proposes the following: “Kant’s answer...is that what makes the object of rational choice good is that it *is* the object of rational choice....His idea is that rational choice has what I will call a value-conferring quality” (Korsgaard, *CKE* 122). Wood endorses this conclusion, with an important qualification (more on this to follow): “[the] rational choice of ends is the act through which *objective* goodness enters the world” (Wood, *KET* 129).

²¹ In the same article Kerstein puts forward a bolstered version of the argument by elimination in support of FH which draws from other parts of Section II of the *Groundwork*. As a stand-alone or “isolated” argument, however, he concedes that it is “open to challenge” (Kerstein, “Deriving the Formula” 202).

If this is correct, then by P1 Kant means that when we set an end “we act as if our choice were a sufficient condition for the goodness of its object” (Korsgaard, *CKE* 123). Of course, this does not mean that everything we desire is good: as we saw in Kant’s argument by elimination we are free to reject certain inclinations, as well as their objects, as bad. The key point, more pronounced in Wood’s account than Korsgaard’s, lies in the capacity to set *rational* ends. Thus, Wood tells us, “to set an end is to attribute objective goodness to it and...we can regard this goodness as originating only in the fact that we have set those ends according to reason” (Wood, *KET* 129).

For the sake of clarity, let us retrace Korsgaard’s argument. She begins with the value we place on the ends we set. An empirical “regress on conditions” yields no satisfactory answer as to what the ultimate source of this value is. Korsgaard concludes: “it must be that we are supposing that rational choice itself *makes* its object good” (Korsgaard, *CKE* 122). If this is correct, then rational nature has a value-conferring quality. As such, rational nature is the prescriptive *source* of all objective value, and the unconditioned end in itself.

In an effort to bolster Korsgaard’s reconstruction of Kant’s proof, Wood highlights two essential points necessary for her argument to go through:

- (1) “First, it [the regressive argument] requires us to concede that setting an end for ourselves involves ascribing *objective* goodness to it.”²²
- (2) “Second, it involves an inference from the *objective* goodness of the end to the *unconditional* objective goodness of the capacity to set an end.” (Wood, *KET* 127)

Both Korsgaard’s reconstruction, as well as Wood’s main contention (that, for Kant, when we set an end we ascribe objective goodness to it) have been met with various objections. Before we turn to these, however, it is worth asking what textual evidence there is for attributing the above positions to Kant. Since Korsgaard’s proof is vague on this front (in fact, it seems to arrive at its conclusion *via* Kant’s

²² Barbara Herman refers to Wood’s first interpretive point, somewhat disparagingly, as “bootstrapping value” (Herman, “The Difference That Ends Make” 104). We review objections to this idea in Section V.

initial argument by elimination—which, as we saw, was not very compelling as a stand-alone argument), some commentators have denied that these positions are attributable to Kant. Kerstein, for example, writes: “I do not deny that it was open to Kant to appeal to the regressive argument or that the regressive argument is philosophically interesting. It is just not, I think, an argument that Kant himself unfurls” (Kerstein, “Deriving the Formula” 201-2). Can we find in Kant’s texts the grounds for a “regress on conditions”?

In the *Groundwork*, Kant equates the will with practical reason. An act of will or practical reason is one done “according to the representation of laws” (G 4:412). As we saw in Chapter 1, § V, all practical reason, for Kant, involves the setting of ends.²³ Hence it should not surprise us that Kant characterizes rational nature strictly by virtue of this capacity: “rational nature is distinguished from others by this, that it sets itself an end” (G 4:437).

To set an end is to place oneself under a normative commitment or plan for achieving that end. Kant refers to the setting of ends in conformity with objective laws as necessitation. These commands of reason which always express an “ought” he calls imperatives.

Kant views all imperatives as *normative* in nature and, as such, each imperative, including hypothetical ones, aims at something *good*—that is, something which we have a practical reason to pursue even in the face of contrary inclinations.²⁴ “they say to do or to omit something would be good, but they say it to a will that does not always do something just because it is represented to it that it would be good to do it” (G 4:413). In the crucial passage that follows, Kant clarifies what he means by “good”: “practically *good*...is what determines the will by means of representations of reason, hence not from subjective causes, but **objectively**, i.e. from grounds that are **valid for every rational being**, as such” (G 4:413, emphasis added).

²³ “An **end** is an *object* of free choice, the representation of which determines it to an action....Every action, therefore, has its end” (MS 6:384-5).

²⁴ “[T]he will is a capacity to choose *only that* which reason, independently of inclination, recognizes as practically necessary, i.e. as good” (G 4:412).

If Kant defines “practically good” as that which is “objectively...valid for every rational being” then to set an end according to the commands of reason, for Kant, is to pursue an end which is *valid for every rational being*—it is to ascribe objective goodness to that end. This is not limited to moral ends, since *all* imperatives, which Kant defines as formulas of the commands of reason, “say to do or omit something would be good” (G 4:413).

This is the fundamental claim, crucial to Korsgaard’s “regress on conditions,” which Wood correctly attributes to Kant: “Kant thinks that when we set an end for ourselves, and consequently regard it as good, we thereby regard it as something of value for all rational beings” (Wood, *KET* 127). Kant then infers from the fact that rational nature is the ground through which objective value enters the world, that it is the “subject of all possible ends” to be esteemed in every willing since “it cannot, without contradiction, be ranked lower than any other object” (G 4:437).

There is, nevertheless, an important qualification vis-à-vis Kant’s inference: the inference is *not*, following Korsgaard’s Kantian constructivism, one that *confers* value on, or “*makes* its object good” (Korsgaard, *CKE* 122). As G. A. Cohen has persuasively argued in his response to Korsgaard’s *Sources of Normativity*, this constructivist reading of Kant (especially given Korsgaard’s focus on practical identity) introduces an element of contingency into her account that is decidedly un-Kantian:

Suppose I ask: *why* should I obey myself? Who am *I*, anyway, to issue a command to *me*? Kant can answer that question. He can say that, although you *legislate* the law, the content of the law, the content of the law comes from reason, not [as Korsgaard would have it] from anything special about you, or *your* reason, or even *human* reason, but from reason as such. And, when that is so, then...reflective endorsement of the law is inescapable. (Cohen, “Reason, Humanity, and the Moral Law” 173-4).

Arguing along similar, realist lines, Wood insists that “Kant’s argument does not involve saying...that setting an end *confers* value on the end. On the contrary, setting an end is an exercise of *practical reason* only to the extent that we think there is already *some good reason* for us to set that end” and this reason, exists “prior to our rational choice” (Wood, *KE* 92). This means that rational nature does not

make its object good, as Korsgaard contends, but is the source through which objective goodness enters the world. Kant stresses this point when he issues the following, terse caveat in the second *Critique*: “we are indeed lawgiving members in a kingdom of morals possible through freedom...but we are at the same time subjects in it, not sovereigns” (*KpV* 5:82). In other words, to say that the goodness of all ends is grounded in the fundamental value of rational nature as an end in itself, is not to say that the latter confers goodness on the former.

Moreover, if Korsgaard’s constructivism were true, then rational nature as the end in itself would be contingent on our practical identity, or any other means by which we confer value on the ends we set. But Kant, as we have seen, requires that the supreme determining ground of the moral law be both objectively and universally valid. Therefore, it cannot be contingent on any aspect of practical identity for the simple reason that “the liaison between morality and practical identity is questionable, since commitments that form my practical identity need not be to things that have the *universality* characteristic of law” (Cohen, “Reason, Humanity, and the Moral Law” 174-5). Now we are in a position to see what Kant means by P1.

The Crucial Premises Explained

As we saw earlier, the crucial premises or steps in Kant’s derivation of FH are the following:

The ground of this principle is: *a rational nature exists as an end in itself*.

- P1:** That is how a human being by necessity represents his own existence; to that extent it is thus a *subjective* principle of human actions.
- P2:** But every other rational being also represents its existence this way, as a consequence of just the same rational ground that also holds for me. (*G* 4:429)

If the regressive argument is correct, then P1 means that in setting an end, rational beings are necessarily committed to representing their existence, however tacitly, as being an end in itself. Why?

Because we cannot regard the capacity for setting ends as capable of determining which ends are good, without esteeming the ground of those capacities as “the subject of all possible ends” (G 4:437).

Wood likens this attitude (of esteeming our rational nature) to the one we take towards an authority figure “as a source of recommendations or commands” (Wood, *KET* 130). If we do not esteem this ground as an authority, we are left with approximations of the truth, or worse yet, mere suggestions. Thus when we set an end according to reason, we are already acting under the *presupposition* that rational nature is an end in itself, capable of determining the good, both as a means and as an end.

Notice that P2 requires precisely this objective reading of P1 which we have been defending (“...a human being *by necessity*”) to go through, since a merely “*subjective* principle of human action” cannot be generalized to “every other rational being” (G 4:429).

Therefore, P2 tells us that rational nature is an end in itself not only in my person but in *every person* by virtue of the “same rational nature that also holds for me” (G 4:429). Once again, if rational nature as an end in itself was the capacity to legislate moral ends (the predisposition of personality), the move would be invalid, since this capacity does not belong to all rational beings. If, on the other hand, the capacity is that of setting ends (the predisposition of humanity), then “it does belong to every rational being as such, to others as well...to stupid and clever people exactly as much as to clever and virtuous ones” (Wood, *KE* 92). From P1 and P2, Kant infers rational nature as an objective principle in P3, which he formulates as FH in the conclusion.

IV. Objections to FH

In the previous sections we were concerned with laying the *foundation* for FH as the central formula, and basic substantive value, at the heart of Kant’s ethics. In effort to present Kant’s arguments as cogently as possible, we temporarily deferred any objections that might arise in favour of exposition. We turn to these objections now, in the order they would have appeared in.

Objection 1: 'Rational nature is ambiguous.'

Kant is unequivocal in his assertion that “*rational nature exists as an end in itself*” (G 4:429). In this there is little, if any, disagreement among Kantian scholars. But as Richard Dean has pointed out, “the term ‘rational nature’ is ambiguous” (Dean, “The Formula of Humanity” 84-5). And since we are arguing for a form of *value monism*, to borrow Wood’s term,²⁵ in which rational nature acts as the single, fundamental value or supreme determining ground in Kantian ethics, we should strive for some clarity in this regard.

Complicating matters further are the highly technical definitions of reason that abound in rational choice theory and other contemporary fields (concentrated, quite tellingly, in economics), as well as Kant’s distinction between theoretical and practical reason, and *Willkür* and *Wille*.

First, both Wood and Thomas Hill argue that Kant’s notion of reason [*Vernunft*] aligns itself, in the spirit of the Enlightenment, with the common, daily use of reason over the rarefied, technical meaning we have come to attribute to it today (Wood, *KE* 12-23; Hill, *HW* 125-63). Understood broadly, reason is the faculty [*Vermögen*] or power [*Kraft*] to think and act according to norms—active primarily in our ordinary, day-to-day affairs. This is why the setting of ends according to reason always involves placing ourselves under a *normative* commitment or plan, and why we can fail to abide by these commands of reason, which take the form of imperatives. Because the faculty of reason is normative through-and-through, it animates all imperatives, including the instrumental and prudential ones which occupy our daily life.

Second, Kant not only distinguishes between theoretical reason and practical reason, which is concerned with the “general determination the will” (*KpV* 5:19), but he also argues for the *primacy* of practical reason over the former (*KpV* 5:119-22). He insists further that “one cannot require pure

²⁵ “For Kant, as for Mill, it is important not only that ethics should rest on objective value but also that it should rest on a *single*, objective value. ‘Value monism’ (as I will call it) naturally recommends itself as the only way of providing a single coherent framework for ethical theorizing, at least on the model of ethical theory represented by Kant and Mill” (Wood, *KE* 59).

practical reason to be subordinate to speculative reason and so reverse the order [of importance], since *all interest is ultimately practical* and even that of speculative reason is only conditional and is complete in practical use alone” (*KpV* 5:120, emphasis added). Therefore rational nature as an end in itself is not concerned with “norms of logic and general principles of empirical understanding, but...with norms of practical reason” (Hill, *HW* 153).

Third, in the *Metaphysics of Morals*, Kant distinguishes between two aspects of the will: *Willkür* and *Wille*. *Willkür* is the capacity to choose and set ends or, in Kant’s (uncharacteristically) breezy definition, “to do or to refrain from doing as one pleases” (*MS* 6:213). The *Wille*, on the other hand, is “the faculty of desire considered not so much in relation to action (as choice is) but rather in relation to the ground determining choice to action (*MS* 6:213). Therefore, the *Wille* is, according to Kant, what legislates or provides us with moral guidance.

Richard Dean identifies rational nature with the *Wille* and concludes, as we have seen, that the “the power of legislating unconditional moral demands is the aspect of reason that marks a rational being as an end in herself” (Dean, “The Formula of Humanity” 86). But Kant equates rational nature with the predisposition of humanity—not personality (*R* 6:26; *G* 4:428-31). And what “characterizes humanity,” according to Kant, is “the capacity to set oneself an end—any end whatsoever...” (*MS* 6:392). This aligns rational nature with *Willkür* (the capacity to set ends), and not *Wille* (the legislative function).

If these three points are correct, then: (a) rational nature is not as ambiguous as its critics claim, and (b) “the Kantian model [of practical reason] which is often dismissed because of its alien terminology or its (supposed) dubious metaphysical commitments, turns out to represent something closer to our ordinary, pretheoretical views than it is usually thought” (Hill, *HW* 128).

Objection 2: 'FH is formalistic.'

One of the problems that plagued FUL was the charge of “empty formalism” which said that the moral law lacks determinate content and therefore cannot provide us with any moral guidance regarding the kinds of actions we ought to engage in or ends we should pursue.

A similar charge can be levelled at FH: despite the fact that the formula is concerned with the matter or content of moral law, it can easily be seen as highly abstract or recondite, especially in its high-sounding injunction to “*use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means*” (G 4:429).²⁶ This is the criticism Sartre wields against FH in his famous example of a young French youth torn between tending to his sick mother and joining the Free French Forces in England:

Kantian morality instructs us to never treat another as a means, but always as an end. Very well; therefore, if I stay with my mother, I will treat her as an end, not as a means. But by the same token, I will be treating those who are fighting on my behalf as a means. Conversely, if I join those who are fighting, I will treat them as an end, and, in so doing, risk treating my mother as a means. If values are vague and if they are always too broad in scope to apply to the specific concrete cases under consideration, we have no choice but to rely on our instincts... (Sartre, *Existentialism is a Humanism* 31-2)

If this charge holds water, then FH is, like FUL, too indeterminate and formal to furnish us with any concrete moral prescriptions.

One way to approach the objection is to view it in light of an example Michael Slote puts forwards against FH in his article “The Problem We All Have with Deontology.” Using the familiar “inhospitable hospital” thought experiment in which a brilliant surgeon can save the lives of five automobile victims by harvesting the organs of a derelict asleep in the emergency corridor, Slote asks,

²⁶ Kant anticipated this objection when he wrote, “in this idea of the absolute worth of a mere will...there is something so strange that, regardless of all the agreement with it even of common reason, a suspicion may yet arise that it might perhaps covertly be founded merely on some high-flown fantastication” (G 4:394).

“Does the surgeon use the derelict solely as a means?” His answer: “I don’t think she does, or at least I don’t think she necessarily does...” (Slote, “The Problem We All Have” 263).

If we apply to the “inhospitable hospital” example the idea proposed by Wood that “every argument from FH depends on an intermediate premise, logically independent of FH itself, which tells us what the action expresses or fails to express concerning the worth of humanity in someone’s person” (Wood, “Humanity As End in Itself” 180)—we get the following Kantian argument:

P1: FH: “*So act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means.*” (G 4:429)

P2: The act of killing a derelict in order to harvest his organs to save the lives of five others is inconsistent with treating humanity in his person as an end in itself.

C: Therefore, one must not kill a derelict in order to harvest his organs to save the lives of others.

Slote offers his counterexample (to the Kantian argument above) to show that FH, like FUL, “offers us no way of to justify the deontological refusal to kill the derelict in such a case” and, as such, “the Formula of Humanity cannot be used to justify deontology” (Slote, “The Problem We All Have” 264).

Once the argument is before us in standard form, however, we see that what is being challenged by Slote is *not* FH, but the intermediate premise, P2: Slote believes that killing the derelict in order to harvest his organs in the given example *is* consistent with treating humanity in his person as an end. The surgeon, he tells us, “might still value the derelict’s life” and, on many previous occasions, “may have brought him food and a blanket” (Slote, “The Problem We All Have” 263). But FH is logically independent of P2 (since we can, as Slote does, reject the conclusion simply by rejecting P2). Thus disagreements about P2 or which actions are consistent or inconsistent with treating humanity as an end in itself do not, *ex hypothesi*, impugn FH as a “high-flown fantastication” (G 4:394).

We can choose to regard the controversies surrounding the *application* of FH as reflections on the indeterminacy of the formula itself but “a more plausible way to look at the matter...is to insist that

the meaning of the FH is clear and determinate, because the concept of humanity as an existent end in itself possessed of a dignity is clear and determinate” (Wood, “Humanity As End in Itself” 181)—and that the disputes that arise do so because every application of FH requires an *intermediate premise*, which is, and always should be, subject to critical revision.

Moreover, as Wood shrewdly notes, FH cannot be formalistic or incapable of providing us with moral guidance, because its own content and prescriptions are used as fodder against it. For example, those who believe that rationality does *not* afford human beings more worth than animals or that rational nature does *not* place us on an equal footing with other people—all deny FH (Wood, “Humanity as End in Itself” 182).

Consider Tom Regan’s rejection of Kant’s “indirect duty view” towards animals: “on Kant’s view...beings who exist but are nonrational have ‘only relative value’ and thus fail to be made ends in themselves. Because they fail to have independent value, we have no direct duty to treat in accordance with the Formula of End in Itself...” (Regan, *The Case for Animal Rights* 177). Regan ultimately rejects FH because he believes that animals fall outside its purview. The rejection is based not on FH’s formalism, but on the *moral implications* that follow from its adoption.

Moreover, Wood suspects that the charges of vagueness and formalism against FH are often “far more problematic than an honest demand for philosophical clarity” (Wood, *KE* 57). As an example, he cites Bush’s response to a question about the Geneva Convention prohibition on “outrages upon human dignity”: “That’s very vague! What does that mean?” (qtd. in Wood, *KE* 57).

Finally, in an eloquent passage responding to the formalism critique while underscoring the radical nature of FH, Wood suggests that

our impression that FH is too abstract to be easily applicable might best be understood by seeing that FH is rather like the Sermon on the Mount, the principles of communism, or other admirable utopian visions of human society, whose demands require such a radical departure from our customary practices and

accepted attitudes toward ourselves and others that we are at first perplexed when we try to apply them.

(Wood, *KET* 139)

As we shall see in our discussion of the kingdom of ends in Chapter 4, § V, if adhered to universally, the implications of FH are both radical and far-reaching.

Objection 3: 'FH is not egalitarian.'

This objection can be made on two fronts: (a) Kant's own views (particularly on gender and race) run afoul of the ostensible egalitarian spirit of FH, and (b) FH, despite its claims to moral equality, excludes nonrational members from moral consideration.

The first version of the objection conflates Kant's ethical doctrines with his own moral errors. There is no doubt that Kant, for example, believed that women were intellectually inferior to men and that they required a "natural curator" in the form of husband or father to manage their civil affairs (VA 7:209).²⁷ Similarly, Kant's anthropological views on race committed him to a Eurocentric hierarchy with the white race perched at the top. The problem arises when we brandish these repugnant views as means of dismissing Kant's ethical insights regardless of their merits. Wood captures this point well when he writes:

[G]reat figures in the history of philosophy are often great precisely because their insights into highly abstract matters of principle far outran the capacity of their own time—and often enough, also their own capacity—to understand fully what these insights mean in practice. To see this gap...as a case of simple hypocrisy is to misunderstand badly the relation of important philosophical principles to the historical conditions of their genesis. (Wood, *KE* 9)

Moreover, focusing on Kant's personal shortcomings or less palatable views conveniently *overlooks* the significance of his contributions in that field. For instance, Kant's harsh views on sexuality,

²⁷ "Woman regardless of age is declared to be immature in civil matters; her husband is her natural curator....just as it does not belong to women to go to war, so women cannot personally defend their rights and pursue civil affairs for themselves, but only by a means of a representative" (VA 7:209)

invariably the subject of ridicule, have often obscured the fact that many feminist thinkers have found an ally in Kant. We revisit this idea in more detail in Chapter 3, § VI.

The second version of the objection covers a wide range of criticisms that center on FH's exclusion of children,²⁸ marginal cases and nonhuman animals from the sphere of moral consideration. Without delving into the particulars of the topic, many modern Kantians have made the case for the moral status of nonrational beings based on Kantian principles that deviate from Kant's original accounts.²⁹

Neither is FH silent on the matter: because our ideas of what constitutes rational nature are *corrigible*, the disputes among Kantians concerning what duties we have towards nonhuman animals (and how they are to be secured) often reflect interpretive disagreements about the *application* of FH.

Objection 4: 'End-setting does not determine objective goodness.'

Wood's contention that, for Kant, when we set an end, we must ascribe objective goodness to it has been challenged by critics and sympathizers alike. Robert Pippin, in his review of Wood's book, *Kant's Ethical Thought*, writes:

the claim that 'reason sets ends on the basis of determining objective goodness' is misleading and cannot be Kant's position, because I take 'objective goodness' in what is its common cognitivist meaning (that some end is objectively better or objectively worse than others, full stop). Whereas Wood wants Kant to say: take out the full stop, and you will see that reason can determine what is objectively *better or worse* for me, and it will also be easier to see that nothing else could determine which end I pursue other than reason in that sense. At which point I insist that this is far afield from objective goodness....What makes rodeo cowboying better for me must be completely relativized to what I regard as more important in life (fame, the truth, the will of God), and then, even my determination of what should be the most important

²⁸ Kant's application of FH in these cases does little to quell the charge of inconsistency: in the Doctrine of Right, he states that children *are* persons, despite their inability to set ends (*MS* 6:280-2).

²⁹ See, for example, Christine M. Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals," *The Tanner Lectures on Human Values*, vol. xxv, ed. G. B. Peterson (Salt Lake City: Utah University Press, 2005): 77-110; Jens Timmermann, "When the Tail Wags the Dog: Animal Welfare and Indirect Duty in Kantian Ethics," *Kantian Review* 10 (2005): 128-49; and Patrick Kain, "Duties Regarding Animals," *Kant's Metaphysics of Morals: A Critical Guide*, ed. Lara Denis (New York: Cambridge University Press, 2010): 210-33.

in *that* sense must again be relativized to considerations that I count as weighty in this sense. (Pippin, “Kant’s Theory of Value” 255).

Kerstein lodges a similar complaint. In his example, Sally, who wants to see Empire State Building, formulates her maxim as follows: ‘If you want to visit the Empire State Building, then you must travel to New York.’ If Wood is correct, then “in setting herself the end of visiting New York, Sally must embrace the view that the end is valuable to all rational agents” (Kerstein, “Deriving the Formula” 208-9). It is, however, difficult to see how others could be rationally compelled to see the trip as valuable *for them*.

Using the terminology of modern ethics, both objections center on Wood’s idea that in setting an end we have an *agent-neutral* reason to pursue that end. This notion puzzles Pippin and Kerstein, since the reasons are not “relativized” to the agent (i.e., they are not agent-relative): it seems to suggest that the ends in questions—“rodeo cowboying” and visiting New York—are ones that *others* have a rationally compelling reason to adopt.

First, Kerstein misunderstands Kant’s hypothetical imperatives when he claims that the “oughtness” of going to New York is conditional upon you wanting to visit the Empire State Building. The imperatives are universally binding in the sense that they “tell *any* rational being that *if* it sets the end of bisecting an angle or curing disease *K*, then it ought to constrain itself to perform the prescribed action” (Wood, *KET* 64). In other words, hypothetical imperatives are addressed to us as a rational beings even *prior* to our adoption of the relevant end. That you ought to travel to New York is conditional upon you *setting the end* of wanting to visit the Empire State Building, but the “oughtness” of the imperative rests *analytically*³⁰ on the fact that going to New York is an *indispensable means* to visiting the Empire State Building. This is true, and hence universally binding on all rational agents, whether or not they adopt the relevant end in the same way that the injunction “keep your promises” is binding as a moral rule even on people who have not made any promises (Wood, *KET* 350, n. 16).

³⁰ “As far as willing is concerned, this proposition is analytic; for in the willing of an object, as my effect, the cause is *already thought...*” (G 4:417, emphasis added).

Wood's general response to doubts about the idea of end-setting determining objective goodness is to point out that agent-relative reasons alone are not "*justifying* reasons" since they lack universal grounding (Wood, *KET* 128). Only agent-neutral reasons can secure the latter. And because "what we are to call good must be the object of the faculty of desire in the judgement of *every reasonable human being*" (*KpV* 5:61, emphasis added)—setting an end requires justifying reasons.

Oddly enough, Wood does not embrace the idea originally developed by Nagel in *The Possibility of Altruism* (and later abandoned by Nagel himself) that *all* reasons must ultimately be agent-neutral, but Korsgaard does.³¹ Not only does she conclude that "the only reasons that are possible are the reasons we share" (Korsgaard, *CKE* 301)—she also embraces Nagel's early notion in *The Possibility of Altruism* that our subjective values commit us, on the pain of solipsism, to objective ones.

Finally, the idea that in setting an end we ascribe objective goodness to it is not as puzzling or rarefied as its detractors would like it to be. Even Sartre who explicitly rejects any notion of an *a priori* good or pre-established value seems to unwittingly corroborate Kant's beleaguered thesis:

"...there is not a single one of our actions that does not at the same time create an image of man as we think he ought to be. Choosing to be this or that is to affirm at the same time the value of what we choose."

"...if I decide to marry and have children—granted such a marriage proceeds solely from my circumstances, my passions, or my desire—I am nonetheless committing not only myself but all humanity to the practice of monogamy."

"A man who commits himself, and who realizes that he is not only the individual that he chooses to be, but also a legislator choosing at the same time what humanity as a whole should be, cannot help but be aware of his own full and profound responsibility" (Sartre, *Existentialism is a Humanism* 24-5).

³¹ See Christine Korsgaard, "The Reasons We Can Share: An attack on the Distinction between Agent-Relative and Agent-Neutral Values," *Creating the Kingdom of Ends* (New York: Cambridge University Press, 1996): 275-310.

The key point is, of course, not the Sartre *endorses* Kant's thesis in his lecture, but that he comes remarkably close to capturing the spirit of it without recourse to an *a priori* good or the conception of rational nature available to Kant. If this move is open to Sartre without such appeals, then the notion of ascribing objective value to the ends we set might be less puzzling than the standard cognitivist picture suggests.

Objection 5: 'The Regressive Argument relies on a faulty inference.'

Even if we grant Kant the idea that we ascribe objective value to the ends we set, the regressive argument involves a further inference from the objective goodness of the ends in question to the unconditional objective goodness of the rational nature that set them. But why should this be? For one, if we value something, we need not value its source. And even if we do value the source why should we value it *unconditionally*?

This objection misfires in one of two ways: it either assumes that rational nature is the *source* of good things, that is, the source of their existence, or it assumes that in setting an end, rational nature *confers* value on that end.

We have already seen that the second Korsgaardian point is false: things are not made good by us choosing them, "on the contrary, we choose...goods because they fulfill our needs or contribute to our happiness. It is not the case that our choosing them *brings it about* that they fulfill our needs or make us happy" (Wood, *KE* 92). In other words, the exercise of practical reason already presupposes (a) their existence and (b) that there is, *prior to our choice*, some good reason to choose them. It is *this* reason (which encapsulates the goodness of the ends in question) that is grounded in the value of rational nature as an end in itself. Wood states the point as follows:

Ends to be produced will usually have value, for instance, because they fulfill the needs, or enrich the lives, or contribute to the flourishing and the happiness of rational beings, and so setting and achieving those ends shows respect and concern for the value of those rational beings. But to hold that the worth of

other goods is derivative from or dependant on the worth of rational nature in this way is not is not at all the same as saying that they have their goodness *conferred* on them by the choices of rational beings (Wood, *KE* 92).

With every end we set, we tacitly affirm rational nature as “the subject of all possible ends” (*G* 4:438). The authority of practical rationality, and hence its unconditional goodness, however, does not lie in *making* its objects of choice good, but rather in *determining* which ends are good. Kant’s inference here is not from the objects of choice to their source, but from the objective goodness of the ends we set to the rational nature capable of setting them. Because this capacity is both active and authoritative in the ends we set on a daily basis, the act of end-setting *already shows us to be committed to* esteeming rational nature as an end in itself.

What the objection serves to highlight is the *nature* of Kant’s proof. The “regress on conditions” does not provide us with a straightforward argument for establishing rational nature as an end in itself. It functions by showing that when we set ends according to reason we are presupposing, or already committed to the view, that rational nature is an end in itself. Because, for Kant, the grounds of ultimate value are indemonstrable, this manner of proceeding is to be expected.

V. Applying FH: The Four Examples Revisited

Having established FH as the central formula in Kantian ethics, we are now in a position to revisit the four examples we surveyed using FUL, and its variant FLN, in the previous chapter. If what we have been arguing so far is correct, FH should be able to deliver the conclusions that FUL/FLN could not.

1. *Suicide*. Since we will be addressing this subject in some detail (as the first duty of virtue) in Chapter 3, § VI, we will keep our comments brief. Someone who “destroys himself,” according to Kant, debases humanity in his person by using it “merely as *a means*” to end his life when it promises more pain than pleasure (*G* 4:429). The interesting point, which Kant appeals to, time and time again, in his treatment of perfect duties to oneself is that humanity in our person, which has absolute worth, is not

something over which we have ownership: “Man cannot dispose over himself, because he is not a thing. He is not his own property....and cannot do as he pleases with his body” (VE 27:386-7). Accordingly, FH forbids us from maiming, defiling, or killing ourselves.

2. *False Promises*. In Chapter 1, § VII, we saw that both the “Logical” and “Practical” universalizability tests (proposed as competing explanations of the wrongness of making false promises) yield false negatives, pointing to structural flaws in FUL/FLN. The application of FH is far more intuitive: he who makes a false promise “use[s] another human being *merely as a means*” since they cannot “contain the end of his action” (G 4:430). The other cannot “contain the end of our action” because if they knew that we had no intention of repaying the loan as promised, they would not agree to it. We fail to treat humanity in their person as an end when we *deceive them* in order to secure the loan we desire.

3. *Resting Talents*. Previously, we maintained that Kant’s application of FUL to the third example involved a *premature* appeal to rational nature as an end in itself. Kant’s argument from FH not only resolves this difficulty, but it also foreshadows Kant’s positive duties of virtue by arguing that “it is not enough the action not conflict with humanity in our person, as an end in itself, it must also *harmonize with it*” (G 4:430). In other words, to neglect our talents “would perhaps be consistent with the *preservation* of humanity, as an end in itself, but not with the *advancement* of this end” (G 4:430). This means that FH involves not only respecting humanity in our person, but also promoting and furthering it: in Kant’s words, “there is bound up with the end of humanity in our own person...the duty to make ourselves worthy of humanity by culture in general, by procuring or promoting the *capacity* to realize all sorts of possible ends...” (MS 6:392). This idea will be important once we turn to Kant’s imperfect duty to develop our natural powers in Chapter 3, § VI.

4. *Refusing to Succour Those in Need*. Kant, once again, emphasizes the need for a *positive* duty based on FH:

Now, humanity could indeed subsist if no one contributed anything to the happiness of others while not intentionally detracting anything from it; but this is still only a negative and not positive agreement with *humanity, as an end in itself*, if everyone does not try, as far as he can, to advance the end of others. (*G* 4:430)

~

...a human being is an end for himself as well as others, and it is not enough that he is not authorized to use either himself or others merely as a means (since he can still be indifferent to them); it is in itself his duty to make man as such his end. (*MS* 6:395)

Therefore, refusing to succour those in need expresses contempt for humanity because it remains indifferent to their ends. Since rational nature as an end in itself admits of absolute, and hence equal, worth in my person and *the person of any other*, “insofar as their existence contains the same rational ground for respect, it is equally impossible for me to will that I should not extend the same help to them” (Wood, *KET* 150). We will have more to say about this when we discuss the duty of beneficence in the following chapter.

So far we have been making the case for FH and the idea of rational nature as the ultimate value underpinning Kantian ethics. In Chapter 3 we turn to FH’s positive role in furnishing obligatory ends and ethical duties in the Doctrine of Virtue, and argue that Kant’s richest ethical insights are to be found in the *Metaphysics of Morals* which constitutes the final form of his ethical philosophy.

Chapter 3: The Doctrine of Virtue

I. Expressive Ethics

In Chapter 2, we examined the foundational arguments for FH and made the case for rational nature as the fundamental value in Kantian ethics. The wide-ranging implications of FH and its positive role, however, only come to the fore in Kant's *Metaphysics of Morals*, specifically in the second half, known as the Doctrine of Virtue [*Tugendlehre*].

Before we turn to these, it is worth asking how Kantian ethics looks through the lens of FH, especially in light of the familiar characterizations of the former as “deontological” or “rule-based.”

Broadly speaking, consequentialist theories are those that derive the value of actions from the value of their consequences and, as such, are concerned with promoting desirable states of affairs. Deontological theories, on the other hand, are said to be about the rightness or wrongness of actions, whereas virtue ethics is primarily concerned with agents, with virtues and vices playing a central role.

If we interpret Kantian ethics exclusively through FUL, it is easy to see how Kant's theory is classified as deontological. If, however, we adopt FH as the primary formula, we end up with fundamentally different conception of ethics that does not seem to fit (at least not neatly) into one of the three main ethical camps.

FH tells us to “*so act that you use humanity, in your person as well as in the person of any other, always at the same time as an end, never merely as a means*” (G 4:429). In Chapter 2, § 1, we outlined the three value conceptions that characterize rational nature, or humanity, as an end in itself. One of these was that humanity was an *existent end*—that is, an end to be respected, furthered and “esteemed in every willing” (G 4:437).

Kant defines respect, against the conventional meaning it inherits from our unsociable sociability, by appealing directly to FH: respect “is not to be understood as the mere *feeling* that comes

from comparing our own *worth* with another's....It is rather...the *maxim* of limiting self-esteem by the dignity of humanity in another person" (MS 6:449). Given this fact, what FH enjoins us to do is to *express* the proper respect for the worth of humanity wherever it is found, while forbidding conduct that fails to afford humanity the respect it is due: "Morally good conduct expresses respect for humanity as an existent end, while bad conduct...expresses disrespect or contempt for humanity" (Wood, KET 117).

Nearly all the ethical duties, as well as the two obligatory ends, found in the Doctrine of Virtue are derived from FH in accordance with this ethical injunction. This results in an expressive ethics "largely overlooked in moral theory" (Wood, KET 141). The ethics of FH is "expressive" because it involves an *expression* of esteem directed at the categorical value of humanity as an end in itself, whether this takes the form of respecting, preserving, or promoting that value.

Scepticism about expressive ethics comes in three forms: the first is scepticism about the notion of existent ends, that is, ends that are not brought about or effected by us, but for the sake of which we nevertheless act (by esteeming, preserving or furthering them). To allay this metaphysical suspicion, Wood offers a number of examples from our daily lives: "people kneel, bow their heads, or doff their hat to something (such as a flag or religious object), they "shake hands, congratulate or condole with others, and say 'Please,' 'Thank you' and 'You're welcome'" (Wood, KET 116, 141-2). In all these cases, there is no end to be effected—only the gesture of esteem, gratitude or reverence towards an object of value. For example, in bowing our head before a religious symbol, we manifest reverence towards a holy figure or sacrament. FH differs from these examples only in its bold assertion that the fundamental value to be esteemed is *humanity*, which is always found in a person.

The second form of scepticism expresses doubts about the idea that a *single* value (humanity) can ground an entire ethical system. One response to this objection is to point out that utilitarianism, one of the most influential ethical theories, is founded precisely on such a value: while Kant derives FH from the fundamental value of rational nature as an end in itself, Mill derives the principle of utility from

the fundamental value of happiness or pleasure. Another response is to simply demonstrate that FH *can* furnish us with obligatory ends, various positive duties and a robust taxonomy of virtues. In fact, as we proceed our aim will be not only to expound Kant's intricate theory of virtue and ethical duties in the *Metaphysics of Morals*, but also to show how they are derived, almost entirely, from FH.

The third and final misgiving about expressive ethics assumes that because FH commands us to express the proper *respect* or *esteem* for humanity as an end in itself, it asks only that we put ourselves in a certain state of mind. As we shall see, Kant's own examples and applications of FH all belie this idea. Moreover, in his *Lectures on Ethics*, Kant explicitly rejects this subjective reading of FH, even in our duties of love:

[W]e must see to it that our inclination to love the other, and wish for his happiness, are not idle longings, or desires with no outcome, but practical desires. A practical desire is one that is directed not so much to the object as to the actions whereby this object is brought about. We should not only take satisfaction in the welfare and happiness of others, but this satisfaction should relate to the effectual actions that contribute to this welfare. (VE 27:421)

Therefore, according to FH, "I should not [merely] wish, when the other is in misery, that he might be rescued from it, but should attempt, rather, to rescue him" (VE 27:421).³²

So far, our arguments have centered on defending the primacy of FH and the expressive ethics it gives rise to against FUL, and the familiar rational-formalistic reading of Kant it entails. We have yet to address the place of the frequently overlooked *Metaphysics of Morals* in relation to the far more influential *Groundwork*.

³² "The heart...is only a good heart insofar as it is able to contribute something to the other's happiness, and not when it merely wishes for that. People pride themselves on having a kind heart, when they merely wish that everyone might be happy. But the only one to have a kind heart is he who contributes something to that happiness" (VE 27:421).

II. The Metaphysics of Morals

The *Groundwork of the Metaphysics of Morals* was first published in 1785, and was intended, as the title suggests, as an antecedent to the *Metaphysics of Morals*:

Intending, then, to publish some day a Metaphysics of Morals, I issue this Groundwork in advance....I find it useful to separate from [the Metaphysics of Morals]...this **preparatory work of laying its foundations** so that I may omit the subtleties it unavoidably contains from **more accessible doctrines** in the future. (G 4:391-2, emphasis added)

The long-promised “more accessible doctrine” would come 12 years later in 1797 and stand among Kant’s last published works.³³ It would contain both Kant’s Doctrine of Right (Part I) and his Doctrine of Virtue (Part II). Schopenhauer, who was not impressed (to put it mildly) by Kant’s Doctrine of Virtue in the *Metaphysics of Morals* wrote of the former: “[in] this counterpart of his deplorable *Rechtslehre* [Doctrine of Right], the effects of his weakness brought on by old age are predominant” (qtd. in Kuehn, “Kant’s *Metaphysics of Morals*” 9-10).

Despite Kant’s own declarations, the standard conception of Kantian ethics, both popular and scholarly, has been derived almost entirely from the pages of the *Groundwork* and Anglophone scholars have been reluctant to let the *Metaphysics of Morals* influence their interpretation of his moral philosophy. If they are, along with Schopenhauer and Kant’s other contemporary critics, justified in writing off the *Metaphysics of Morals* as a confused product of senescence, then Kantian ethics has no place for virtues and ends, and the primacy of FH is cast into doubt. If, on the other hand, we find in the *Metaphysics of Morals* a cogent theory of virtues and ethical duties derived from FH, then the “familiar image of Kantian ethics is in serious error on some fairly basic points” (Wood, “The Final Form” 4).

Manfred Kuehn, in defending the importance of the *Metaphysics of Morals* compared to Kant’s (two) other late works, adduces three reasons in support of his claim: (1) the other works closely follow

³³ The only two works published after the *Metaphysics of Morals* were *The Disputes of the Faculties* and the *Anthropology from a Pragmatic Point of View*.

Kant's lectures, whereas the *Metaphysics of Morals* has a logic and structure unique to it, (2) it was a work that Kant had been planning for a long time (thirty long years)³⁴ and thus not a late idea, marred by old-age, and (3) it sums up or represents the answers to Kant's most pressing ethical concerns in his philosophical system. Kuehn concludes, albeit cautiously: "So, the view that the *Metaphysics of Morals* should actually inform our view of Kant's ethics as it is expressed in the *Groundwork* and the second *Critique* and not the other way around is not entirely implausible..." (Kuehn, "Kant's *Metaphysics of Morals*" 11). But we can be significantly *bolder* in our conclusion, and argue, along the lines of Marcia Baron and Melissa Seymour Fahmy that the *Metaphysics of Morals*

is a major work of Kant's and is the culmination of his ethical thought...because most of the myths about Kant's ethics—e.g., that it concerns actions, not character or how to live; that is all about applying a rule to a generate a clear decision about how we should act; that it is rigid, leaving no room for hard cases; that it is not sensitive to particular situations and to the nuanced character of moral life; that the Categorical Imperative not only is not based on anything empirical but is supposed to be applied in such a way as to ignore empirical facts—lose whatever semblance of plausibility they might otherwise have once one reads the *Metaphysics of Morals* (in particular, Part II). (Baron and Fahmy, "Beneficence and Other Duties" 211)

As we turn to Kant's Doctrine of Virtue, we will see how these respective "myths" give way to a novel and often neglected conception of Kantian ethics. In order to fully appreciate Kant's rich and nuanced arguments, we will also draw extensively from his *Lectures on Ethics*, where he expounds the duties of virtue in greater detail.

III. Virtue

In the second *Critique*, Kant describes virtue [*Tugend*] as the "moral disposition *in conflict*" (*KpV* 5:84). More specifically, in the "struggle of inclination with the moral law," virtue is "the constant disposition

³⁴ According to Wood, Kant was working on a system of ethics called the "metaphysics of morals" as early as 1768 (Wood, "The Final Form" 1). The *Metaphysics of Morals* was published in 1797—roughly thirty years later.

(*intentio constans*)” to carry out one’s duties (VE 27:492). Kant’s formal definition, however, is found in the *Metaphysics of Morals*: “Virtue...is the moral strength of a *human being’s* will in fulfilling his *duty*, a moral *constraint* through his own lawgiving reason, insofar as this constitutes itself an authority *executing law*” (MS 6:405).

This has two implications: (1) there is no strict duty or obligation to possess virtue, since we *require virtue* to place ourselves under any obligation at all (MS 6:405). As we shall see in Section VI, we have a wide or imperfect duty to increase our moral perfection which includes the cultivation of virtue, but failure to achieve this end is not blameworthy so long as we are striving towards it. And (2) owing to “the frailty (*fragilitas*) of human nature” we can never *reach* moral perfection (MS 6:446). Hence virtue is an “unending” or “continual” progress towards moral perfection which “can never be completed” (KpV 5:32-3).

The best way to appreciate this last point is to contrast Kant’s notion of virtue with “*holiness*” [*Heiligkeit*], which lies in the “possession of a complete *purity* of dispositions of the will” (KpV 5:84). Holiness, for Kant, is the ability to act from the moral law alone as a sufficient motive. Because holiness is an unattainable practical idea for finite rational beings, we can never be holy:³⁵ “man can only get so far as to be virtuous” (VE 27:465). Therefore virtue, as the moral strength of the will in carrying out its duty, consists in an “endless progress toward holiness” (R 6:46-7). This idea of holiness as an ideal or archetype to approximate will be important once we see what *disposition* Kant attaches to the holy will.

Finally, “virtue as the will’s conformity with every duty” naturally corresponds to different morally prescribed ends which Kant refers to “*duties of virtue (officia honestatis)*” (MS 6:395). As a result, there are many duties of virtue which we ought to adopt as ends and our commitment to them may vary in strength from one duty to the next.

³⁵ “...an ideal of holiness is not attainable by any creature but is yet the archetype which we should strive to approach and resemble in an uninterrupted but endless progress” (KpV 5:83).

IV. Obligatory Ends: 'The Ends That Are Also Duties'

Contrasting ethics with the Doctrine of Right, Kant maintains that “ethics goes beyond this and provides a *matter*...an **end** of pure reason which it represents as an end that is also objectively necessary, that is, an end that, as far as human beings are concerned, it is a duty to have” (MS 6:380). It is interesting to note that Kant formulates the charge of “empty formalism” long before his critics when he argues that

there must be such an end [that is also a duty] and a categorical imperative corresponding to it....For were there no such ends, then all ends would hold for practical reason only as means to other ends; and since **there can be no action without an end, a categorical imperative would be impossible**. This would do away with any doctrine of morals. (MS 6:385, emphasis added)

In other words, “only the concept of an *end* that is also a duty, a concept that belongs exclusively to ethics, establishes a law for maxims of actions by subordinating the subjective end (that everyone has) to the objective end (that everyone ought to make his end)” (MS 6:389).

The *positive* role of FH in furnishing these ends that are also duties is immediately highlighted by Kant: “a human being is an end for himself as well as others, and it is not enough that he is not authorized to use either himself or others merely as means (since he could still be indifferent to them); it is in itself *his duty to make man as such his end*” (MS 6:395, emphasis added).

In response to the question “What are the ends that are also duties?” Kant offers a single, compendious sentence: “*one’s own perfection and the happiness of others*” (MS 6:385). Before clarifying the two obligatory ends, he warns us that they cannot be interchanged (so that we adopt our *own* happiness and the perfection *of others* as duties). We will see why this is and what this asymmetry reveals about his doctrine as we examine the two ends individually.

More importantly, Kant maintains that these ethical duties are “wide” or “meritorious” duties (MS 6:390-4). They are wide in the sense that it our duty to adopt them as ends, but they leave us with “playroom (*latitudo*)” regarding how, and to what extent, we should fulfil them (MS 6:390). They are

meritorious in that “fulfilment of them is *merit*...but failure to fulfill them is not in itself *culpability*” unless we make it our principle not to adopt them (*MS* 6:390). As we shall see, not all of our ethical duties in Kant’s final taxonomy of virtues, which is grounded on the two obligatory ends, are wide: perfect duties to oneself and duties of respect to others are narrow or “*limiting* (negative) duties” which forbid us to act contrary to the positive ends in question (*MS* 6:419).

One’s Own Perfection

Kant includes both natural and moral perfection under the first obligatory end, which he derives entirely from FH:

The capacity to set oneself an end—any end whatsoever—is what characterizes humanity (as distinguished from animality). Hence there is also bound up with the end of humanity in our own person the rational will, and so the duty, to make ourselves worthy of humanity by culture in general, by procuring or promoting the *capacity* to realize all sorts of possible ends, so far as this is to be found in the human being himself. (*MS* 6:392)

We saw this passage in Chapter 2, § II, when we identified humanity with the capacity to set ends. Here our focus lies on the second half of the passage in which the duty of one’s own perfection enjoins us to “make ourselves worthy of humanity” in our person by “promoting the capacity to realize all sorts of possible ends” (*MS* 6:392).

In their article, Baron and Fahmy characterize Kant’s denial of a duty to perfect others as evidence of his “anti-paternalism,” that is, the idea that “one must initiate the process and take charge” solely within the bounds of self-governance (Baron and Fahmy, “Beneficence and Other Duties” 213). In spite of their parenthetical qualification,³⁶ this has the effect of enforcing the specious view of Kantian ethics as overly individualistic and hostile to aid, especially in the form of succour from without.

³⁶ “The idea in holding that one cannot be perfected by others presumably is not that one could not get some help from one’s friends....Friends can help but only by helping the agent attain her ends” (Baron and Fahmy, “Beneficence and Other Duties” 213).

Kant's reasons have more to do with the *nature* of ethics and its relation to setting ends. The first distinction to note is the one Kant makes between the Doctrine of Right (juridical duties) and the Doctrine of Virtue (ethical duties): "What essentially distinguishes a duty of virtue from a duty of right is that external constraint to the latter kind of duty is morally possible, whereas the former is based only on free self-constraint" (*MS* 6:383). In other words, juridical duties are duties that can be coercively enforced. But the setting of an end, which lies in practical reason, is always an act of freedom. Therefore ethics as "the system of *ends* of pure practical reason....contains duties that one cannot be constrained by others (through natural means) to fulfill...since *coercion* to ends (to have them) is self-contradictory" (*MS* 6:381).

It follows from this that Kant's denial of a duty to perfect others is not reflective of a monadic individualism or atomistic attitude—it is predicated instead on the idea that setting an end for another is *self-contradictory*: "it is a contradiction for me to make another's perfection my end....For the perfection of another human being, as a person, consists just in this: that he himself is able to set the end...and it is self-contradictory to require that I do...something that only the other can himself do" (*MS* 6:386).

Put differently, Kant's denial of a duty to perfect others does not tell us that we *should not* set an end for others because in doing so we fail to afford them the respect they are due as rational beings by violating their autonomy—it says, first and foremost, that we *cannot* conceptually "set an end" for others without moving from the sphere of ethical duties (which involve free self-constraint) to the sphere of juridical duties (which involve coercion from without).

The Happiness of Others

In Chapter 1, § V, we saw that, for Kant, the end of happiness can be presupposed "safely and a priori in every human, because it belongs to his essence" (*G* 4:415). This translates into a duty to promote the happiness of others, but not our own (*MS* 6:388).

We are, accordingly, faced with three questions: (1) how is “happiness” to be understood; (2) how do we promote it; and (3) why is promoting our own happiness not a duty for us?

Kant conceives of happiness as the idea of “an absolute whole...a maximum of well-being, in my present and every future condition” (*G* 4:399; *MS* 6:387; *VE* 27:276, 29:599), and maintains that we have a prudential imperative to pursue our own happiness. The problem is that “the concept of happiness is so indeterminate a concept that, even though every human being wishes to achieve it...he can never say determinately and in agreement with himself what he actually wishes and wants” (*G* 4:418).

This means that promoting the happiness of others, for Kant, always involves adopting *their* (permissible) ends as my own: “It is for them to decide what they count as belonging to their happiness” (*MS* 6:388). Here Baron and Fahmy are right to highlight Kant’s anti-paternalism, since he openly acknowledges that “I cannot do good to anyone in accordance with *my* concepts of happiness” (*MS* 6:454). Even though the concept of happiness belongs to the essence of every human, we can only benefit others by making their morally permissible ends our own.

A final remark: due to its utilitarian pedigree, the phrase “promoting happiness” can be easily misconstrued as an injunction to *maximize* happiness. But as we saw earlier, the duty to promote the happiness of others is a wide or meritorious one. Therefore we “deserve no blame for failing to promote the end on any given occasion, and *a fortiori* no blame for not promoting it maximally” (Wood, *KET* 328).

Kant does not, however, believe that we have a *moral* duty to promote our own happiness:

For his own happiness is an end that every human being has (by virtue of the impulses of his nature) but this end can never without self-contradiction be regarded as a [moral] duty. What everybody already wants unavoidably, of his own accord, does not come under the concept of *duty*, which is *constraint* to an end adopted reluctantly. (*MS* 6:386)

This “self-other asymmetry” prompts Michael Slote to charge Kantian ethics with devaluing agents (see Ch. 1, § V): “Of course, viewed from the recipient of acts...Kantian morality regard[s] the

well-being of every person as having a positive and fundamental valuational significance” but it assigns “no such significance to...the well-being of the act’s agent” (Slote, *From Morality to Virtue* 15-6).

In response, we argued that counsels of prudence or assertoric imperatives, not moral ones, serve to ground the rational obligation to pursue one’s own happiness. In the *Metaphysics of Morals*, Kant, once again, denies any *moral* duty to promote the latter. But this denial does not translate into the fearful asymmetry Slote suggests. Kant’s chief concern is that prudential considerations do not usurp categorical ones, aside from which they are given free rein: “Pure practical reason merely *infringes upon* self-love, inasmuch as it restricts it, as natural and active in us even prior to the moral law, to the condition of agreement with the law, and then it is called *rational self-love*” (KpV 5:73).

Self-love, for Kant, is *rational* “to the extent that with respect to the end only what is consistent with the greatest and most abiding well-being is chosen....Reason here only occupies the place of the servant of natural inclinations; the maxim one adopts has absolutely no relation to morality” (R 6:45n). Thus it is fallacious to infer from the absence of a (direct) *moral* duty to promote one’s own happiness that Kant places no value or significance on the latter. Kantian ethics only asks that self-love not become self-conceit by arrogating for itself the *lawgiving* function of morality: “Let this maxim [of self-love], however, become an unconditional principle of the power of choice, and it is the source of incalculably great resistance to morals” (R 6:45n).

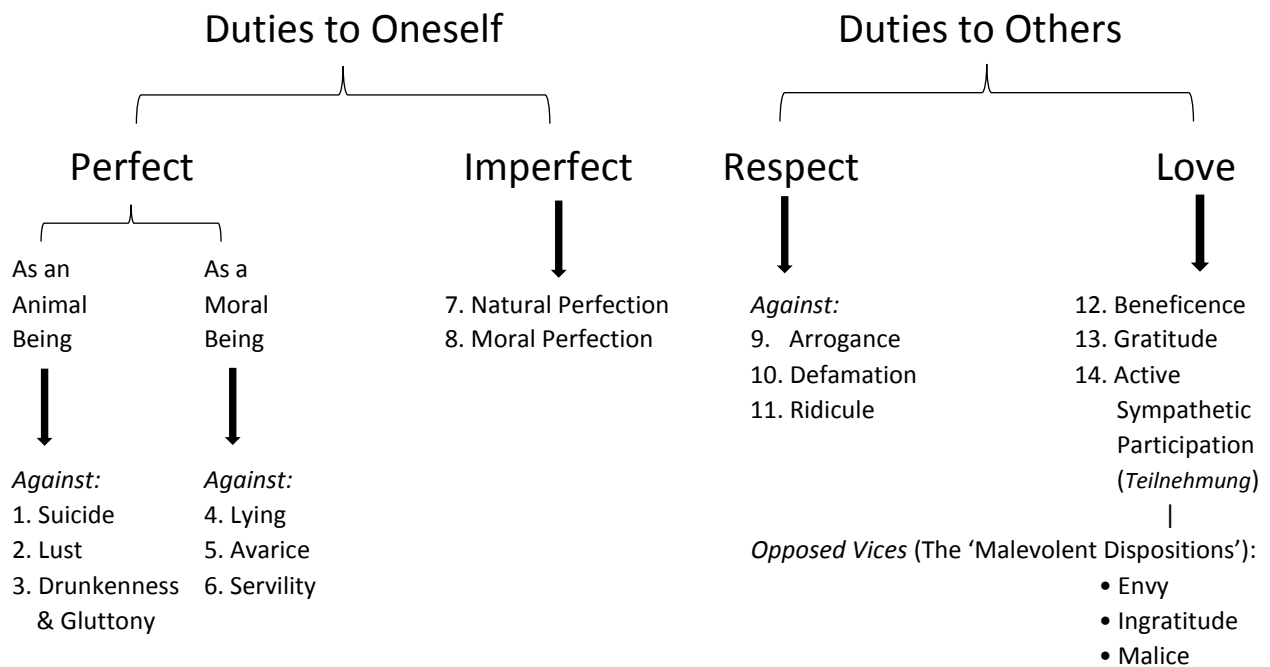
In addition to the imperatives of prudence, Kant also believes that we have an *indirect* duty to make our own happiness an end, “for lack of contentment with one’s condition, in the trouble of many worries and amidst unsatisfied needs, could easily become a great *temptation to transgress one’s duties*” (G 4:399). In the *Metaphysics of Morals*, Kant offers a similar argument listing “adversity, pain and want” as temptations to violate one’s duty, which can be checked by “prosperity, strength, health and well-being in general” (MS 6:388).

V. The Taxonomy of Ethical Duties

In the *Groundwork* Kant relies on a provisional division of duties, stating in a footnote that “I reserve the division of duties entirely for my future *Metaphysics of Morals*” (G 4:422n).

The primary division in Kant’s taxonomy of ethical duties, which is founded on the two obligatory ends, is between duties to oneself and duties to others. Under the first heading, Kant distinguishes between *perfect* (strict or narrow) duties and *imperfect* (wide or meritorious) duties. Under duties to others the division is similarly between duties of *respect*, which are narrow in scope, and duties of *love*, which are wide. A complete taxonomy of Kant’s duties of virtue is provided in the table below:

DUTIES OF VIRTUE



Kant's taxonomy enumerates fourteen duties of virtue, but it is significant to note that these are not all the ethical duties we have. Because the *Metaphysics of Morals* "cannot dispense with the principles of application" it takes as its object "the particular *nature* of human beings" (MS 6:217). Subsequently, it is concerned only with "those duties that are generated by applying the principle of morality to empirical human nature in general" (Wood, KE 165). These duties, in turn, play a *foundational* role in any social or institutional virtues that might arise, without establishing them in advance. This plurality of (prospective) virtues, coupled with our options and latitude in pursuing them, will be crucial once we turn to our duty to perfect our natural powers in the following section.

VI. Duties of Virtue

The centrality of FH is most prominent in Kant's discussion of perfect duties to oneself. These are not duties about our own welfare or self-interest, as they are traditionally understood in moral philosophy, but about our self-perfection (Wood, "Duties to Oneself" 235). Kant divides perfect duties into duties to oneself as an animal being and duties to oneself as a moral being.

It is important to stress that we are not here concerned with mounting an elaborate defence of Kant's more controversial duties (which are tellingly all negative duties)—our aim is merely to show that (a) they are derived, almost entirely, from FH and (b) that, along with the positive ends they protect, they give rise to a richer, more nuanced conception of Kantian ethics that belies many of the charges of its critics.

Perfect Duties to Oneself as an Animal Being (Against):

1. ***Suicide***. Kant sets aside the idea that killing oneself might be regarded as a violation of one's duty to others (e.g. family, friends, or citizens). As we saw in Chapter 2, § V, suicide violates a perfect duty to oneself because "disposing of oneself as a mere means to some discretionary end is debasing humanity in one's person" (MS 6:423). In his *Lectures on Ethics*, Kant relies on FH to similar effect:

“...by sacrificing one’s condition to abandon at a stroke all the pains and hardships of life...humanity is subordinated to animal nature” (VE 27:1427-8).

“By the rule of prudence, it would often be the best course, to remove oneself from the scene; but by the rule of morality it is not allowed under any condition, because it is the destruction of humanity” (VE 27:373).

“...suicide is contrary to the concept of the right of humanity in my own person; and humanity is in itself an inviolable holiness...” (VE 27:628).

Wood has argued that “the position Kantian ethics should take [on suicide] is quite distinct from (in many cases diametrically opposed to) the position Kant himself takes” (Wood, KE 173). The important point for our purposes is that Kant demonstrates an understanding of the nuanced problems that attend his position. For example, in the attached “*Casuistical Questions*” at the end of his discussion of suicide, he leaves the reader with the following *unanswered* questions: is it wrong to “hurl oneself to certain death” in order to save one’s country (MS 6:423)? Is it wrong for a man who, having been bitten by a rabid dog and beset by hydrophobia, to take his own life “lest he harm others...in his madness” (MS 6:423-3)? In his *Lectures of Ethics*, Kant argues that “a worthless man values his life more than his person” and that “a man of inner worth would sooner sacrifice his life than commit a disreputable act” (VE 27:376). The fact that Kant poses these difficult questions (that problematize his own conclusions) even in the case of *perfect* duties should make us think twice before labelling his ethics “rigoristic.”

2. Lust. Kant’s thoughts on sexuality have often been the subject of both amusement and derision. There is, of course, a great deal to balk at: Kant’s hostility towards sexual pleasure, his harsh condemnation of “unnatural” sex acts, etc. Wood and Herman are nevertheless right to point out that “the noise in our heads provoked by Kant’s wrongheadedness on the surface too easily prevents us from listening for the ways he gets things right at a deeper level” (Wood, “Duties to Oneself” 238).

“Defiling oneself by lust” in the context of Kant’s perfect duties to oneself is often translated as carnal self-defilement (*MS* 6:424-5)—a slightly less ambiguous (but no less cumbersome) epithet. Once we set aside Kant’s teleological arguments, the violation of a duty to oneself lies in the fact “man surrenders his personality (throwing it away), since he uses himself merely as a means to satisfy an animal impulse” and this amounts to a “violation of humanity in one’s own person” (*MS* 6:425).

Kant’s meaning becomes clearer, and arguably more palatable, when we turn to his discussion of the sexual impulse directed towards another:

Since the sexual impulse is not an inclination that one human has for another, *qua* human, but an inclination for their sex, it is therefore a *principium* of the debasement of humanity....The desire of a man for a woman is not directed to her as a human being; on the contrary, the woman’s humanity is of no concern to him, and the only object of his desire is her sex. So humanity here is set aside... (*VE* 27:385)

The sexual impulse, for Kant, is not a species of love since “love as a human affection, is the love that wishes well, is amicably disposed, promotes the happiness of the other and rejoices in it” whereas “those who merely have sexual inclination love the person from none of the foregoing motives...and will plunge them into the greatest unhappiness, simply to satisfy their own inclination” (*VE* 27:384). Thus, the sexual impulse *imperils* humanity—not only by using it a means, but by disposing over it as a thing, which FH prohibits in one’s person and the person of any other.

However unappealing Kant’s views (or his application of them) may seem, there is surely some truth to the notion “that there is something about what happens in human sexual relations that...[may] lead to a condition compromising the moral standing of the partners” (Herman, “Could It Be Worth” 59). What offends modern readers of Kant, however, is not that the idea that the sexual impulse might jeopardize the moral standing of the partners (which is, after all, hardly controversial), but rather Kant’s outright denial that there can *be* a sexual *relationship*, since “in loving from sexual inclination” we “make the other person into an object of...appetite. As soon as the person is possessed, and the appetite sated,

they are thrown away, as one throws away a lemon after sucking the juice from it" (VE 27:384). There can be no sexual relationship, for Kant, because when we act from "sexual inclination" we dispose over the other as an *object* of appetite.

We find a modern equivalent to Kant's shocking denial in Lacan's notorious claim, *il n'y a pas de rapport sexuel* [there is no sexual relationship]. Here is how Alain Badiou explains Lacan's controversial phrase:

Jacques Lacan reminds us, that in sex, each individual is to a large extent on their own...Naturally, the other's body has to be mediated, but at the end of the day, the pleasure will always be your pleasure. Sex separates, doesn't unite. The fact that you are naked and pressing against each other is an image, an imaginary representation. What is real is that pleasure takes you a long way away, very far from each other. What is real is narcissistic....love is what fills the absence of the sexual relationship. Lacan does not say that love is a disguise for sexual relationships; he says that sexual relationships don't exist, that love is what comes to replace that non-relationship. (Badiou, *In Praise of Love* 18-9)

Badiou's explanation mirrors Kant's, without the strident language and jarring images of spent lemons cast aside. Kant's emphasis on objectification explains why (certain aspects of) his thoughts on sexuality have appealed to feminist thinkers. Herman, for example, who once thought that "Kant's views on sex, women and marriage would best be forgotten by anyone who wanted to take Kant seriously" finds a surprising ally in Kant (Herman, "Could It Be Worth" 53). In "Could It Be Worth Thinking About Kant on Sex and Marriage?" she offers the following anecdote:

A year and a half ago I was teaching Kant's political philosophy and participating in a study group that was reading Andrea Dworkin and Catharine MacKinnon [the well-known feminist writers jointly responsible for the 1983 Anti-pornography Civil Rights Ordinance]. In the midst of a class discussion I found that I could paraphrase Kant's views on sexuality using Dworkin and MacKinnon's analysis. It seemed at once a perverse and right thing to do. (Herman, "Could It Be Worth" 71, n. 2)

Herman goes on to draw strong parallels between Kant's views on sexuality and passages from Dworkin's book, *Intercourse*. As a testament to Kant's influence, in the *Stanford Encyclopaedia of Philosophy*, under the article "Feminist Perspectives on Objectification" by Evangelia Papadaki we find the first section dedicated to Kant's ideas on sexuality. Finally, highlighting the radical import of FH in her essay "Autonomy-Denial in Objectification" Rae Langton outlines Kant's contribution as follows:

The thought that there might be something *wrong* with treating a human being as a living tool is one that gains eloquent expression...in the work of Immanuel Kant. For Kant, moral wrong-doing consists in a failure to treat humanity 'always as an end and never as a means only'—a failure to respect that humanity 'by virtue of [which] we are not for sale at any price'. This historically Kantian idea has gained new impetus in recent applications by feminist thinkers, who have observed its relevance to oppression, and to the varied ways that women have been treated as means only, and sometimes put for sale. (Langton, SS 223)

Although many feminists may still "have few warm words" for Kant (Langton, SS 241), to view his thoughts on sexuality, however misguided his proposed solutions may have been, as *inimical* to the promotion of equality and emancipation of women is to mistake one of their closest allies for one of their staunchest enemies.

3. ***Drunkenness and Gluttony***. In "stupefying oneself by the excess use of food or drink" we show a disregard for humanity by impairing, or wholly suspending, its powers and capacities (MS 6:427). Kant's discussion of the psychology behind two cases is revealing. He admits that drinks and narcotics, such as opium, "are seductive because, under their influence people dream for a while that they happy and free from care" but "dejection and weakness [soon] follow" coupled with the need for higher doses (MS 6:427).

Kant seems to acknowledge that this carefree play of the imagination which intoxication gives rise to is, at least independently, a good thing, since he regards gluttony as the more vicious of the two because "it only lulls the senses in a passive condition, and unlike drunkenness, does not even arouse

imagination to an *active* play of representations; so it approaches even more closely the enjoyment of cattle” (MS 6:427; VE 27:380-1). This also explains why Kant goes on to rank opium “closer to being a base act than the use of wine” since it has the effect of stupefying humanity in us by making the user “silent, reticent and withdrawn” whereas wine may enliven us in conversation and endear us to company (MS 6:428).

This social argument is not incidental. In his *Lectures on Ethics*, Kant praises wine for its power to sharpen the senses and facilitate communication as well as the enjoyment of social pleasures by causing individuals to overlook each other’s weaknesses and become more good-natured, gregarious and open-hearted (VE 27:633). Absent this social aspect, however, it loses its advantage over gluttony, since “drinking and drunkenness in solitude are...shameful; because there *we no longer have the factor that raised them above gluttony*” (VE 27:381, emphasis added).

It is interesting to note that, far from the picture of Kant as a solitary, donnish curmudgeon, his observations seem to have their basis in experience. For example, in Roger Scruton’s short introduction, we find this description of Kant’s daily routine:

He would invariably have guests at his midday meal, inviting them on the same morning lest they should be embarrassed by the need to refuse some other invitation, and providing for each a pint of claret, and, if possible, some favourite dish. He conversed, to the delight and instruction of his companions, until three, endeavouring to end the meal in laughter... (Scruton, *Kant* 8)

It is no wonder, then, that Kant cites Seneca approvingly—“*virtus ejus incaluit mero* [his virtue was kindled by wine]”—and offers us the following explanation: “It [winebibbing] awakens much activity, and is to that extent acceptable to morality; indeed owing to the advantages that *temulentia* [conviviality] exhibits in regard to social life, promotion of this tendency...seems impossible to quarrel with” (VE 27:633). What is remarkable about Kant’s statement is not his celebration of (the powers of)

wine, but the nature of his argument: it is, first and foremost, a *social* argument, predicated on the *promotion of conviviality*.

Perfect Duties to Oneself as a Moral Being (Against):

4. **Lying.** Kant deviates from the standard accounts of the wrongness of lying by insisting that it is not located in the harm that it causes others (in which case it would be a violation of a duty to them), or in the harm that it brings upon the liar (in which case it would be a prudential, not moral, violation). Instead Kant argues that “by a lie, a human being throws away and, as it were, annihilates his dignity as a human being” (*MS* 6:429). Therefore “every *falsiloquium* [false statement], every knowing deception is impermissible” because “there is always a violation in this of humanity in our own person” (*VE* 27:700).

Kant distinguishes between two types of lies and “identifies the deception of others with the deeper and more serious action of *self-deception*” (Potter, “Duties to Oneself” 376). According to Kant, these “inner lies” which mar our self-confessions and obfuscate our true motives by falsifying them or making them opaque are the wellspring of “the ill of untruthfulness” that spreads into our relations with others (*MS* 6:430-1). It follows from this that

the ultimate wisdom, which consists in the harmony of a human being’s will with its final end, requires him first to remove the obstacle within...and then to develop the original predisposition to a good will within him, which can never be lost. (Only the descent into the hell of self-cognition can pave the way to godliness.) (*MS* 6:441)

This makes the inner lie, or self-deception, a direct violation of our duty to ourselves as moral beings, since “know[ing] your heart—whether it is good or evil, whether the source of your actions is pure or impure, and what can be imputed to you...belong[s] to your moral *condition*” (*MS* 6:441).

5. **Avarice.** Wood is right to find in Kant’s treatment of avarice “some perceptive remarks about the psychology of this brand of self-contempt” which prefigure Marx’s critique of the fetishism of commodities (Wood, *KE* 173-4). According to Kant, there are two forms of avarice: “*greedy avarice*” and

“miserly avarice” (MS 6:432). Kant is not concerned with the first of these since it constitutes a “violation of one’s duty (of beneficence) *to others*” (MS 6:432)—not a violation of a duty to oneself.

It is worth pausing here to appreciate the implications of Kant’s distinction: Kant defines *greedy avarice* as “acquiring the means to good living in excess of one’s true needs” (MS 6:432). Kant views the adoption of this end as a violation of our duty to others. Thus Wood is not mistaken when he exclaims: “how ‘un-American’ Kantian ethics is—the *mere end* of having more than others have is contrary to ethical duty!” (Wood, KE 173).

Miserly avarice, on the other hand, involves “restricting *one’s own* enjoyment of the means to good living so narrowly as to leave one’s own true needs unsatisfied” (MS 6:432). The miser violates humanity in his own person because he “possesses money, and thus a means of attaining all ends, but does not attempt thereby to become more perfect, since he renounces every use of it” (VE 27:606).

In prescient language that would find its way into Marx’s *1844 Manuscripts* and early essays, Kant locates this breach of humanity in the miser’s fetishistic deification of money as an end in itself (when it is only a means) and the degradation of humanity to a mere means (when it is the end in itself):

since a man, in and for himself, is supposed to be an end for his humanity, the miser assuredly violates humanity in his own person, in that he puts out of sight the end prescribed to him, and looks upon the means given for that purpose as though he were himself a means in that regard. (VE 27:659-60)

Avarice stands the proper relation between humanity and Mammon on its head: money becomes an end in itself or a “visible divinity” (Marx, *Economic and Philosophical Manuscripts* 139)—while the miser is reduced to “a custodian....an instrument, a mere means to no end, like a watchdog...” (VE 27:659).

Kant argues that money, pursued for its own sake, amounts to a form of *alienation*: Misers “nourish themselves on the thought of the enjoyment they have in their power; they all go about in fine clothes, ride in a carriage with six horses and eat twelve-course dinners every day; but all this in thought only....Possession of the wherewithal serves them in place of the real possession of all pleasures” (VE

27:400). And a few pages later: “The miser literally takes an immediate pleasure in money though it is nothing but a mere means. It is simply a crazy dream of possibility, to make use of it. The advantage of using the money is never realized” (VE 27:402). As a result, the miser “possesses money in *potentia* only, not in *actu*. Imagination, for him, replaces enjoyment” (VE 27:606).

Here is Marx roughly half a century later:

Everything which the political economist takes from you in life and in humanity, he replaces for you in *money* and *wealth*; and all the things which you cannot do, your money can do. It can eat and drink, go to the dance hall and the theatre; it can travel, it can appropriate art, learning, the treasures of the past, political power—all this it *can* appropriate for you—it can buy all this for you: it is true *endowment*. Yet being all this, it is *inclined* to do nothing but create itself, buy itself; for everything else is after all its servant. And when I have the master I have the servant and do not need his servant. (Marx, *Economic and Philosophical Manuscripts* 119).

The “great attraction” of avarice for Kant resides in the fact money consists “in mere alienation; that taken for itself, as money, it has no specific use or utility, but can be set in relation to all things alienable, and employed in every transaction, as a means thereto” (VE 27:658-9). Unlike (the young) Marx, however, Kant does not view the alienating effect of money as *inherently* degrading: it becomes so only when it assumes the place of humanity as an end in itself by demoting it to a mere servant.

Marx’s verdict is far more damning: “money is the universal, self-sufficient *value* of all things. Hence it has robbed the whole world, the human world as well as nature, of its proper worth. Money is the alienated essence of man’s labor and life, and this alien essence dominates him as he worships it” (Marx, “On the Jewish Question” 24).

Nevertheless, the difference between the two positions should not be overstated: in the *Anthropology*, Kant classifies the “mania for possession” as a dangerous social passion that is “if not *always morally reprehensible*, completely banal” (VA 7:274, emphasis added). We will have more to say about this when we discuss the passions in Chapter 4, § I.

6. **Servility.** Respect for oneself, as opposed to servility, is a duty “with reference to the dignity of humanity in us” (MS 6:436). In Chapter 1, § IV, we saw how the pursuit of power, wealth and honour in society causes individuals to constantly seek a comparative advantage over others. Kant term for this is “*unsociable sociability*” (I 8:21). At the same time, humanity or rational nature as an end in itself, possesses “a *dignity*” or “absolute inner worth” that places individuals on equal footing with each other (MS 6:434-5).

This results in a tension between the two outlooks: “All human beings share alike and equally in this incomparable worth, yet we have a powerful natural tendency to self-conceit, to value ourselves, our welfare, and our inclinations above others, and to treat...[them] as a mere means to our own ends” (Wood, KE 174). Therefore Kant believes that we should value ourselves “by a low and a high standard” (MS 6:435).

With regard to others, we should (a) not rank our worth above theirs or try to “equal or surpass others in this respect, believing that in this way one will get an even greater inner worth” (MS 6:435). Kant refers to this as “*ambition*” and considers it a direct violation of our duty to others (MS 6:435). Once again, it is worth noting that, for Kant, having *the mere end* of surpassing others in order to gain an advantage over them is contrary to duty. And (b) we should not belittle our own worth by flattery, obsequiousness or sycophancy in order to curry favour or ingratiate ourselves to others. This servility (false humility) is “contrary to one’s duty to oneself since it degrades one’s personality” (MS 6:436). Thus “moral self-esteem, which is founded on the worth of humanity, must never be based on a comparison with others, but only on comparison with the moral law itself” (VE 27:349).

With regard to the dignity or absolute worth of humanity vis-à-vis the moral law, we should (a) not rank our moral worth above the law, or fail to compare ourselves with it altogether. This is known as “*moral arrogance*” (MS 6:435). Nor should we (b) “waive any claim to moral worth” in ourselves before

the law, whether this takes the form of an abject humility, or the idea that in doing so we acquire a borrowed worth from it—since this amounts to a “morally-false *servility*” (MS 6:435).

True humility (*humilitas moralis*), for Kant, lies in the “sincere and exact comparison of ourselves with the moral law” as its *subjects*, while simultaneously esteeming “our capacity for internal lawgiving” as its *authors* (MS 6:436). Wood organizes Kant’s duty against servility into five general requirements in the order that they appear in at the end of Kant’s discussion:

- (1) Be no man’s lackey. — Do not let others tread with impunity on your rights.
- (2) Contract no debt for which you cannot give full security.
- (3) Do not be a parasite or a flatterer.
- (4) Complaining and whining, even crying out in bodily pain, is unworthy of you.
- (5) Kneeling down or prostrating oneself on the ground, even to show your veneration for heavenly objects, is contrary to the dignity of humanity. (MS 6:436)

Let us pause briefly to consider the last of these. Much is made of Kant’s Pietistic³⁷ upbringing, fuelling the charges of “rigorism” and entrenched religious prejudices in his moral philosophy. For example, in a passing remark, MacIntyre notes that “Lutheran pietists brought up their children to believe that one ought to tell the truth to everybody at all times, whatever the circumstances or consequences, and Kant was one of their children” (MacIntyre, “The Nature of the Virtues” 130). Goethe, upon reading Kant’s *Religion within the Boundaries of Mere Reason*, in a famous letter to Herder, accused Kant of appeasing his Christian critics by reintroducing the notion of original sin into his philosophy: “Kant, who required a long life to purify...his philosophical mantle from all sorts of slovenly prejudices, has wantonly tainted it with the stain of radical evil, so that Christians too might be attracted to kiss the hem” (qtd. in Louden, *Kant’s Impure Ethics* 134).

³⁷ Pietism was a 17th century movement within the Lutheran church that stressed biblical study, work, duty and prayer. A Pietist school was established in Königsberg, where Kant was sent at the age of eight.

There are two reasons to reject these charges in all their forms. The first is that Kant's religious writings landed him in trouble with the authorities, precisely for the Enlightenment spirit he is accused of betraying. An official letter from King Frederick William II accused Kant of misusing his philosophy to "distort and disparage many of the cardinal and basic teaching of the Holy Scriptures and Christianity" and exacted from him a written guarantee that he would no longer write or lecture on religious matters (qtd. in Wood, General Introduction xx-xxii). It is the second reason, however, that interests us.

Deeply anticlerical and a fierce critic of the institutional orthodoxy of his day, Kant regarded all religious creeds, rituals, prayers and devotional observances, including prostrating oneself before heavenly objects as idolatry since "you...humble yourself, not before an *ideal* represented to you by your own reason, but before an *idol* of your own making" (*MS* 6:437). This species of servility stems from our tendency to anthropomorphize God and subsequently picture our duties towards Him as duties to others (*R* 6:168-9; *VE* 27:327-8).

Therefore, in the *Religion*, we find the following claim "as a principle requiring no proof":
"Apart from a good life-conduct, anything which the human being supposes that he can do to become well-pleasing to God is mere religious delusion and counterfeit service to God" (*R* 6:170).

Kant argues that religions have sacrifices, fasts, pilgrimages, penances and castigations, "whereby their members demonstrate that they are ready to obey orders. They are mere observances, which have no goodness whatever and are no help to anyone. All religions are full of them" (*VE* 27:328). And "the more the *cultus* [worship] is overloaded with observances, the emptier it is of moral training" (*VE* 27:329).

Thus, Kant shrewdly observes, "*cultus* is an invention of men, and then having two ways of pleasing God, by morality and by *cultus*, they resort to the latter to replace the former; for if men are not punctilious in regard to morality, they are all the more so in regard to *cultus*" (*VE* 27:329-30). They may, like the miser, offer God "sacrifice by lip service" (which costs them nothing), or give up worldly

goods in the manner of hermits and monks, or mortify themselves as ascetics do—offering everything to God, *but their moral disposition* (R 6:172-3).

Kant offers the example of a pharisaical shopkeeper who observes his daily devotions in full, only to cheat his first customer of the evening. What good, asks Kant, are “a couple of pious ejaculations as he passes a church door” (VE 27:332)? “Such religion,” he concludes, “is more harmful than every kind of irreligion” (VE 27:332).

Imperfect Duties to Oneself:

7. **Natural Perfection.** Kant divides the imperfect duties we have to ourselves into duties of natural and moral perfection. Unlike the perfect duties we have been discussing, these are wide or meritorious duties whose omission, on any given occasion, is not morally blameworthy unless we fail to adopt the end in question altogether.

Under natural perfection, Kant outlines three powers as “a means to all sorts of possible ends”: (1) “*powers of spirit*,” (2) “*powers of soul*” and (3) “*powers of the body*” (MS 6:444-5).

Powers of spirit involve ratiocination and include “mathematics, logics and the metaphysics of nature” as well as “theoretical philosophy” in general (MS 6:445). Powers of soul are concerned with the understanding and include “memory, imagination...learning [and] taste,” while powers of the body are somatic in nature, comprising fitness, health, vitality and the “*basic stuff* (the matter) in a human being, without which he could not realize his ends” (MS 6:445).

Kant, once again, relies on FH to ground our duty to promote these natural powers in ourselves: “a human being has a duty to himself to be a useful member of the world, since this also belongs to the worth of humanity in his own person, which he ought not to degrade” (MS 6:446). This is done by cultivating his natural powers since “he owes it to himself (as a rational being) not to leave idle and...rusting away the natural dispositions and capacities that his reason can someday use” (MS 6:444).

These are two illuminating points in Kant's treatment of our duty to natural self-perfection.

First, being a wide duty,

which of these natural perfections should take *precedence*, and in what proportion...it may be a human being's duty to himself to make these natural perfections his end, are matters left for him to choose in accordance with what sort of life he would like to lead and whether he has the powers necessary for it.
(MS 6:445)

This wide latitude (both in the *content* and *degree* of the ends we should promote) means that our "ground projects" or those activities and pursuits which give our life meaning are in no way jeopardized by the duty to cultivate our natural powers or the spectre of an impartial morality: "we may take up weight-lifting, learn to play the violin, study quantum physics, or learn how to interpret the poetry of Wallace Stevens" (Wood, *KET* 328). Kantian ethics does not dictate which natural perfections we should promote, let alone how and to what extent we should promote them.

This leads us to a second, more curious point: Kant includes "powers of the body" as one of the three capacities we have a duty to cultivate in ourselves. Combined with the latitude we have in choosing and promoting these natural perfections, there is *no* precedence given to the other two powers over the former. If this is correct, then the standard image of Kantian ethics as hostile towards sensuality and the body, compared to reason (powers of spirit) and the understanding (powers of soul), is cast into doubt.

"Many visionary moralists think", Kant tells us, "[that] by weakening and removing the body's sensuality, to renounce everything that its sensuous enjoyment promotes, so that thereby the animal nature of the body would be suppressed...spiritual life...might already be anticipated here" (VE 27:379).

For example, in Epictetus's *Enchiridion*, we find the following passage: "It shows a lack of refinement to spend a lot of time exercising, eating, drinking, defecating or copulating. Tending to the body's needs should be done incidentally, as it were; the mind and its functions require the bulk of our

attention" (Epictetus, *Discourses and Selected Writings* 240). Epictetus is mild in his exhortation, whereas Kant has in mind not only acts which involve an abjuration (or gradual divestment) of sensuality, but also self-mortification. For Kant,

all such practices, which include...fasting and [corporal] chastisements, are fanatical and monkish virtues, which merely emaciate the body. The perfection of bodily discipline consists in a man being able to live in accordance with his vocation. The body must be certainly subjected to discipline, but it must not be destroyed by men, nor must its forces be impaired. (VE 27:379)

Kant's main concern revolves around the ability of the "powers of the body" to realize possible ends. Sensuality, in this context, is only censured if it *impairs* this ability: we must not etiolate our body by starving it, for example, nor should we "go too far the other way; for flabbiness represents an incapacity" (VE 27:380). Kant's emphasis on the ability to realize possible ends, results in a two sub-duties of bodily discipline: "moderation in its diversions, and sufficiency in regard to its genuine needs" (VE 27:380). Needless to say, this is far from the *disdain* for the body often attributed to Kant.

8. Moral Perfection. Kant divides the duty to increase our moral perfection into two parts. The first involves "the *purity (puritas moralis)* of one's disposition to duty" and has to do with our ability to act from the motive of duty (MS 6:446). The second part is related to "fulfilling all one's duties and...attaining completely one's moral end with regard to oneself" (MS 6:446).

The most striking, and often overlooked, element of Kant's duty of moral self-perfection is that it is a *wide* or *meritorious* duty. This translates into a duty to "*strive* for...perfection, but not to *reach* it" (MS 6:446). Applying this notion of an "imperfect duty to be perfect" to the two parts outlined above, yields some surprising results (MS 6:447).

It means (a) that "a person who does his narrow duties from some motive other than duty, and whose striving to improve himself on this point is only minimal and even unsuccessful, is not blamable" and (b) "we are not blamable for remaining morally imperfect, and our efforts to improve ourselves

morally are meritorious rather than strictly required” (Wood, “Duties to Oneself 243-4). This underscores our contention in Chapter 1, § I, that a good will is not necessarily one that acts from duty. We have an *imperfect* duty to strive to make duty alone a sufficient motive for our actions, but if we fail to achieve this end we incur no blame, and our efforts remain meritorious.

Yet another surprising consequence is that we can draw on other incentives in fulfilling our duty (as long as we are actively striving to make duty alone a sufficient motive), since “virtue here has the abiding maxim of *lawful* actions, no matter whence one draws the incentives that the power of choice needs for actions” (R 6:47). This stems from the fact that “a human being who incorporates...purity into his maxims, though on this account still not holy [i.e., pure in motive]...is nonetheless on the endless progress toward holiness” (R 6:46-7).

This concludes our survey of Kant’s duties to oneself. Before moving on, there are two key points to note regarding these. First, all duties to oneself, being either perfect or imperfect, “refer to the right of humanity in a person, or to the end of humanity in him” (VE 27:603). In both cases, they are derived from FH. Nelson Potter puts the point as follows:

The relation of DOS [duties to oneself] to the second formulation of the CI is more direct. The doctrine of DOS is a *spelling-out* of the nature of the moral self, which gives us a deeper understanding of the idea of respect for persons as mentioned in the second formulation of CI. The list of duties to oneself can be read simply as a list of ways one ought to act so as to respect one’s own moral person. Hence the Kantian doctrine of DOS can be taken as a *direct application* of the idea of respect for persons, where the person to be respected is the agent’s own. (Potter, “Duties to Oneself” 378, emphasis added)

Second, Kant views duties to oneself as the basis for all other ethical duties. In the *Metaphysics of Morals*, he argues if there were no duties to oneself “then there would be no duties whatsoever, and so no external duties either” (MS 6:417). Because the free self-constraint required for setting moral ends “proceeds in every case from my own practical reason” humanity grounds our obligations to others “only insofar as I at the same time put myself under obligation” (MS 6:417-8). Not only does humanity

obligate us in this dual sense—it also acts as the *wellspring* for our duties to others: the obligatory end of our self-perfection helps promote our duties of respect and love by realizing all sorts of possible ends in accordance with virtue. We turn to these duties now.

Duties to Others: Respect (Against)

Kant divides duties to others into duties of *respect* and duties of *love*. Duties of respect correspond to perfect (strict or narrow) duties—fulfilling them is a duty “*that is owed*” (MS 6:448-50), while duties of love correspond to imperfect, or wide, duties whose performance is *meritorious*.

Earlier we saw that Kant appeals directly to FH when he defines respect as “the *maxim* of limiting self-esteem by the dignity of humanity in another person” (MS 6:449). “Every human being,” according to Kant, “has a legitimate claim to respect from his fellow human beings and is *in turn* bound to respect every other....he is under obligation to acknowledge, in a practical way, the dignity of humanity in every other human being” (MS 6:462). Because respect is grounded in the absolute worth of rational nature as an end in itself, it is owed to every individual by virtue of humanity: “Humanity...is worthy of respect, and even though somebody may be a bad man, the humanity in his person is entitled to respect” (VE 27:373). Duties of respect are expressed *negatively* through the prohibition of the following vices: arrogance, defamation, and ridicule (MS 6:465). Like all social vices born of our unsociable sociability, these revolve around our pursuit of power, wealth and honour in order to gain precedence over others.

9. **Arrogance.** Kant defines arrogance (*superbia*) as “the inclination to be always *on top*,” a kind of ambition “in which we demand that others think little of themselves in comparison with us” (MS 6:465). The first distinction to note is between arrogance and pride proper (*animus elatus*). Pride proper is a love of honour that “will yield nothing of one’s dignity in comparison with others” (MS 6:465). Arrogance, however, not only “demands from others a respect it denies them” (MS 6:465)—it insists that they share in our *contempt* for their worth in comparison to ours. Thus “contempt,” for Kant,

stands diametrically opposed to respect since it “is based on a judgement that denies to the other any legitimate moral worth” (VE 27:709).

In his *Lectures on Ethics* Kant draws a further distinction between arrogance and moral self-love (*philautia*). Moral self-love is “an inclination to be content with one’s perfections” whereas arrogance “makes an unwarranted pretension to merit” (VE 27:357). Because the first takes pride in its perfections, while the second sees in itself *no* imperfections whatsoever, Kant views arrogance as “a far more damaging defect” (VE 27:357).

Arrogance is a violation of one’s duty to others, because (a) it debases humanity in their person by treating it with *contempt* (in order to elevate oneself, by comparison, above them) and (b) its presumption to worth is not grounded in *equality* with others by virtue of their humanity, but by seeking precedence “in things of no account”: “pre-eminence in trifles”, “smart clothes”, “fine carriage[s]”, “titles”, “positions”, “appear[ing] genteel”, etc. (VE 27:457-8).

In the *Anthropology*, Kant classifies arrogance as a type of “*buffoonery*” that frustrates its own ends, since “it is foolish to expect others to attach little value to themselves in comparison with me” and the more we insist on the latter, the lower we fall in their estimation (VA 7:210-1). Hence the arrogant person is despised as a “*conceited ass*” for the *same* self-important pretensions he imagines will secure his superiority (MS 6:465). In a direct reference to our unsociable sociability, Kant concludes that “the fault [of arrogance] always lies in this, that the agent makes his self-estimate on a comparison, not with the law, but with his fellow-man” (VE 27:708).

10. **Defamation.** Defamation (*obtretractio*) or “backbiting” involves an “immediate inclination, with no particular aim in view, to bring into the open something prejudicial to respect for others” (MS 6:466). Wood simplifies Kant’s definition by characterizing the former as “the desire to blame others and expose them to blame” (Wood, “Duties to Oneself” 246). Kant distinguishes defamation, which is the

spreading of *true* information that is injurious to others, from slander (*contumelia*), which involves the spreading of *false, calumnious* reports about others.

There are two types of defamation: (1) the intentional spreading (*propalatio*) of something that detracts from another's honor" and (2) "a mania for spying on the morals of others (*allotrio-episcopia*)" (MS 6:466). Both of these are contrary to the respect owed to humanity, the first, because it *diminishes* respect for humanity and is, on a broader scale, conducive to misanthropy—the second, because we take "malicious pleasure" in monitoring and exposing the faults of others in order to elevate ourselves above them by comparison (MS 6:466).

Kant argues that our duty against defamation comes in three forms: (a) a duty not to take malicious pleasure in the shortcomings of others, but to "throw a veil of benevolence over their faults" by (b) *softening* our judgment of them and (c) keeping our disapprobative views to ourselves (MS 6:466).

Notice that even within the confines of a strict or negative duty, Kant is concerned not only with the respect owed to others (which we can discharge by simply abstaining from defamation), but also with their *well-being* and self-worth.³⁸ This is evident in his insistence that we have a duty to another "not to censure his errors by calling them absurdities, poor judgement and so forth, but rather to suppose that his judgement must yet contain some truth and to seek this out...to preserve respect for his own understanding" (MS 6:463).

Once again, this leads Kant to some unexpected conclusions: "We must be blind to the other's faults, for otherwise he sees that we have lost respect for him" which, in turn, might undermine his own confidence and self-respect (VE 27:452). Moreover, if we *must* call attention to another's faults "kindness, benevolence and respect must prevail" (VE 27:452). For Kant, all faultfinding that is not mediated by love and good-will is captiousness.

³⁸ "The happiness of others also includes their *moral well-being* (*salubritas moralis*) and we have a duty, but only a negative one, to promote this" (MS 6:394).

11. **Ridicule.** Ridicule whether it be “*wanton faultfinding*” or “*mockery*” is “the propensity to expose others to laughter...[by] making their faults the immediate object of one’s amusement” (*MS* 6:467). Ridicule differs from defamation in finding *amusement* in the public disclosure of another’s faults.

Kant distinguishes ridicule from banter, which is free of malice and found among friends who target each other’s personal habits and quirks in good-natured badinage. Unlike banter, ridicule is a “kind of malice” that rejoices in exposing the real faults of others for our own amusement and thus it “has something of fiendish joy in it” (*MS* 6:467). For this reason, Kant views ridicule as one of the more serious violations of our duty of respect to others.

Kant also grounds our duty against ridicule on reasons which anticipate his notion of a shared kingdom of ends. In his *Lectures on Ethics*, he argues that

ridicule and *reductio ad absurdum*...obviously damage the possibility of communication in human relations, and a *spiritus causticus* [a sarcastic outlook] is thus objectionable since it at once combines with its judgment the inclination to make mock of the other’s action. (*VE* 27:705)

The threat posed by ridicule to the “possibility of communication” amounts to a threat to reciprocity. And, as we shall see in Chapter 4, § V, “people who enter into relations of reciprocity must be prepared to share their ends and reasons; to hold them jointly; and to act together” (Korsgaard, *CKE* 196). Ridicule, in this sense, not only alienates us from the ends of others which we transform into a form of voyeuristic amusement—it precludes the very possibility of reciprocity and, with it, friendship and the shared kingdom of ends.

Both defamation and ridicule involve the spreading of true information that detracts from another’s respectability. Kant refers to these offenses against respectability as “*scandals*” (*MS* 6:464). “To *give scandal*,” according to Kant, is “contrary to duty” (*MS* 6:464). It is significant that Kant’s injunction against giving scandal amounts to (a) withholding *truthful* information, in cases in which not

doing so would diminish another's self-respect and (b) refraining from partaking in, or tempting another to, actions "for which his conscience could afterwards pain him" (MS 6:394).

This means that our duty to refrain from scandal is not limited to our personal interactions with others, but instead covers a wide range of cases, from censuring "the fiendish joy" we derive from watching less than talented *American Idol* contestants flounder on national television, to eschewing lurid tabloids and other forms of yellow journalism that seek to profit from the misfortunes of others.

Duties to Others: Love

The ethical duties that pose the most compelling challenge to the traditional view of Kantian ethics are Kant's duties of love, which he divides into beneficence, gratitude and active sympathetic participation (*Teilnehmung*). Like their predecessors, the opposing vices of envy, ingratitude and malice also have their basis in our unsociable sociability, displaying a "veiled" hatred for human beings (MS 6:458-61). However, before we turn to these as well as Kant's wide duties to others, we must enquire into his account of love.

Under the section "*Love of Human Beings*," Kant argues that "*love* is a matter of *feeling*, not of willing, and I cannot love because I *will* to, still less because I *ought* to (I cannot be constrained to love); so a *duty to love* is an absurdity" (MS 6:401). This gives rise to an immediate problem: if we cannot will to love, or be constrained to love by duty, how can we have duties of love towards others?

The answer lies in Kant's characterization of our duty to others as a duty of *practical love*: "**love** is not to be understood as a 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)" (MS 6:449). Lending credence to this account is the fact that, throughout his ethical writings, Kant associates the duty of practical love with the Christian commandment to love one's neighbour (G 4:399; KpV 5:83; MS 6:402; VE 27:335, 720-7).

Commentators have usually understood these points to mean that practical love, for Kant, refers to a “policy of benefiting others on moral principle from the motive of duty” (Wood, *KET* 270). Melissa Seymour cites Paul Guyer and Allen Wood as proponents of this reductive view, which she calls the policy of beneficent action or PBA account of practical love.

PBA Account: The duty of practical love is a duty to adopt a general policy of unselfishly advancing the welfare and happiness of others. (Seymour, *DL* 29)

Seymour admits that certain passages³⁹ in the *Groundwork* lend support to the PBA account, but offers three arguments against this reductive reading of practical love. The first draws on a passage from the second *Critique*, in which Kant identifies practical love not with beneficent action, but with performing one’s duty *gladly*:

[To] Love God above all, and your neighbour as yourself...cannot be commanded, for it is not in the power of any human being to love someone merely on command. It is therefore only *practical love* that is understood in that kernel of all laws. To love God means, in this sense, to do what He commands *gladly*; to love one’s neighbour means to practice all duties toward him *gladly*. But the command that makes this a rule cannot command us to *have* this disposition in dutiful actions but only to *strive* for it. (*KpV* 5:83)

Support for Kant’s identification of practical love with the affective disposition involved in performing one’s duty gladly can also be found in his *Lectures on Ethics*:

“Heart is the *principium* of moral disposition. Hence the disposition to performance of duties is to be cultivated, and this is what the teacher of the gospels says: that we should do everything from the love of God. But to love God is to do His commandments gladly” (*VE* 27:274).

“To love God in a practical sense is to obey his commandments gladly” (*VE* 27:335).

“...he who loves another wishes him well, but without owing it to him: he acts, rather, from a willing disposition, gladly, and from his own impulse” (*VE* 27:413).

³⁹ See, for example, *G* 4:399.

In all four examples, Kant's emphasis is on the *affective disposition* associated with practical love. Loving one's neighbour "in a practical sense" is not understood in terms of benefiting her or advancing her welfare (although these are undoubtedly present), but in doing so gladly, from a willing disposition. Of course, practical love cannot command us to have this disposition; it can command only its *cultivation*. This means that whatever affective account we give of practical love will not be a success account but will, like our pursuit of virtue, involve an endless progress towards this disposition.

Seymour's second argument concerns a short passage in the *Metaphysics of Morals* in which Kant seems to explicitly deny the PBA identity between practical love and unselfish benevolence: "...benevolence (*amor benevolentiae*), as conduct, can be subject to a law of duty. However, unselfish benevolence toward human beings is often (though very inappropriately) also called *love*" (MS 6:401). Clearly, if it is "very inappropriate" to conflate love with unselfish benevolence, then the PBA account cannot be correct.

This highlights a third, and more troubling, aspect of the PBA account. In addition to practical love, Kant includes beneficence, gratitude and active sympathetic participation as specific duties of love. Kant defines beneficence as "promoting...the happiness of others in need without hoping for something in return" (MS 6:453). But this results in a redundancy of duties: both practical love and beneficence enjoin us to benefit others unselfishly. As a result, Paul Guyer, who supports the PBA account, admits that it renders the two duties "virtually identical" (qtd. in Seymour, *DL* 33-4). The fact that Kant has a duty of beneficence is not open to dispute (MS 6:452-4). Therefore, if we accept the PBA account, "we must accept the unattractive and implausible view that Kant assigned two different names to the same duty without ever acknowledging their equivalence" (Seymour, *DL* 34).

Following Seymour, we can reject the PBA account of practical love in favour of one involving a direct duty to cultivate a benevolent disposition. Seymour refers to this as the cultivation account.

Cultivation Account: The duty of practical love is to cultivate a benevolent disposition and practical desires. (Seymour, *DL* 40)

A “benevolent disposition” is one which “finds satisfaction in the well-being (*salus*) of human beings considered simply as human beings, [and] for whom it is *well* when things go well for every other” (*MS* 6:450). Kant refers to someone who harbours such a disposition as a “*friend of humanity*” (*MS* 6:450). Seymour defines “practical desire,” which makes up the second component of the cultivation account, as the “inclination to perform beneficent actions” (Seymour, *DL* 40).

There are two important points to note regarding the cultivation account. First, it involves a duty to *cultivate* or *strive for* a benevolent disposition and practical desires, since “obligations with regard to moral feeling can be only to *cultivate* it and strengthen it” (*MS* 6:399). We saw in our discussion of virtue, how holiness serves as an archetype of moral perfection for virtue to approximate. There we were concerned with the aspect of the holy will related to the purity of its disposition in acting from the moral law. But there is another feature of the holy will that practical love commands us to approximate: it fulfills the moral law *gladly*, with a joyous heart.

“The law of all laws,” Kant tells us, “presents the *moral disposition* in its complete perfection, in such a way that as an ideal of holiness it is not attainable by any creature but is yet the archetype we should strive to approach in an uninterrupted but endless progress” (*KpV* 5:83, emphasis added). In striving towards the complete perfection of the holy will, virtuous individuals strive to act from *duty*, friends of humanity, to act from *love*. These are merely two aspects of the same holy will, which acts *joyously* from a pure disposition in accordance with the moral law. Therefore, an agent who acts from duty comes close to approximating a holy will, but an agent who acts from duty *gladly* comes even closer (Seymour, *DL* 130).

The second point follows on the heels of the first: if practical love commands us to cultivate a benevolent disposition in an endless effort to approximate the disposition of a holy will, it presents us

with a *direct* duty to do so since “a heart joyous in the *compliance* with its duty...is the sign of genuineness in virtuous disposition” (R 6:23n). Many commentators, such as Nancy Sherman⁴⁰ and Henry Allison, have argued for an *indirect* duty to cultivate positive feelings including sympathy and love as a means of “facilitating our ability to act from duty” (Allison, *Kant’s Theory of Freedom* 166).

The cultivation account parts ways with the former by maintaining that “cultivating the disposition prescribed by practical love is simply constitutive of genuinely embracing the obligatory end as one’s own” (Seymour, *DL* 44). If this is correct, then, under the aegis of practical love, “good progress [in approximating a holy will] must...effect a joyous frame of mind, without which one is never certain of having *gained* also a *love* for the good” (R 6:25n). In other words, a failure to strive for a benevolent disposition amounts to a failure to genuinely adopt the happiness of others as one’s end.

So far we have been concerned with differentiating Kant’s account of practical love from his duty of beneficence by highlighting the affective nature of the former. This is not meant to suggest that the two duties are unrelated or disparate: the cultivation account of practical love directs us to cultivate practical desires in the form of inclinations to perform beneficent actions. Similarly, the duty of beneficence is not limited to well-doing “but also to loving others with well-wishing, and well-liking, too” (VE 27:418). We turn to this duty now.

12. Beneficence. Kant distinguishes *benevolence* which he describes as the “satisfaction in the happiness (well-being) of others” from *beneficence* which is “the maxim of making others’ happiness one’s end” (MS 6:452). To be beneficent, according to Kant, is “to promote according to one’s means the happiness of others in need, without hoping for something in return” (MS 6:453).

This gives rise to three questions: (1) how much of our welfare should we sacrifice to others; (2) how are we to understand the qualifier “in need”; and (3) how should we go about promoting this end?

⁴⁰ See Nancy Sherman, “The Cultivation of Emotions as Supports for Duty,” *Making a Necessity of Virtue* (New York: Cambridge University Press, 1997): 141-58.

The question of how much one should sacrifice to promote the happiness of others “must be left to each person to decide...for himself” and according to his “true needs in view of his sensibility” (MS 6:393). Because Kant’s duty of beneficence is wide or meritorious duty, “it is impossible to assign determinate limits to the extent of this sacrifice” (MS 6:393). This does not mean, however, that we are free to promote the happiness of others in cases where doing so violates a *strict* duty to ourselves.

Kant outlines three such cases: (a) as we saw in Kant’s duties to oneself, *miserly avarice* which “restrict[s] *one’s own* enjoyment of the means to good living so narrowly as to leave one’s own true needs unsatisfied” is contrary to our duty to ourselves as moral beings (MS 6:432). Similarly, (b) “depriving oneself (fanatically) of enjoyment of the pleasures of life by exaggerated *discipline* of one’s natural inclinations” is slavish and a violation of humanity in our person (MS 6:452). And (c) any “maxim of promoting others’ happiness at the sacrifice of one’s own happiness, [or] one’s true needs, would conflict with itself if it were made a universal law” (MS 6:393). In other words, we *cannot* sacrifice so much of our welfare to promote the happiness of others that we severely restrict our own happiness, enjoyment of the means to good living or pleasures of life in the process.

Notice that, coupled with the playroom attached to wide duties, this qualification clears Kant of the charge that Kantian ethics devalues agents because it attaches no significance to their welfare. We may not have a *moral* duty to promote our own happiness, but we cannot, according to Kant, advance the happiness of others at the expense of our own.

Kant argues that we have a duty to promote the happiness of others “in need” (MS 6:453). How should we understand this qualifier? Are we to help others meet *basic* needs (e.g. food, shelter, clothing) or should we aid only those in *desperate* need? Baron and Fahmy argue against both these restrictive readings of need. For one, Kant identifies true needs with the “enjoyment of the means to good living” including “the pleasures of life” (MS 6:432, 452).

In his *Lectures on Ethics*, Kant maintains that “we ought not deprive ourselves of all amenities and pleasures, and enjoy none whatever; that would be a monk’s virtue, to forgo everything that is *proper to human life*” (VE 27:393, emphasis added). In the same passage, Kant distinguishes “amenities” from “necessities” and argues that while the latter is required for one to live, the former enables one to live *comfortably* (VE 27:393). The fact that both amenities and pleasures are “proper to human life” suggests that Kant understands “needs” in a broad, not restricted, sense.

A broad reading of “needs” identifies them with the subjective ends of others, namely, those ends attached to their personal projects and endeavours. This is why beneficence directs us to adopt the (permissible) ends of others as our own (MS 6:388, 450). Understood broadly, needs stem from the fact that “happiness consists in...enjoying and possessing the means to procure ourselves all ends that are even possible, and thereby to satisfy all wishes” (VE 27:497). This also explains why, when we aid others in securing *any one* of their permissible ends, we contribute to their happiness.

Kant’s most perceptive observations, however, are reserved for the question of how we should go about promoting the happiness of others. Kant outlines two cases in which he believes the duty of beneficence can no longer be regarded as meritorious (despite being a wide duty), to the point of doubting whether it can even be called beneficence at all. Both cases involve the rich or well-off.

According to Kant, “someone who is *rich* (has abundant means for the happiness of others, i.e., means in excess of his own ends) should hardly even regard beneficence as a meritorious duty on his part, even though he also puts others under obligation by it” (MS 6:453). The rich benefactor must (a) avoid humiliating the other in the process of rendering him a benefit and (b) “show [instead] that he himself is put under obligation by the other’s acceptance or honored by it” (MS 6:453). Thus, the benefactor must regard his beneficence as something *owed* to the other or, better still, confer the benefit in complete secrecy.

Commenting on the same passage, Baron and Fahmy draw parallels between Kant's remarks and "the Aristotelian idea that acting well involves not just doing a virtuous action, but doing it in the right way, and with the right tone and gesture. It is not enough to render aid; we need to do it well" (Baron and Fahmy, "Duties of Beneficence" 215). We will have more to say in support of this claim in Chapter 4, § III.

In the second case, Kant acknowledges that

having the resources to practice such beneficence as depends on the good of fortune is, for the most part, a result of certain human beings being favored through the injustice of the government, which introduces an inequality of wealth that makes others need their beneficence. Under such circumstances, does a rich man's help to the needy, on which he so readily prides himself as something meritorious, really deserve to be called beneficence at all? (*MS* 6:454).

Kant concludes that, if the unequal distribution of wealth, which condemns many to poverty, is due to certain individuals being favoured through *unjust* political structures or the "general injustice" then

if we...do a kindness to an unfortunate, we have not made a free gift to him, but repaid him what we were helping take away through the general injustice. For if none might appropriate more of this world's goods than his neighbour, there would be no rich folk, but also no poor. Thus even acts of kindness are acts of duty and indebtedness, arising from the rights of others. (*VE* 27:416).

We have not made a gift to him because "the provision made for us is universal" (*VE* 27:414).

Accordingly, "if I find...a table laden with food in the forest, I am not to suppose that it is solely for myself; I can partake of it, but must also be mindful of leaving something for others....I have an obligation to limit my consumption, and to bear in mind that nature has made these arrangements for everyone" (*VE* 27:414).

Comments such as these not only anticipate Kant's kingdom of ends, in which rivalry, inequality and social antagonism are replaced by cooperation, universal provisions and the pursuit of common ends (see Ch. 4, § V), but they also problematize the typical libertarian readings of Kant's political philosophy.

13. **Gratitude.** Gratitude is a “*sacred duty*” which “consists in *honoring* a person because of a benefit he has rendered us” (MS 6:454-5). Kant outlines two types of gratitude: *active* gratitude and *affective* gratitude (MS 6:455).

Active gratitude instructs us “in the least degree...to render *equal* services to the benefactor if he can receive them (if he is still living) or, if he cannot, to render them to others” (MS 6:456). Kant insists, however, that no matter what how great the services we provide in return, we can never discharge our obligation for the kindness received “since the recipient can never win away from the benefactor the *priority* of merit, namely having been the first in benevolence” (MS 6:455). Seymour finds Kant’s account troubling for two reasons: (a) it encourages us to think of gratitude in terms of bookkeeping and debt repayment and, paradoxically, (b) it places us *forever* under a debt that cannot be repaid (Seymour, DL 53-6).

Seymour comes close to recognizing the motivation behind Kant’s account of active gratitude, when she considers it to be “in part, evidence of Kant’s sensitivity to the ways in which social hierarchies create anxiety and the potential for humiliation” (Seymour, DL 55). Nevertheless, she finds his remarks by turns “disquieting” and “angst-ridden” concluding that, in the final analysis, his position is “untenable” (Seymour, DL 55).

Kant’s insistence that the recipient “stands a step lower than his benefactor” (MS 6:456, 458) stems from his unwavering commitment to *equality*. Kant’s remedy, in cases in which equality is compromised, rests in the idea of reciprocity. In the case of sexual relations, this takes the form of marriage⁴¹ in which “I hand over my whole person to the other, and thereby obtain the person of the other in place of it, I get myself back again” (VE 27:388). In the case of receiving a benefit, it involves rendering an equal or greater benefit to the benefactor.

⁴¹ As both Herman and Korsgaard have pointed out, “marriage as it has usually existed has hardly been a solution to *this* problem [of inequality]. The equality necessary for reciprocity is far more likely to be distanced even further by marriage, which has usually given the husband rights over his wife *additional* to those that accrue from the superior social position he has held as a man” (Korsgaard, CKE 195).

However, the most compelling evidence against Seymour's conclusion comes from Kant's own observation that active gratitude

involves *not* regarding a kindness received as a burden one would gladly be rid of...but taking even the occasion for gratitude as a moral kindness, that is, as an opportunity given one to unite the virtue of gratitude with love of man, to combine the *cordiality* of a benevolent disposition with *sensitivity* to benevolence...and so to cultivate one's love of human beings. (MS 6:456, emphasis added)

Here, we see that active gratitude is not an angst-ridden form of debt cancellation, but an occasion to demonstrate our appreciation for an act of moral kindness, without begrudging our benefactor her assistance. In other words, active gratitude is merely an expression of *affective* gratitude, which serves to draw us closer to others by cultivating our love of humanity. Kant refers to the duty to cultivate a grateful disposition, which is called for even in response to a "mere heartfelt benevolence," as *appreciativeness* (MS 6:455). Along with practical love, we now have two duties with a clear, affective component. In the third, this element is even more pronounced.

14. Active Sympathetic Participation (Teilnehmung). Seymour and Wood both argue that "the duty of 'sympathetic participation' deserves special mention...because appreciating its role in Kantian ethics will help to correct important elements in the prevailing false image of Kantian ethics" (Wood, *KE* 176). Traditionally, Kant's third duty of love has been understood as an *indirect* duty "we need to rely on in fulfilling our other obligatory ends" (Sherman, *Making a Necessity of Virtue* 142). Following Seymour, we reject this view in favour of a closer reading of *Teilnehmung* as a direct duty of active sympathetic participation with others.

The first problem we encounter is definitional. Mary Gregor translates Kant's third duty of love, *Teilnehmung*, as "sympathy." Therefore she renders the title of section, *Teilnehmende Empfindung ist überhaupt Pflicht* as "*Sympathetic Feeling is Generally a Duty*" (MS 6:456). But, as Seymour points out, in the opening passage of the following section Kant draws a clear distinction between sympathetic

sadness and joy (*Mitleid* and *Mitfreude*), and *Teilnehmung* when he writes “while it is not in itself a duty to share in the sufferings [*Mitleid*] as well the joys [*Mitfreude*] of others, it is a duty to sympathize actively [*Teilnehmung*] in their fate” (MS 6:467).

This means that we cannot equate *Teilnehmung* with sympathy as such since Kant conceives of the former as *distinct from* sympathetic feelings. This is evidenced by the fact that we have only an *indirect* duty to cultivate our sympathetic feelings of sadness and joy towards others, but a *direct* duty of active *Teilnehmung* in their fate.

Before settling on a more faithful translation, Seymour surveys the various meanings of *teilnehmen* (the verb from which *Teilnehmung* is derived): “be present at, attend, take part in, participate in, take an interest in, share in, and sympathize with” (Seymour, DL 61). Central to her understanding of *Teilnehmung* is Kant’s distinction between *humanitas aesthetica* and *humanitas practica*.

Kant defines *humanitas aesthetica* as “receptivity, given by nature itself, to the feeling of joy and sadness in common with others” (MS 6:456). *Humanitas practica*, on the other hand, “can be located...in the *capacity* and *will* to *share in others’ feelings*” (MS 5:456). Seymour, once again, takes issue with Gregor’s translation, specifically the word “sharing” [*mitteilen*], refining the definition to read: *Humanitas practica* “can be located in the capacity and will to **communicate** with each other in view of (with respect to) one’s feelings” (Seymour, DL 66, emphasis added). According to Kant, we have an obligation only to the latter.

The important point for our purposes is that *humanitas aesthetica* is unwilling and, therefore, unfree, while *humanitas practica* is practical in nature and thus furnishes us with a duty of love. But what exactly does this duty entail? Whereas *humanitas aesthetica* is passive in its receptivity to sympathetic feelings, *humanitas practica* prescribes an “**active sympathetic participation** in the fate of others” (Seymour, DL 73, emphasis added).

The duty is “active” because it is grounded in practical reason and, as such, must be willed by the agent. That it involves sympathetic “participation” follows from the fact that it is not enough to simply *harbour* a sympathetic feeling—one must actively share or “participate” in the joys and sorrows of others. We find strong evidence for this in Kant’s discussion of the vices opposed to our duties of love in his *Lectures on Ethics*:

Such vices [of hatred for human beings] are contrary to humanity; for just as the latter involves a participation in the person and state of the other, and is evinced in well-wishing, so these three vices involve a lack of participation....Hence they...are directly opposed to moral sympathy and indicate *inhumanity*. (VE 27:692)

Following her translation of the word *mitteilen*, Seymour locates sympathetic participation in human communication: “verbal, facial expression, gestures and even actions” (Seymour, DL 73). Kant’s emphasis on communication, fellowship and sharing can be seen in his explicit characterization of the former as “*communio sentiendi liberalis*” [the free communion of sentiment] (MS 6:456). The reciprocal enjoyment of humanity requires individuals to “stand together, to communicate not only their feelings and sensations to one another, but also their thoughts” (VE 27:677).

Kant approvingly quotes Terence’s well-known line: *homo sum humani a me nihil alienum puto* [I am a human being; nothing human is foreign to me] to capture the idea that what befalls others also befalls me (VE 27:677; MS 6:460). Accordingly, “I cannot limit my well-wishing to myself only, and must show it to be *active* in regard to every other” (VE 27:677, emphasis added). As such, Kant’s third duty of love anticipates his treatment of friendship and the kingdom of ends, which we will examine more closely in Chapter 4, § V.

If this account of *Teilnehmung* is correct, then we are also required to cultivate our sympathetic feelings actively *through* our participation with others, even if this causes us pain or discomfort (not to

adopt this end is contrary to duty). In a moving passage, Kant is clear that *active sympathetic participation* is

a duty not to avoid the places where the poor who lack basic necessities are to be found but rather **to seek them out**, and not to shun sickrooms or debtor's prisons and so forth in order to avoid sharing painful feelings one may not be able to resist. For this is still one of the impulses that nature has implanted in us to do what the representation of duty alone might not accomplish. (MS 6:457, emphasis added)

The last line of Kant's paragraph should not surprise us, since we saw earlier that, tied to the end of our moral perfection, we have only a wide or meritorious duty to act from duty alone as a sufficient motive. We are thereby permitted to draw from incentives to aid us in doing our duty—as long as we do not abandon this end altogether.

Perhaps the best proof that Kant's duties of love are principally social and affective in nature (that is, concerned with the cultivation of certain dispositions) is that the "loathsome family" of vices that oppose them all take the form of "malevolent disposition[s]" which have their basis in our unsociable sociability (MS 6:458-61).

The Malevolent Dispositions: 'Vices of Hatred for Human Beings'

There are three vices directly opposed to the duty of love: envy, ingratitude and malice.

Envy. Kant defines envy, the vice opposed to beneficence, as the disposition or "propensity to view the well-being of others with distress, even though it does not detract from one's own" (MS 6:459). Thus envy is distinct from "grudging" in which we begrudge others an advantage they have over us, since it is grieved merely at the thought of others *sharing our happiness* (VE 27:438-9). This makes envy "utterly detestable" (VE 27:440), "the basis of mistrust and reserve...the greatest obstacle to friendship" and an enemy of happiness throughout the world (VE 27:679).

When envy "breaks forth into action" it is called "envy proper" but Kant views the disposition itself, which involves "a reluctance to see our own well-being overshadowed by another's," as *devilish*

(MS 6:458; VE 27:440). The fact that envy is a disposition is supported by Kant's advice that "to keep the soul free from the vice of envy, we must try to bear every hardship, and, once it has befallen us, to extract from it the advantage that always resides in misfortune. It is up to us to put ourselves in a certain mood, which is a *voluntary chosen disposition*..." (VE 27:368, emphasis added).

Ingratitude. Kant's commitment to the love of humanity is underscored by his condemnation of ingratitude, which he defines as "*ingratitude proper*" when it extends to hatred for one's benefactor, and mere "*unappreciativeness*" when it does not (MS 6:459).

The nature of Kant's objection is informative. He argues that ingratitude "is a vice that shocks humanity" not because it fails to discharge one's debt to a benefactor (a failure of active gratitude), but because it "stands the love of humanity on its head...and degrades absence of love into an authorization to hate the one who loves" (MS 6:459). In other words, it amounts to a failure of *affective* gratitude. While "grateful dispositions are extremely lovable" (VE 27:443), *ungrateful* dispositions are "utterly repugnant to human nature" since they (a) "hate and persecute those who have done one a kindness," (b) deter individuals from well-doing, and (c) inspire misanthropy (VE 27:439-40).

Malice. Malice or *Schadenfreude* [malicious glee] involves taking an immediate pleasure in the misfortunes of others, and is directly opposed to active sympathetic participation. When malice goes so far as to *bring about* the downfall of others or when it plays an active role in the ills that befall them, it is "*malice proper*" (MS 6:460).

Malice, for Kant, represents the height of inhumanity and the most "cruel disposition" of the three vices: malicious individuals "laugh when others weep, and feel pleasure when others feel pain" (VE 27:443-4). Worse still, is malice proper which "does not remain a mere spectator of the other's sufferings, but...participates in them" and rejoices gleefully at their misery and plight (VE 27:697). Consequently, Kant regards the vice of malice as "the farthest removed from humanity" (VE 27:697). Malice not only violates our duty of active sympathetic participation, but also our duties of practical love

and beneficence since, rather than cultivating a benevolent disposition and loving our neighbours, we secretly pray for their downfall and endeavour to bring it about.

Kant's examples of malice range from trying to stir up strife within a marriage, to schoolchildren jabbing each other with pins or inflicting distress on hapless animals (by twisting their tails, for example), to cruel kings who derive pleasure from torturing their subjects (VE 27:440-4). In the Doctrine of Virtue, Kant also includes our "*desire for revenge*" as a species of malice (MS 6:460).

Kant's contention that three vices of hatred have their origin in our unsociable sociability can be easily gleaned from his comments: Envy is "a reluctance to see our well-being overshadowed by another's because *the standard we use to see how well off we are is not the intrinsic worth of our own well-being but how it compares to others*" (MS 6:459, emphasis added). Similarly, when one has rendered us a benefit, it arouses ingratitude

for all his *merita* in regard to our person, or fortunes, bring it about that we are obliged to him on that account...he has advantage over us, whereby he is elevated above our worth, and we, on the contrary have become *inferiores* in his regard; for ***assuming that the estimation of our self rests on comparative judgement with the other's worth***, this degradation displeases us. (VE 27:695-6, emphasis added)

Finally, Kant's explains the psychology of malice as follows: "we feel our own well-being and even our good conduct more strongly when the misfortune of others or their downfall in scandal is *put next to our condition, as a foil to show it in so much the brighter light*" (MS 6:460, emphasis added).

Kant's constant references to our unsociable sociability are important for two reasons: (1) anticipating our discussion in Chapter 4, § IV, they demonstrate that the true enemy of virtue does not lie in the natural inclinations, and (2) they allow us to better understand the role Kant's duties of love are meant to play in securing the kingdom of ends in which love of humanity deposes these social vices.

VII. How Much Latitude Do Wide Duties Permit?

So far we have been expounding Kant's duties of virtue in order to show that criticisms of Kantian ethics from the field of virtue ethics misfire when they claim that Kantian ethics has no place for sociality, character and the emotions. With Kant's rich taxonomy of virtues before us, we are now in a position to respond to these charges in Chapter 4. Before we turn to this task, however, there is one important question we have yet to address. This is the question of latitude.

Among the charges levelled against Kantian ethics, is the well-known charge of "rigorism," which suffers a severe blow with the introduction of Kant's wide (or imperfect) duties of virtue. Kant argues that ethics "can prescribe only the maxims of actions, not actions themselves, [and] this is a sign that it leaves a playroom (*latitudo*) for free choice in following (complying with) the law" (MS 6:390). Hence, the law cannot specify (a) what we do and (b) how much we do it. Far from being disarmed, however, the threat of "rigorism" re-emerges immediately in Kant's subsequent caveat:

But a wide duty is not to be taken as permission to make exceptions to the maxim of actions but only as permission to limit one maxim of duty by another (e.g., love of one's neighbour in general by love of one's parents), by which in fact the field for the practice of virtue is widened. (MS 6:390)

The question then becomes: how much latitude do wide duties permit? Following the secondary literature there are two possible answers: the first is based on a *rigorist* interpretation of Kant's warning, the second, on a *permissive* one.

In the *Groundwork*, Kant understands a perfect duty as one that "allows no exception to the advantage of inclination" (G 4:422n). The threat of "rigorism" introduced by Kant's caveat is that, despite duties of virtue being wide (or imperfect) duties, they might *not* allow any exception to the advantage of inclination. A rigorist interpretation is, therefore, one "which denies a general permission to make 'exceptions in favor of inclination'" (Seymour, *DL* 80), whereas a permissive interpretation allows for such exceptions. In terms of ends, to make an exception to the advantage of inclination simply

means “to decline an opportunity to promote an obligatory end from any motive that is neither moral nor malicious” (Seymour, *DL* 80).

The Rigorist Interpretation: The only permissible reason for not performing a wide ethical duty is (a) that one is performing another wide ethical duty, or (b) that one is performing an act required by a strict or narrow duty.

The rigorist interpretation severely restricts the latitude we have in carrying out our wide ethical duties to the promotion of (a) the happiness of others or (b) our own perfection (the two obligatory ends).

Since we have been arguing that the duties of virtue involve an endless progress towards moral perfection, “they are not the kinds of ends we can ever hope to achieve and be done with” (Seymour, *DL* 87). In light of these unremitting duties which can never be “done with” or discharged, the rigorist interpretation seems to leave us with no room to pursue our own projects and other *non-moral* commitments.

It should come as no surprise to us by now that the textual evidence weighs heavily against this interpretation. The first passage, which we have already seen, explicitly rejects these rigoristic implications:

Providing for oneself to the extent necessary just to find satisfaction in living (taking care of one’s body, but not to the point of effeminacy) 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 these are opposed to a human being’s duty to himself. (*MS* 6:452)

Not only does a “maxim of promoting others’ happiness at the sacrifice of one’s own happiness, [and] one’s true needs...conflict with itself if it were made a universal law” (*MS* 6:393), but

lawgiving reason, which includes the whole species (and so myself as well) in its idea of humanity as such, includes me...with all others in the duty of mutual benevolence, in accordance with the principle of

equality, and *permits* you to be benevolent to yourself on the condition of your being benevolent to every other as well. (MS 6:451)

However, the most compelling evidence against the rigorist interpretation (as well as *any* rigorist reading of Kantian ethics), which allows nothing to be morally indifferent, is Kant's unequivocal denunciation of "fantastic virtue":

[T]hat human being can be called fantastically virtuous who allows *nothing to be morally indifferent (adiaphora)* and strews all his steps with duties, as with mantraps; it is not indifferent to him whether I eat meat or fish, drink beer or wine, supposing that both agree with me. Fantastic virtue is a concern with petty details which, were it admitted into the doctrine of virtue, would turn the government of virtue into tyranny. (MS 6:409)

Kant condemns not only "fantastic virtue" but also the "micrological conscience" which attends it, and "is burdened with many small scruples on matters of indifference" (VE 37:356). We can, accordingly, reject the rigorist interpretation, which threatens to "turn the government of virtue into tyranny" by "*allowing nothing to be morally indifferent*" in favour of the permissive one.

The Permissive Interpretation: A wide duty allows us to do as we please on certain occasions, even if this action is not prescribed by an ethical duty or fails to instantiate a specific ethical duty as long as (a) we do not relinquish the *maxim* of the obligatory end in question, or (b) violate a strict or narrow duty.

Here, Kant's warning that "a wide duty is not to be taken as permission to make exceptions to the maxim of actions" (MS 6:390), is understood to mean that we cannot relinquish the *maxim* of the obligatory end in favour of specific instantiations of that end—not that we must *always* be instantiating these ends. For example, I "may focus primarily on the happiness of a few others [e.g. my parents, spouse, children, etc.], but I may not have a policy of ignoring the needs of all others" (Baron, KEWA 94).

The idea is as follows: narrow duties notwithstanding, Kant's warning is *not* that our wide duties can only be limited by other wide duties, but that we cannot pick a specific *instantiation* of an obligatory

end as pretext for dismissing the *maxim* (or all other instantiations) of that obligatory end. This is why Kant says that “love of one’s neighbour in general” (the end of promoting the happiness of others) can only be *limited*—not relinquished—by “love of one’s parents” (an instantiation of *that* obligatory end) (*MS* 6:390).

In other words, we are free, on almost any occasion, not to instantiate a wide duty as long as we continue to hold the maxim of the relevant obligatory end. What we cannot do is regard specific instantiations of that wide duty as making exceptions to, or *releasing us from*, these obligatory ends. Why we would be tempted to make such exceptions to the two obligatory ends in the first place, especially given their wide or meritorious nature, and the role moral feelings play in obviating this possibility, is the question we turn to next.

Chapter 4: From the Inclinations to the Kingdom of Ends

In the previous chapter we examined Kant's complete taxonomy of virtues, and argued against any "rigorist" interpretation of Kant's wide or imperfect duties. So far we have (a) responded to charges of "empty formalism" in Chapter 2 by defending rational nature as the basic value grounding Kantian ethics, and (b) shown that FH and the *Metaphysics of Morals* yield a robust taxonomy of virtues in Chapter 3 that takes us far from the "severe Kant" of the *Groundwork* (Langton, SS 201). Although we had a great deal to say about the affective virtues in Kant's duties of love and the joyous *disposition* of the holy will which the virtuous individual tirelessly seeks to approximate, we are now in a position to address the issues of sociality, character, and the emotions directly. We begin by discussing the role of the emotions, which will lead us to the idea of a kingdom of ends. Having finally arrived the "*complete determination*" of the moral law (G 4:436), we will be able to see what role friendship and character play in Kant's "*kingdom of virtue*" (R 6:95), by responding to specific charges from Bernard Williams and Rae Langton.

I. Desire, Affects, Passions, Delusions

Desire. Kant defines desire [*Begehrung*] as "the self-determination of a subject's power through the representation of something in the future as an effect of this representation" (VA 7:251). In footnotes scattered throughout his other works, Kant describes desire as "a being's *faculty to be by means of its representations the cause of the reality of the objects of these representations*" (KpV 5:9n; KU 5:178n).

We can simplify Kant's definition as follows: Kant calls the "capacity for having pleasure or displeasure in a representation...*feeling*" (MS 6:211). To desire an object, for Kant, is to have a representation of it accompanied by a feeling of pleasure. Here, Kant is concerned only with "*practical pleasure*" which is "necessarily connected with desire for an object" (MS 6:212).⁴² In the case of *empirical* desire, the feeling of pleasure *precedes* the rational determination of the will (to bring about

⁴² Pleasure which is not connected with any desire for the object, but attached only to the representation by itself, Kant calls "contemplative pleasure" or "*inactive delight*" (MS 6:212).

the object of representation). In the case of rational desire, however, the feeling of pleasure is produced in us *after* the rational determination of the will (MS 6:212-3). Therefore, all actions, for Kant, involve both feelings and desires.

As we saw in Chapter 1, § III, if this empirical desire becomes *habitual*, Kant refers to it as “inclination” (MS 6:212; R 6:29; VA 7:251). There are two types of inclinations which threaten to “shut out the sovereignty of reason” (VA 7:251), and hence the ability required by virtue to govern oneself (MS 6:407-8): affects and passions.

Affects. An affect [*Affekte*] is “surprise through sensation, by means of which the mind’s composure (*animus sui compos*) is suspended” (VA 7:252). Therefore affects are “precipitate” or “rash” (MS 6:407). We get a better idea of what Kant means when we turn to his examples in the *Anthropology*: grief, fright, anger and shame are all affects (provided they occur suddenly), but so are exuberant joy, astonishment and laughter, which Kant considers beneficial since they can “cheer us up” (VA 7:254-6). While the majority of affects pose a notable threat to reason (as well as their own ends), due to their *fitful* nature, “the tempest quickly subsides” (MS 6:408). This is not the case with passions.

Passions. A passion [*Leidenschaft*] is “a sensible *desire* that has become a lasting inclination” (MS 6:408). In passions we find a more insidious threat to both prudence and morality since passions are not thoughtless or transitory, like affects, but can coexist with the “calmest reflection” (VA 7:265).

Kant’s edifying contrast between affects and passions borders on the poetic:

Affect works like water that breaks through a dam; passion, like a river that digs itself deeper and deeper into its bed. Affect works on our health like an apoplectic fit; passion, like consumption or emaciation....Affect is like *drunkenness* that one sleeps off; passion is to be regarded as a *dementia* that broods over a representation which nestles itself deeper and deeper (VA 7:252-3)

Kant divides passions into (a) passions of natural inclination (which include *freedom* and *sex*), and (b) passions of human culture. In the social passions arising from culture, we find the

anthropological underpinnings of Kant's idea of unsociable sociability: "the *manias for honor, dominance, and possession*" (VA 7:268). Kant argues that the capacity for "getting other human being's inclinations into one's power" is the same as "possessing others as mere tools of one's will" (VA 7:271). Fuelled by the differences between people, which become exacerbated in the civilized stage of society, and coupled with our proclivity to self-conceit, this hunger for dominance and power takes the form of "*honor, authority, and money*" (VA 7:271). If these inclinations develop into passions, they become *manias* for honor, domination and possession.

The mania for honor is not a love of honor, according to Kant, but merely a "striving after the *reputation of honor*, where semblance suffices" (VA 7:272). As we saw in the last chapter, Kant aligns this mania with the vice of arrogance (a violation of our duty of respect to others) which, in its relentless pursuit of vainglory, surrounds itself with flatterers and yes-men until it brings about its own ruin. This is ironic given the fact that arrogant individuals attempt to exercise influence over others through *opinion*, only to be *hated* for their efforts.

The mania for domination begins as a fear of being dominated by others, but quickly transforms into a desire to dominate them. It is both imprudent, since it antagonizes those it wishes to control, and immoral, since it uses humanity as a means to its own ends (VA 7:273). Because peremptory individuals exercise their dominion through explicit force they are often *feared*.

The mania for possession, which we examined at length in Chapter 3, § VI, "contains a power [to use others merely as a means to one's interests] that people believe satisfactorily replaces the lack of every other power" since money acts as a "universal means of exchange" (VA 7:274). Kant characterizes this mania as "always morally reprehensible," "completely banal" and, once it has set in, the most difficult to rectify (VA 7:274). Because it apotheosizes money in the place of humanity as an end in itself, avarice (a) makes individuals stranger to themselves by alienating them from their humanity, and (b) "swallows up all other vices" making it "incorrigible" (VE 27:402).

Kant notes that, since people gripped by the mania for possession attempt to influence others by accumulating wealth, they are invariably “*despised*” by those less fortunate individuals who continue to gauge their own worth against theirs (VA 7:274). Kant is clear that

the greatest source of happiness or unhappiness, of faring well or ill, of content or discontent, lies in our relationship to other people. For if everyone alike in the town is eating rotting cheese, I eat it too, with satisfaction and a cheerful mind, whereas if everyone else were well-fed, and I alone in sorry circumstances, I would deem it a misfortune. (VE 27:366-7)⁴³

This introduces a psychological concept of relative deprivation which, once again, prefigures Marx’s observation that individuals are more responsive to disparities of wealth rather than their basic standard of living—however “adequate” it may be:

A house may be large or small; as long as the neighbouring houses are likewise small, it satisfies all social requirements for a residence. But let there arise next to the little house a palace, and the little house shrinks into a hut...and however high it may shoot up in the course of civilization, if the neighbouring palace arises in equal or greater measure, the occupant of the relatively little house will always find himself more uncomfortable, more dissatisfied, more cramped within his four walls. (Marx, *Wage-Labour and Capital* 33)

Of course, Kant’s solution is *moral* in nature and, as such, involves overcoming our unsociable sociability by insisting on the absolute and equal worth of all human beings as ends in themselves, while Marx is more concerned with the effects of relative deprivation on the formation of class consciousness. But in terms of practical anthropology, both observations are psychologically astute.⁴⁴

⁴³ See also VE 27:438-9.

⁴⁴ For example, the *Oxford Dictionary of Psychology* identifies two forms of relative deprivation, each corresponding to one of the conceptions outlined above: (1) *egoistic relative deprivation* which “arises from unfavourable comparisons between one’s individual circumstances and those of a comparative reference group” resulting in feelings of “personal dissatisfaction and unhappiness” and (2) *fraternal relative deprivation* which arises from unfavourable comparisons between the circumstances of one’s group (race, class, sex) and a comparative reference group resulting in “protest behaviour, rebellion and...revolution” (Colman, “Relative Deprivation” 648-9).

Delusions. The last passion we will examine plays an important role in Kant's conception of evil. Kant defines the passion of delusion [*Wahn*] as "the inner practical illusion of taking what is subjective in the motivating cause for objective" (VA 7:274). Delusion completes Kant's account of the social passions (despite not being one of the three manias) since the domination and subjection of others is grounded not only "on the fear, interest, or opinion of the dominated, but also on their delusory hopes, when they fancy themselves on the way to achieving a dominant position in relation to someone else" (Wood, *KET* 267). As Wood keenly notes, this explains the "vicious delusion" of the American Dream in which the indigent and dispossessed are willing to tolerate egregious inequalities of wealth in the hope that they will one day occupy the top strata, looking down on their less fortunate neighbors (Wood, *KET* 267).

Following Rousseau, Kant believes that the proclivity to dissimulation and prevarication increases with the advancement of civilization: "On the whole, the more civilized human beings are, the more they are actors" (VA 7:151). This artifice and guile are so entrenched in the civilized condition that "the illusion of affection, of respect for others, of modesty, and of unselfishness" does not deceive anyone at all, since "it is understood by everyone that nothing is meant sincerely by this" (VA 7:151).

There are two important points to note about the passion of delusion. First, its definition mirrors that of self-conceit. Kant defines delusion as "taking the subjective for the objective" (VA 7:275). In the second *Critique*, Kant offers a virtually identical account of self-love as "the propensity to make oneself as having subjective determining grounds of choice into the objective determining ground of the will in general" (*KpV* 5:74). If the former assumes a fixed or unconditional lawgiving function, Kant refers to it as self-conceit. This reinforces the idea we have been arguing for that the true enemy of virtue is not inclination, but our tendency to invert the rational hierarchy of imperatives so that subjective considerations take on objective import. For Kant, this subversion of the moral law is always *delusory*, born in the service of self-love.

Second, while Kant insists that if virtue is habit, it is a “*free*” one, that is, not one that proceeds from “a uniformity in action that has become a necessity through frequent repetition” (MS 6:407)—he nevertheless suggests that our endless pretences at virtue and propriety in civilized society may serve to “deceive the deceiver in ourselves” by becoming ingrained in us as second nature (VA 7:151).

What interests us is that this stratagem is based on an idea of habituation that approximates Aristotle’s: “For when human beings play these roles,” Kant tells us, “eventually the virtues, whose illusion they have merely affected for a considerable length of time, will gradually really be aroused and merge into the disposition” (VA 7:151).⁴⁵ This leads Kant to the conclusion that virtue must be acquired “and the way to acquire it is to enhance the moral *incentive*...by *practicing* virtue” (MS 6:397). Of course, there are marked differences, as well as similarities, between Aristotle’s conception of virtue and Kant’s. We will have more to say about these in Section III.

II. Moral Feelings

After discussing the importance of the emotions to various aspects of morality in his essay “Morality and the Emotions,” Bernard Williams says the following:

It is time, finally, to face up to Kant. For, if one is going to suggest that those things a man does as the expression of certain emotions, can contribute to our view of him as a moral agent...[and] that he should be disposed to certain kinds of emotional response, and not to others, one has to answer the very powerful claim of Kant that this is impossible. (Williams, “Morality and the Emotions” 226)

Is Williams right in his assertion that Kantian ethics denies the emotions a (substantive) role in morality?

The answer to this question lies in two parts. The first concerns Kant’s account of moral feelings, the second, which we will examine in detail in the following section, the role of emotional dispositions in virtuous conduct.

⁴⁵ Consider the following analogous passage from Aristotle’s discussion of habituation in the *Nicomachean Ethics*: “it is through acting as we do in our dealings with human beings that some of us become just and others unjust....for the varieties of these [daily interactions] are reflected in our dispositions” (NE 1103a15-b26).

In the *Metaphysics of Morals*, Kant outlines four “aesthetic preconditions”⁴⁶ [*ästhetische Vorbegriffe*] that lie at the basis of morality as “*subjective* conditions of receptiveness to the concept of duty” (MS 6:399). Like virtue, it cannot be a duty for us to possess them, since it is by virtue of them that we can be put under any obligation at all. Secondly, “consciousness of them is not empirical in origin...[but] follows from consciousness of a moral law as the effect this has on the mind” (MS 6:399). They are: “moral feeling” [*Das moralische Gefühl*], “conscience” [*Gewissen*], “love of human beings” [*Menschenliebe*] and “respect” [*Achtung*] (MS 6:399-403).

This is already an improvement from Kant’s scant footnote on “respect,” the feeling most associated with Kantian ethics in the *Groundwork*. In the same footnote Kant anticipates the objection that will continue to vex both his account of respect and, by extension, the three other moral feelings he outlines in the *Metaphysics of Morals* 12 years later: “I might be accused of using the word *respect* to seek refuge in an obscure feeling, instead of giving distinct information about the matter in question” (G 4:401n). Making matters worse, in the *Groundwork*, Kant seems to conceive of respect *epiphenomenally*, denying it any causal role in the production of moral actions, when he writes, “respect...signifies merely the consciousness of the *subordination* of my will to the law, without mediation of other influences on my senses” (G 4:401n). As we shall see, this conception is short-lived, and undergoes a notable transformation in the *Critique of Practical Reason*.

In light of these problems, our task will be to: (a) give a distinct account of the four aesthetic preconditions, and (b) show how these are not obscure or rarefied feelings which can (c) create an *empirical* incentive to comply with the moral law. Let us begin with the first of these.

Moral Feeling. Kant defines moral feeling as “the susceptibility to feel pleasure or displeasure merely from being aware that our actions are consistent or contrary to the moral law” (MS 6:399). In

⁴⁶ This is Paul Guyer’s translation of Kant’s section title from his essay “Moral Feelings in the *Metaphysics of Morals*” (Guyer, “Moral Feelings” 137). For reasons that will become clear, it is preferable to Mary Gregor’s circuitous translation of the same clause, which reads: “Concepts of What Is Presupposed on the Part of Feeling” (MS 6:399).

Section I, we identified two types of desire which Kant reiterates here in moral terms: a feeling is *pathological* if it precedes the representation of the law, and *moral* if it succeeds it. Kant insists, however, that in *both* cases, “every determination of choice proceeds *from* the representation of a possible action *to* the deed through the feeling of pleasure or displeasure, taking an interest in the action or its effect” (MS 6:399). This is significant for two reasons.

First, it shows that moral feeling is not a “retrospective response” to an action we have already performed, but is concerned, instead, with possible action. And, more importantly, it “makes it indisputable that moral feeling plays a causal role in the etiology of particular actions” (Guyer, “Moral Feelings” 140). Despite not knowing, at this stage, how this causal role comes about, Kant’s maintains that the feeling of pleasure or displeasure acts as an *intermediary* in both cases. We will revisit this idea later on in the section.

Conscience. Conscience, for Kant, is “practical reason holding the human being’s duty before him for his acquittal or condemnation in every case that comes under the law” (MS 6:400). Thus a ruling of conscience is always directed at the individual “to affect moral feeling by its act” (MS 6:400). The analogy to a court of law or judicial proceeding is not accidental since “consciousness of an *internal court* in man (‘before which his thoughts accuse or excuse one another’) is **conscience**” (MS 6:438). Accordingly, conscience is not itself a feeling, but “it causes or stimulates ‘affects’” in the subject that act as triggers to action (Guyer, “Moral Feelings” 143). The idea of conscience as an internal court which compels the subject via moral feeling is captured best by a passage from Kant’s *Lectures on Ethics*:

The inner tribunal of conscience may aptly be compared with an external court of law. Thus we find in us an accuser, who could not exist however, if there were not a law....Now this moral law underlies humanity as a holy or inviolable law. In addition, there is also at the same time in man an advocate, namely self-love, who excuses him and makes many an objection to the accusation, whereupon the accuser seeks in turn to rebut the objections. Lastly we find in ourselves a judge, who either acquits or condemns us. (VE 27:354)

Kant's portrayal of the moral law in advocative conflict with self-love will be important once we turn to the causal role of moral feelings in the production of actions. Finally, Kant argues that the duty to cultivate our conscience involves sharpening the voice of the inner judge and using every means to obtain a hearing, however unfavorable the verdict.

Love of Human Beings. In the previous chapter, we saw that love, as a feeling, cannot be willed (G 4:399; MS 6:401). This meant that whatever affective account we gave of practical love could not be a success account, but had to involve a wide duty to *cultivate* a certain (benevolent) disposition. As such, both practical love and beneficence fell under Kant's duties to others. But there is another *felt* love in Kant's theory that is neither empirical in origin, nor a duty that can be commanded, because if we were not susceptible to it as moral agents, we could not be put under any obligation at all.

This is love of human beings or philanthropic love. Love of human beings is a feeling produced in us directly by reason: it is the love we feel towards others directed at their humanity, or rational nature, as an end in itself. It is important to note, however, that love of human beings is not the product of a rarefied intellect: "*It already lies in human nature, to love something outside oneself, and especially another human being. It would be unnatural coldness in a man to feel no [such] interest...*" (VE 27:682).

There are, of course, many reasons to love others. Our unsociable sociability offers up some popular candidates: we may love others for their celebrity, for their power and clout, or for their wealth. Less cynically, we may love them for their intelligence, beauty, or companionship. But absent all these qualities, when Yeats's yellow-haired girl objects that she wants to be loved for herself alone, and not her yellow hair—"But I can get a hair-dye / And set such colour there, / Brown, or black, or carrot, / That young men in despair / May love me for myself alone / And not my yellow hair"—the poet answers her: "I heard an old religious man / But yesternight declare... / That only God, my dear, / Could love you for yourself alone / And not your yellow hair" (Yeats, "For Anne Gregory" 245). The Kantian need not offer the same reply.

It might be objected that the notion of humanity as an end in itself is an *impartial abstraction* which blinds us to the individuality or uniqueness of the other. In the words of Robin Dillon:

[A]lthough the Kantian formula of persons as ends in themselves is claimed to regard persons as irreplaceable, there is a sense in which Kantian respect does in fact view persons as intersubstitutable, for it is blind to everything about the individual except her rational nature, leaving each of us indistinguishable from the other. Thus, in Kantian-respecting someone, there is a real sense in which we are not paying attention to *her*... (qtd. in Baron, *KEWA* 10, n. 9)

As David Velleman points out, this objection “confuses judgement and appreciation”: it is true that Kant *judges* all persons as having absolute, and hence equal, value by virtue of the rational nature they share, but Kant insists that we must *appreciate* this value as a dignity that sits above any comparison or price (Velleman, “Love as a Moral Emotion” 102-6). Everything, for Kant, “has either a **price**, or a **dignity**” (*G* 4:434). That which has a “*market price*” admits of equivalence (and is therefore intersubstitutable), whereas what has dignity, as we saw in Chapter 2, § I, is “elevated above any price, and hence allows of no equivalence” (*G* 4:434-5). This explicitly bars us from drawing comparisons between persons, let alone replacing one rational nature with another. It is precisely the inviolability of rational nature which guarantees our individuality and protects us from the fungibility of the market.

Respect. In the *Groundwork*, Kant defines respect, or *reverentia*, as “the representation of a worth that infringes on my self-love. Thus it is something that is considered an object neither of inclination, nor of fear, even though it is at the same time somewhat analogous to both” (*G* 4:401n). Respect is the mixture of humility and awe we feel in response to the moral law. Earlier, we saw that in the *Groundwork*, Kant seems to offer an epiphenomenal account of respect when he describes the feeling as merely the effect, or by-product, of the subordination of the will to the moral law (*G* 4:401n). In the second *Critique*, this conception has already begun to change:

[T]he representation of the 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 superiority of its objective law to the

impulses of sensibility...(with regard to the will affected by impulses) is produced in the judgement of reason through the removal of the counterweight. And so respect for the law is not the incentive to morality; instead it is morality itself subjectively considered as an incentive inasmuch as pure practical reason, by rejecting all the claims of self-love in opposition with its own, supplies authority to the law. (KpV 5:75-6)

There are two focal points in Kant's passage: (1) respect is now produced, not by the subordination of the will to the moral law, but by "the mere *representation* of the moral law," and (2) the causal role played by respect (and moral feelings in general) is one of "the *reweighting of empirical feelings*" by removing the hindering influence of self-conceit to the moral law (Guyer, "Moral Feelings" 135). The first point situates respect *in between* the representation of the moral law and the subordination of the will to the law. This is where we should expect moral feelings to be if they are to be causally efficacious. The second asserts that, in the inner tribunal of conscience where the moral law and self-love make their respective claims, moral feelings "reweigh" the odds *empirically* in favour of the former. But how is this possible, if moral feelings do not precede the representation of the law, but succeed it? Guyer adduces the following passage in way of explanation:

...whatever diminishes the hindrances to an activity is a furthering of this activity itself. Recognition of the moral law, however...fails to express its effect in actions only because subjective (pathological) causes hinder it. Therefore respect for the moral law must be regarded as also a positive though indirect effect of the moral law on feeling insofar as the law weakens the hindering influence of the inclinations by humiliating self-conceit, and must therefore be regarded as the subjective ground of activity—that is, as the incentive to compliance with the law. (KpV 5:79)

According to Kant, the etiology of any action (of possible moral import) will involve both the claims of self-love, driven by the inclinations, and those of the moral law, whose representation gives rise to moral feelings. Given what we know about affects, passions, and delusions, especially in light of

our unsociable sociability, Kant believes that these inclinations will often be in the service of self-love. This acts as an impediment to the moral law.

Moral feelings subsequently play a dual role: they add “the expectation of pain to our otherwise pleasurable representation of acting in some way contrary to morality and pleasure to our otherwise painful representation of forgoing that course of action for the sake of morality” (Guyer, “Moral Feelings” 135). In other words, moral feelings act not only as pleasurable counterweights to the inclinations (when these oppose duty), but they also “*weaken*” or “*strike down* self-conceit” by humiliating it before the moral law (*KpV* 5:73). This empirical effect notwithstanding, the fact that these counterweights are rational desires which have their *origin* in reason has led to the suspicion that they are obscure or rarefied feelings, befitting only a Kantian scholar or Stoic sage.

There are three ways to respond to this charge. The first is to highlight Kant’s repeated statements that moral feelings “lie at the basis of morality, as *subjective* conditions of receptiveness to the concept of duty” (*MS* 6:399). In other words, “Kant takes it to be obvious that a human being without feelings and emotions....simply could not be a rational moral agent” (Wood, *KE* 35). This is why, for Kant, it cannot be duty to have moral feelings—rather “every human being has them, and it is by virtue of them that he can be put under obligation” (*MS* 6:399). Kant reiterates this idea in his individual treatment of the four feelings: “there can be no duty to have moral feeling or acquire it; instead every human...has it in him originally” (*MS* 6:399-403). Similarly in his discussion of conscience, Kant writes:

Many have contended that conscience is a product of art and education...but if this were so, the person having no such training and education of his conscience could escape the pangs of it, which is not in fact the case....[T]hough our understanding may be cultivated, conscience does not need to be. It is therefore nothing but a natural conscience. (*VE* 27:355-6)

In other words, if we *all* have these feelings in us “originally,” then they cannot be things we acquire, much less the rarefied or technical products of a moral intellect. This is why Kant refers to the four

aesthetic preconditions as “moral endowments” and “natural predispositions of the mind” (*MS* 6:399), which takes us to the second point.

While Kant argues that moral feelings have their *origin* in reason, their effects play out *empirically* in time, whether they are acting as pleasurable counterweights to opposing inclinations or as phenomenal predispositions to moral action alongside the former. After analysing Kant’s language in his discussion of the moral feelings, Guyer concludes that

by putting the various clues together, it seems clear that what Kant intends to discuss are phenomenal feelings or effects on feeling—“*ästhetisch*” [aesthetic] and both “*natürlich*” [natural] and “*Gemüts*” [mind]—make that clear, especially since Kant uses the last term only to refer to the phenomenal mind or empirical consciousness... (Guyer, “Moral Feelings” 137)

Guyer is wrong to regard moral feelings as “entirely empirical” (Guyer, “Moral Feelings” 139). He is right, however, in noting Kant’s language vis-à-vis their *role* in the production of moral actions: the four aesthetic preconditions are *rational* feelings which act on our sensibility. What Kant’s language demonstrates is that he views this role as “aesthetic” or “natural”—not the magical product of a “rampant platonism”⁴⁷ (to borrow a provocative term from McDowell). This is simply a reflection of Kant’s belief that all actions, including moral ones, involve feelings and desires (*MS* 6:211-3), but “in an action done from duty the feelings and desires are *rational* (not empirical)” (Wood, *KE* 36). Here, we have finally arrived at the heart of the doubt plaguing Kant’s account: scepticism about rational feelings.

Kant’s empiricist critics hold that all feelings and desires are empirical in origin. Because their moral psychologies cannot countenance non-empirical feelings,⁴⁸ Kant’s account is immediately suspect: rational feelings appear *sui generis*, rarefied or occult. This point of contention is not problematic in

⁴⁷ “In rampant platonism, the rational structure within which meaning comes into view is independent of anything merely human, so that the capacity of our minds to resonate it looks occult or magical” (McDowell, *Mind and World* 92).

⁴⁸ Consider, for example, Williams: “...Kant’s work is in this respect a shattering failure, and the transcendental psychology to which it leads is, where not unintelligible, certainly false. No human characteristic which is relevant to degrees of moral esteem can escape being an empirical characteristic, subject to empirical conditions, psychological history and individual variation...” (Williams, “Morality and the Emotions” 228).

itself (as a longstanding philosophical disagreement). The problem arises when Kant's critics attempt to *read him* through an empiricist lens.

The most glaring example of this concerns Kant's denial that actions from duty are done from inclination (G 4:398). As we saw in Chapter 1, § III, his critics translate "inclination" (which Kant defines as habitual *empirical* desire) into "desire" *simpliciter*—and conclude on that basis that actions from duty, for Kant, involve no desire at all. According to Wood,

this seemingly innocent translation (of 'inclination' as 'desire') actually involves a direct begging of the question against the most fundamental tenet of Kant's entire moral psychology. When Kant's assertions are subjected to it, this guarantees from the start that his position will appear repugnant, if not downright nonsensical—which is exactly how Kant himself would regard the "Kantian" position when so translated. (Wood, *KE* 282, n. 6)

Only *rational* desires, according to Kant, can account for the moral disposition in human beings, since they guard against both (a) "the empiricism of pure practical reason" (which reduces the consciousness of virtue to the domain of self-love, i.e., happiness or pleasure) and, more significantly, (b) "the mysticism of pure practical reason," which Kant *rejects* on the grounds that it projects the moral disposition into an "invisible kingdom of God" which "strays into the transcendent" (*KpV* 5:70-1). Nevertheless, in a direct swipe at his empiricist critics, Kant asserts that

it is more important and advisable to guard against *empiricism* [than mysticism], since *mysticism* is still compatible with the purity and sublimity of the moral law...empiricism, on the contrary, destroys at its roots the morality of dispositions...and substitutes for it something quite different, namely in place of duty an empirical interest. (*KpV* 5:71)

Kant's reasons for guarding morality against the claims of self-love, "with which the inclinations generally are secretly leagued" (*KpV* 5:71), will become clearer when once we examine the nature of self-conceit in Section IV.

Actions from duty involve rational desires—moral feeling, conscience, love of human beings and respect—which operate as “effect[s] on feeling and hence on the sensibility of rational being[s]” (KpV 5:76), and without which we could not be moral agents. But Kantian ethics also has a great deal to say about the role of *empirical* desires in virtuous conduct. We turn to these next.

III. Two Conceptions of Virtue

In Chapter 3, § III, we saw that virtue, for Kant, involves (a) the moral strength to carry out one’s duty as a commitment to a rational end, and (b) an asymptotic progression towards holiness or moral perfection. Given this definition, even if we have successfully shown that Kantian ethics has a robust taxonomy of virtues, we are met with another objection centering on the *nature* of Kantian virtue. The charge, commonly voiced by virtue ethicists, takes issue with Kant’s conception of virtue as moral *strength*—a characterization which seems to align it with Aristotelian “continence” or “self-control” [*enkrateia*].

According to Aristotle, although self-control resembles virtue [*aretē*], it is inferior to the latter because in self-control the agent is *distressed* at acting as she should (NE 1145a15-1152a36). Since Aristotle maintains that “the pleasure or pain that supervene on what people do should be treated as a sign of their disposition...someone who holds back from bodily pleasure and does so cheerfully is a moderate person, while someone who is upset at doing so is self-indulgent” (NE 1104b4-8).

In other words, in virtuous conduct there exists a harmony between what an agent thinks she should do and does, and her inclinations—whereas in self-control, while the agent holds the same beliefs regarding what she should do and successfully acts on those beliefs, she does so against *conflicting* inclinations. This is why, despite carrying out the same action, in the first case the agent acts “cheerfully” while in the second she is grieved or “upset.”

Now we can formulate the objection as follows: if Kantian virtue corresponds to Aristotelian self-control, then it is, at best, an inferior species of virtue (or, at worse, not virtue at all). In *After Virtue*,

Alasdair MacIntyre writes, “[v]irtues are dispositions not only to act in particular ways, but also to feel in particular ways. To act virtuously is not, as Kant was later to think, to act against inclination; it is to act from inclination formed by the cultivation of the virtues” (MacIntyre, *After Virtue* 149).

The problem with this line of criticism is that it mischaracterizes both Aristotle’s position and Kant’s. In the case of the former, Wood and Korsgaard each dispute the notion that Aristotelian virtue involves acting “from inclination” (Wood, *KET* 329-33; Korsgaard, “From Duty” 213-27). If by “inclination formed by the cultivation of virtues” MacIntyre means rational desires stemming from the relevant virtues, then Aristotle and Kant are in agreement. But, since MacIntyre draws a stark contrast with Kant, this cannot be his intended meaning. Rather by acting “from inclination” he must mean having one’s inclinations so constituted that they always issue in the right conduct (Wood, *KE* 145-6).

Aristotelian Virtue

Korsgaard points to three fundamental aspects of Aristotelian virtue in the *Nicomachean Ethics*⁴⁹ that cast doubt on this characterization: (a) the possibility of “self-control” [*enkrateia*], (b) the role of “decision” [*prohairesis*] and (c) the idea that virtuous action is done “for the sake of the fine” [*tou kalou heneka*] (Korsgaard, “From Duty” 213).

Self-Control. The first of these we have already touched on. In the case of self-control, we act successfully, and as we think we should, but we do so *against* our inclinations or empirical desires which are pulling us in the opposite direction. Korsgaard concludes: “the bare possibility of continence...shows that Aristotle thinks that human agents have the power to step back from inclinations and decide whether to act on them or not” (Korsgaard, “From Duty” 213). Korsgaard’s aim (in all three cases) is to show that this rational process is also at play in Aristotelian virtue.

⁴⁹ For a similar argument that draws parallels between Aristotle’s idea of acting “for the sake of the fine” and Kant’s notion of acting “from duty” in the *Eudemian Ethics*, see Jennifer Whiting, “Self-Love and Authoritative Virtue: Prolegomenon to a Kantian Reading of *Eudemian Ethics* viii 3,” *Aristotle, Kant and the Stoics: Rethinking Happiness and Duty*, eds. Stephen Engstrom and Jennifer Whiting (Cambridge: Cambridge University Press, 1996): 162-99.

Of course, given what we have said so far, it could still be the case that Aristotle thinks this power or function (of reason) is reserved only for the self-controlled person and not the virtuous one. This suggests that the *difference* between self-control and virtuous action lies in the exercise of reason, but Aristotle makes exactly the opposite claim: “in the self-controlled person it [the non-rational part of the soul] is obedient to reason—and in the moderate and courageous person it is...*still readier* to listen; for in him it always chimes with reason” (NE 1102b26-9, emphasis added). Aristotle goes as far as to say that in the virtuous person the non-rational part of the soul “listens” to and “obeys” the rational element as one does a father (NE 1102b31-1103a4). It is evident from these comments that both self-control and virtuous conduct involve reason. This becomes even clearer in Aristotle’s account of decision.

Decision. Decision, for Aristotle, is a practical “all-things-considered judgement of what to do” (Broadie, Philosophical Introduction 42). There are five important points to note about decision in Aristotle’s discussion (NE 1111b4-1112a17):

- (a) Decision is not found in non-rational animals, children or things done “on the spur of the moment.” As such it is a subclass of voluntary action.
- (b) The self-controlled person acts from decision, not appetite.
- (c) Decision is not concerned with what is painful or pleasant.
- (d) Decision involves a practical immediacy or “ought.”
- (e) Decision is arrived at through deliberation.

We have already seen that, for Aristotle, the self-controlled person acts from decision (NE 1111b15), but what about the virtuous agent? Aristotle not only believes that decision is “highly germane to excellence [*aretē*, or virtue]” but also that it is better than actions for discerning virtuous character (NE 1111b5-6), since it encapsulates the reasoning and deliberation behind any given action.

Despite Aristotle’s claim that “we deliberate, not about ends, but about what forwards those ends” (NE 1112b12-3), he defines virtue as “a disposition issuing in decisions...being determined by

rational prescription in the way in which the wise person would determine it” (NE 1106b36-1107a2).

Thus rational deliberation, for Aristotle, cannot be understood only instrumentally, as something falling outside the province of virtue—or so we have been arguing—since all the virtues “are kinds of decision, or anyway involve decision” (NE 1106a4). If this account is correct, not only is reason active in Aristotelian virtue, or excellence, but it is *constitutive* of it. Aristotle puts the point as follows:

everybody when defining excellence describes the disposition and what it relates to, and then adds ‘according to the correct prescription’....But one must go a little further than this: it is not just the disposition according to the correct prescription, but the disposition *accompanied by* the correct prescription, that constitutes excellence. (NE 1144b22-8)

In his translation, Christopher Rowe uses “prescription” in place of “reason.” The Terence Irwin translation (Hackett, 1999) of the same passage is clearer on this front: “For it is not merely the state in accord with correct reason, but the state involving correct reason, that is virtue” (NE 1144b27-8).

‘*For the Sake of the Fine.*’ In the *Groundwork*, Kant maintains that “in the case of what is morally good it is not enough that it *conform* with the moral law, but it must also be done *for its sake*” (G 4:390). This establishes the basis for his distinction between actions done “*in conformity with duty*” and those done “*from duty*” (G 4:390, 398; KpV 5:81). While the former are “honorable” and deserving of “praise and encouragement,” only the latter, according to Kant, are worthy of proper moral esteem (G 4:398). We find a similar idea in Aristotle’s contention that “actions in accordance with excellence are fine and for the sake of the fine” (NE 1120a23-4).

The best illustration of this is found in Aristotle’s various comments on the virtue of courage, which he employs readily in his examples:

“...the courageous person is as unshakeable as a human being can be. So he will be afraid of those sorts of things [i.e., fearsome things] too, but he will withstand them in the way one should, and following the correct prescription, for the sake of achieving what is fine” (NE 1115b11-14).

“...one should be courageous because it is fine to be so” (NE 1116b3).

“...while death and wounds will be painful to the courageous person...he will withstand them because doing so is fine or because not doing so is shameful” (NE 1117b7-10).

The example of the courageous person who risks his life or suffers grievous wounds for the sake of the fine is especially revealing since it shows that “not all the excellences give rise to pleasant activity, except to the extent that the pleasant activity *touches on the end*” (NE 1117b16-7, emphasis added). Consequently, the value of this end cannot be determined solely by the inclinations. The capacity to choose to act for the sake of the fine is, according to Korsgaard, “the capacity to make a reflective judgement about the value of the action as such and to be moved by that judgment to perform or avoid the action” (Korsgaard, “From Duty” 217).

The parallel between acting “from duty” and acting “for the sake of the fine” gains further credence from Aristotle’s distinction in the *Eudemian Ethics* between the *agathos*, who acts virtuously for the sake of natural goods, and the *kaloskagathos*, who does so because it is both fine-and-good (Whiting, “Self-Love” 165-6). Whiting uses Aristotle’s criticism of the Spartans as an illustration of this point. The Spartans believe they ought to possess virtue, say courage, for the sake of an external good (e.g. military victory). Thus, the Spartans’ conduct only coincides with the fine—they are *agothoi*, since *kalokagathia* requires choosing fine things for their own sake (Whiting, “Self-Love” 166-7). Therefore, for both Aristotle and Kant, it is not enough that our inclinations coincide with, or incline us towards, the right action. Virtuous action is that which is done for its own sake on rational grounds.

Kantian Virtue

So far we have been arguing that reason plays a significant role in Aristotelian virtue. Yet MacIntyre is surely right when he says that virtue, for Aristotle, (also) involves pleasure and pain, specifically in the form of inclinations harmonizing with the demands of morality (NE 1104b4-1105a16). Is this desideratum missing in Kantian virtue?

First, it is not correct to say that Kantian virtue always involves acting “against inclination.” Kant believes that virtue “resist[s] the inclinations *when they invite transgression*” (R 6:57n, emphasis added). Virtue is an endless struggle for Kant because (a) the purity of the moral content of its deeds remains opaque to the agent⁵⁰ and (b) virtue is corruptible, which is to say, it “can never settle down in peace and quiet with its maxims adopted once and for all...[because] if it is not rising, it is unavoidably sinking” (MS 6:409). As we saw in Chapter 3, § IV, Kant thinks we have a duty to cultivate certain inclinations that aid in our capacity to fulfill our duty (G 4:399; MS 6:388; “The End of All Things” 8:338). However, this instrumental conception alone will not do since Aristotle conceives of the virtues not only in terms of “the state involving correct reason,” but also as *dispositions* to feeling (NE 1105b19-1106a13). There are two points to note here.

Regarding the dispositions themselves, Kant’s duties of love—specifically practical love, gratitude and active sympathetic participation—all involve an affective component that is constitutive of the virtue itself. In terms of practical love, this is a direct duty to love and benefit our fellow human beings *gladly*, from a *joyous disposition* (KpV 5:83). Affective gratitude enjoins us to cultivate a “grateful disposition” and a warm *appreciativeness* in response to the “mere heartfelt benevolence” of others (MS 6:455-6). Finally, anticipating Kant’s social kingdom of ends, active sympathetic participation involves the reciprocal enjoyment and sharing in the joys and sufferings of others in the form of a “*communio sentiendi liberalis*” [free communion of sentiment] (MS 6:456). For Kant, the loathsome vices opposed to these virtues are also dispositions, albeit “malevolent” ones, which he refers to as the “vices of hatred for human beings” (MS 6:458-61).

⁵⁰ “For a human being cannot see into the depths of his own heart so as to be quite certain, in even a *single* action, of the purity of his moral intention and the sincerity of his disposition....In the case of any deed it remains hidden from the agent himself how much pure moral content there has been in his disposition” (MS 6:393-4).

Nevertheless, the crux of the challenge to Kantian virtue seems to lie in Aristotle's contention that "the pleasure and pain that supervenes on what people do should be treated as a sign of their dispositions" in *all* the virtues, not just a handful of duties (of love, for example) (*NE* 1104b4-5).

Knowing what we do about the disposition of the holy will, which Kantian virtue tirelessly seeks to approximate, the convergence of Aristotle and Kant on this point should come as no surprise. Following Aristotle and the ancient moral philosophers, Kant insists that only "a heart joyous in the compliance with its duty...is the sign of genuineness in virtuous disposition" (*R* 6:23n). And, without this joyous disposition, or cheerful heart, virtue lapses into ascetical fear, dejection and servility—a frame of mind Kant disparages as "Carthusian" (*R* 6:23n). Here, the textual evidence is unequivocal:

The rules of practicing virtue...aim at a frame of mind that is both *valiant* and *cheerful* in fulfilling its duties (*animus strenuous et hilaris*)....[for] what is not done with pleasure but merely as compulsory service has no inner worth for one who attends to his duty in this way and...is not loved by him. (*MS* 6:484)

~

...*health* is only a negative kind of well-being: it cannot itself be felt. Something must be added to it, something which, though it is only moral, affords an agreeable enjoyment to life. This is the ever-cheerful heart, according to the idea of the virtuous *Epicurus*. (*MS* 6:485)

~

Now, if we ask, "What is the *aesthetic* constitution, the *temperament* so to speak of *virtue*: is it courageous and hence *joyous*, or weighed down by fear and dejected?" an answer is hardly necessary.

The latter slavish frame of mind can never be found without a hidden *hatred* for the law... (*R* 6:23n)

From these passages, we can safely conclude that Kantian virtue cannot be equated with Aristotelian self-control because, in the latter, the agent is grieved or upset at fulfilling her duty. Nor can we say that Aristotelian virtue stems "from the inclinations" if this is understood to exclude reason. If, on the other hand, MacIntyre intends the qualifier "by the cultivation of the virtues" to include rational desires, then Aristotle and Kant are, once again, in agreement.

IV. Are the Inclinations the Enemy of Virtue?

We saw that Kant defines virtue as “the moral disposition *in conflict*” (*KpV* 5:84). The question then becomes: in conflict with what? Traditionally, commentators, both sympathetic to and critical of Kant, have taken his denouncement of the inclinations as a sign that virtue is a struggle against the former.⁵¹

In the *Religion*, however, Kant explicitly denies that the inclinations are the enemy of virtue: “*Considered in themselves* natural inclinations are *good*, i.e. not reprehensible and to want to extirpate them would not only be futile but harmful and blameworthy as well” (*R* 6:58). In other words, not only is attempting to rid ourselves of the inclinations impossible, but “depriving oneself (fanatically) of the pleasures of life by exaggerated *discipline* of one’s natural inclinations....is opposed to a human being’s duty to himself” (*MS* 6:452). Kant considers all such attempts to be servile and monkish, and hence a violation of humanity in our person.

More strikingly, Kant maintains that the ground of evil “cannot...be placed, as is commonly done, in the sensuous nature of the human being, and in the natural inclinations originating from it” since these “*give the occasion for what the moral disposition can demonstrate in its power, for virtue*” (*R* 6:34-5, emphasis added). According to Kant a “human being...is evil only because he reverses the moral order of incentives in incorporating them into his maxim....he makes the incentives of self-love and their inclinations the condition of compliance with the moral law” (*R* 6:36). Kant refers to this as the “radical propensity to evil” (*R* 6:29-32), or “self-conceit” (*KpV* 5:73-4).

Jeanine Grenberg argues that this reading is superior to the standard account, which locates evil in the inclinations, because “inclinations considered as natural impulses have no standing from the practical point of view, the only point of view from which a discussion of virtue can make sense” (Grenberg, “What is the Enemy of Virtue?” 153). In other words, the enemy of virtue must be one we

⁵¹ For passages in Kant supporting this reading see *G* 4:400, 428, 454; *KpV* 5:80, 86, 117-8, 147; *MS* 6:380, 384, 394. A review of the two positions which concludes in favour of the one advanced in this section can be found in Jeanine Grenberg, “What is the Enemy of Virtue?” *Kant’s Metaphysics of Morals: A Critical Guide*. Ed. Lara Denis. (New York: Cambridge University Press, 2010): 152-69.

can combat and be held responsible for, and thus must fall within the sphere of freedom, but “*sensuous nature...contains too little to provide a ground of moral evil in the human being, for, to the extent it eliminates the incentives originating in freedom, it makes the human a purely animal being*” (R 6:35).

This has two important consequences: (1) the enemy of virtue resides in reason and not the inclinations, and (2) this propensity to evil, which is another name for self-conceit, is *socially* conditioned. The first point puts Kant at variance with the Stoics: “those valiant men [the Stoics] mistook their enemy, who is not to be sought in the natural inclinations...but is rather as it were an invisible enemy, one who hides behind reason” (R 6:57).

The Stoics, according to Kant, *mistook* the natural inclinations as the enemy of virtue, and thus wished to be entirely free of them—a position that Kant’s critics today ironically attribute to him. Against the Stoics, Kant locates the “real opponent of goodness” in the corruption of one’s maxims so that a human being must “extricate himself from an evil which is to be sought not in his inclinations but in his perverted maxims” (R 6:57n). The threat to virtue, then, is not the inclinations, *nor* the incorporation of the inclinations into one’s maxims—since Kant believes we have a rational obligation to seek our happiness whenever it does not conflict with duty (G 4:415-9; MS 6:388; VE 27:395)—but the perversion of our maxims so that the moral law is *subordinated* to self-love. Yet if evil or self-conceit lies in reason, which always involves the *free* adoption of maxims, how is it socially conditioned?

Kant’s explanation of this second, Rousseauian point is quite revealing. He begins by arguing that self-love always “*involves comparison* (for which reason is required)” (R 6:27). As we saw earlier in our discussion of relative deprivation, Kant believes that “only in comparison with others does one judge oneself happy or unhappy” (R 6:27). From this comparison, we acquire “the inclination *to gain worth in the opinion of others*” and, eventually, “the unjust desire to acquire superiority for oneself over others” (R 6:27). This constitutes the negative dimension of our unsociable sociability, which manifests itself in

the three manias for honor, domination and possession and gives rise to jealousy, rivalry and inequality in the civilized stage of society.

We are now in a position to understand what Kant means when he says that these vices are “grafted” on the inclinations as “vices of *culture*” (R 6:27).

Like all inclinations, these “acquired” passions of cultures are “inclination[s] whose direction [must be]...approved by reason” (VA 7:267).⁵² Self-conceit is what grants them this (specious) approval by regarding self-love as *legislative* (KpV 5:74). For Kant, this endorsement always involves inverting the rational order of incentives so that our comparative self-worth takes precedence over the absolute worth of humanity as an end in itself.

In a revealing passage inspired by Rousseau that could have been taken directly from his *Discourse on the Origin of Inequality*, Kant concludes:

It is not the instigation of nature that arouses what should properly be called the *passions*, which wreak such great devastation in his [a human being’s] originally good disposition. His needs are but limited, and his state of mind in providing for them moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and despise him for it. Envy, addiction to power, avarice, and the malignant inclinations associated with these assail his nature, which on its own is undemanding, *as soon as he is among human beings*. (R 6:93-4)

So far we have been arguing that reason, not inclination, is the enemy of virtue, and that this enemy takes the form of a socially conditioned self-conceit. Although these points are significant in themselves, what interests us above all is Kant’s *solution* to the problem of moral evil.

If the provenance of self-conceit is social, then it requires a social remedy. In Kant’s words, “inasmuch as we can see...the dominion of the good principle is not otherwise attainable, so far as human beings can work toward it, than through the setting up and diffusion of a society in accordance

⁵² “As Kant sees it, human beings are subject to certain incentives—impulses which present themselves as candidates, so to speak, to be reasons for actions. Among these are our desires and inclinations, as well as respect for the moral law” (Korsgaard, CKE 202).

with, and for the sake of, the laws of virtue" (R 6:94). This establishes two points which orient Kantian ethics in the opposite direction of the familiar individualistic and libertarian readings of his philosophy: (1) it is futile to strive after virtue in isolation, and (2) the highest ideal, as well as clearest illustration, of Kant's ethics is to be found in a free "*ethical community*" (R 6:94-5). Because this community involves a systematic and harmonious unity of all human virtues and ends, Kant refers to it as the kingdom of ends.

V. The Kingdom of Ends

Kant defines "kingdom" [*Reich*] as "the systematic union of several rational beings through common laws" (G 4:433). A kingdom of ends, is one in which these laws "have as their purpose...the reference of these beings to one another, as ends and means" (G 4:433). We can identify three distinct features of the Kantian kingdom of ends: (a) it is *systematic* because it entails the "progressive organization of citizens of the earth into and towards the species as a system that is cosmopolitically united" (VA 7:333); (b) its ends are combined into a *harmonious* and mutually supporting system and are, as a result, (c) *shared* ends pursued by all in common, without conflict or competition.

The importance of the kingdom of ends in Kant's ethics becomes clear when we turn to its corresponding formula in the *Groundwork*:

FKE: "Act according to the maxims of a member universally legislating for a merely possible kingdom of ends" (G 4:439).

In the *Groundwork*, Kant intends to chart the progression of the principle of morality scholastically from its *form*, to its *matter*, to the *complete determination* of the moral law (G 4:436-7). FUL represents the form of the law (see Ch. 1, § VI). FH represents the matter, or fundamental value, underpinning Kant's ethical system (see Ch. 2, § I).

When Kant says FKE represents the "*complete determination*" of the moral law, he has in mind a formula that incorporates both FUL and FH into a "*unity*" or "*totality*" of ends (G 4:436). FKE combines (a) the form of FUL—"since the ground of all practical legislation lies *objectively* in...the form of

universality”—with (b) the matter of FH, which tells us that the “subject of all ends...is every rational being as an end in itself” (G 4:431). Kant clearly believes that FKE “follows indisputably” from the previous two formulas (G 4:438), but his reasoning is far from clear.

We can reconstruct Kant’s argument as follows: the form of law is *universal* legislation (FUL). But the ground of this legislation is that rational beings should be treated as ends in themselves (FH). Treating rational beings as ends in themselves entails bringing our own ends in agreement with theirs (G 4:430; MS 6:395). Therefore, the laws of a possible kingdom of ends are those that combine the idea of rational beings as ends in themselves, as well as the maxims chosen by them, into a systematic unity of ends.

This has far-reaching implications: if we adhered to FH *universally*, both in our own person and the person of any other, competition, enmity and social antagonism would disappear—to be replaced by mutual cooperation, amity, and the pursuit of common ends. Here, Wood’s elegant description of the ramifications of FKE (FRE in his text, since Wood translates *Reich* into “Realm”) is worth quoting in full:

FRE commands us to avoid all patterns of end setting that involve fundamentally competitive relations between ourselves and other rational beings. It forbids us to relate to others in any way that involves the frustration of any person’s deepest ends. Conflict or competition between human ends is compatible with FRE only if it is in the service of a deeper systematic unity among all human ends...in which no member of the realm of ends is left out. (A later formulation of the same idea was: “An association in which the free development of each is the condition for the free development of all.”) (Wood, KE 79)

Given this result, there are two points worth highlighting: (1) the Kantian kingdom of ends is not merely the “liberal ideal” of “a good society” (Hare, *Freedom and Reason*, qtd. in Wood, KET 166), but something far more radical since it requires an organic unanimity of ends, which are to be pursued in common and (2) the fact that the end result is one in which rivalry and conflict disappear confirms our earlier contention that FKE is the answer to self-conceit, which is founded on these social antagonisms. Further proof of this is Kant’s assertion that friendship, which he regards as the clearest model of FKE

available to us, is the process by which “self-love is swallowed up in the idea of generous mutual love” (VE 27:423).

Friendship

We saw in Section I, that even when we enter into society, our unsociable sociability guarantees that “we do not enter completely into society”: “we do not...pour out our feelings, attitudes and judgement; “we are all under constraint”; we “harbour a mistrust of others”, “cover up our weaknesses” and “withhold our opinions” lest others think poorly of us, or use them against us (VE 27:427). On the other hand, Kant argues, “the human being is meant for society (though he is also an unsociable one), and...feels strongly the need to *reveal* himself to others (even with no ulterior purpose)” (MS 6:471).⁵³

Friendship, according to Kant, is what fulfills this need and allows human beings to open their hearts to others and commune with them—for “this is the whole purpose of man, which allows him to enjoy his existence” (VE 27:427). Following Aristotle,⁵⁴ Kant identifies three types of friendship and regards only one of them as the true Idea of friendship. Tellingly, Kant refers to the latter as the “friendship of disposition and sentiment” (VE 27:246-7).

The *friendship of need* is the most rudimentary form of friendship, and is found under primitive conditions in which individuals must bond together to meet some basic need (e.g. hunting in groups). Once this need is met, however, the friendship is immediately dissolved.

The *friendship of taste* is the most common of three types, and “consists in taking pleasure in the company and mutual association of the two parties, rather than their happiness” (VE 27:426). Because friendships of taste are not predicated on a mutual benevolence, Kant regards them at best as “analogue[s] of friendship” (VE 27:426).

⁵³ See also VE 27:427: “We...have a strong impulse to unbosom ourselves and be wholly companionate.”

⁵⁴ NE 1156a6- 1157b6.

The *friendship of disposition and sentiment*, however, is the one which Kant thinks best illustrates the promise of FKE and constitutes the true Idea of friendship. It involves five components, which Kant outlines as follows:

- (a) ***Reciprocal Well-Wishing (or Benevolence)***—“a bilateral bond” in which “the other also views my best interests as his own” (VE 27:675-6);
- (b) ***Equality***—Kant believes that true friendship can only exist between equals since beneficence lapses into “favour” among unequals, prompting him to conclude that “*inter superiores et inferiores* no friendship can occur” (VE 27:676);
- (c) ***Reciprocal Possession***—“the communal possession of one person by another” so that “one belongs to the other...each mutually shar[ing] in every situation of the other, as if it were encountered by himself” (VE 27:677);
- (d) ***Reciprocal Enjoyment of Humanity***—in the form of the “mutual disclosure of thoughts” which is the ground for the communication of feeling necessary for friendship and the “foundation of openheartedness, *animus apertus sinceritas aperta* [a candid and sincere disposition]” (VE 27:677);
- (e) ***Love for Mutual Well-Liking***—the idea that true friendship also requires cherishing the moral value of friendship in addition to its various components, and the social enjoyment they provide (VE 27:680).

In her essay, “Creating the Kingdom of Ends” Korsgaard acknowledges the importance of friendship to Kant’s vision, but focuses almost exclusively on the relationship between reciprocity and *responsibility* (Korsgaard, CKE 188-221), which lends her discussion a dry, academic air. Kant, however, has little to say about responsibility in this regard.⁵⁵ The five components of friendship, which constitute “the moral relation in perfected form” (Korsgaard, CKE 191), give us an intimation of what the kingdom

⁵⁵ Korsgaard admits this much: “Kant is not very interested in praise and blame, and seldom mentions them. And when he does discuss issues of moral assessment, much of what he says favors taking a generous attitude” (Korsgaard, CKE 189).

of ends might look like: a free ethical community of equals, in communal possession of each other so that what befalls one befalls all, their projects and ends mutually shared, each supplying what the other lacks—openhearted, benevolent, joyful, participating in the joys and sufferings of others in the form of a free communion of sentiment.

It might still be objected that, although Kant has a rich account of friendship, it lacks the significant moral dimension which leads Aristotle to describe friendship at the beginning of Book VIII as “a kind of excellence...necessary for living” (*NE* 1155a4-5). In other words, Kant’s account of friendship might be purely anthropological.

Kant, however, regards friendship both as the disposition which (a) makes us worthy of happiness and (b) prefigures the kingdom of ends since it is an “overcoming [of] the constraint that we harbour, from mistrust, towards those we associate with, and in opening up to them without reserve” (*VE* 27:428). In light of these two factors, Kant concludes that “friendship (as a maximum of good disposition toward each other) is a duty set by reason, and no ordinary duty but an honorable one” (*MS* 6:469). Far from the stark, theoretical individualism of *FUL*, *FKE* enjoins us to enter into personal relationships with others which approximate the kingdom of ends, wherein self-love is “swallowed up” by a “generous mutual love” in the service of collective ends (*VE* 27:423). This is an ethical duty for Kant, even if the natural ends of friendship are not moral ends.

Ground Projects

The notion of character, according to Williams, involves “having projects and categorical desires with which that person is identified” (Williams, *ML* 14). Categorical desires are desires that constitute the basis, or reason, for caring about our existence (Williams, *ML* 11). These, in turn, furnish us with ground projects without which one’s life would lack meaning, since they provide “the motive force which propels him into the future, and gives him a reason to live” (Williams, *ML* 13).

In Chapter 3, § VI, we argued that Williams was wrong to claim that an allegiance to an impartial Kantian morality is predicated on an “omission of character” (Williams, *ML* 14), which deprives life of substance by failing to make room for ground projects. Focusing on FUL as the central formula of Kant’s ethics, Williams interprets moral imperatives as side-constraints on those pursuits which give our lives meaning. But the duties of virtue include both negative *and* positive duties.

The duty to promote our natural perfections, for example, not only directs us to take up various ground projects, but, being a wide or meritorious duty, “which of these natural perfections should take *precedence*, and in what proportion...are matters left for him [the individual] to choose in accordance with what sort of life he would like to lead and whether he has the powers necessary for it (*MS* 6:445). Thus, Korsgaard reminds us, “it is worth keeping in mind that natural perfection is a large category, including all the activities that cultivate body and mind” (Korsgaard, *CKE* 153). This would have settled the matter had Williams not developed his objections further in his essay “Moral Luck.”

In “Moral Luck” Williams asks us to imagine a hypothetical Gauguin case, in which the artist’s central ground project is to paint, but if he is to do so, he will neglect his obligations to his family. Here the demands of morality are in *direct conflict* with the ground project that affords the subject meaning. From this Williams concludes that an allegiance to Kantian morality robs the subject both of character and the ability to pursue a meaningful life.

Williams adduces three arguments in support of this conclusion (Williams, *ML* 22-39): (a) adhering to an impartial morality cannot be *reasonable* if it divests the agent of the activity that gives his life meaning; (b) the moral spectator has a reason to be glad that Gauguin succeeded in pursuing his art despite the moral cost since we value his paintings today; (c) in light of his success, Gauguin could not regret his choice, even if it involved abandoning his family.

The difficulty in responding to Williams's charge lies in the fact that "Williams is not misreading Kantian morality here. It does involve a requirement that one be prepared to set aside one's deepest projects if they require impermissible actions" (Herman, *PMJ* 39). Let us begin with Williams's first claim.

There are two responses to this charge. First, in Williams's example, the moral cost *might* strike us as negligible, or at least commonplace, since people regularly neglect their obligations to others. Notice, however that the more delineated the cost, the less plausible is Williams's assertion that an allegiance to an impartial morality is unreasonable. Let us imagine, for example, that in pursuing his art, Gauguin will not only abandon his family, but also plunge them into poverty. As Herman points out, "given the possibility of grossly immoral projects or vile actions taken for the sake of morally neutral projects, it does not seem rational to want it otherwise" (Herman, *PMJ* 39).

More importantly, FKE commands us to pursue those projects which *do not frustrate the deepest ends of others* under the broad categories of "one's own perfection" and "the happiness of others" (*MS* 6:385). Here, Williams's appeal to Tolstoy's Anna Karenina alongside his Gauguin example backfires precisely because he overlooks this point.

Williams argues that Anna's suicide, following her deteriorating relationship with Vronsky, is not a form of repentance (for leaving her husband and son), but stems from the fact that "she could have been justified *only by the life she hoped for*, and those hopes were...refuted by what happened" (Williams, *ML* 27, emphasis added). This may very well be Anna's logic, but Tolstoy's work is a deeply *moral* one: the strength of his novel lies in his highly sympathetic depiction of Anna's plight, while simultaneously concluding that *one's happiness cannot be built upon another's misery*. Tolstoy recognizes Anna's ground project, and what it means to her, as well as her deep-seated unhappiness

with Karenin, and the hypocrisy of the Russian aristocracy (who shun her), but he ultimately rejects her project on moral grounds.⁵⁶

Put differently, we may choose to give our lives meaning by pursuing the ends of self-love *unconditionally*. From a Kantian perspective, all such ground projects are “contrary to reason (even if their goals are pursued in accordance with technical or prudential rationality)” and to the spirit of FKE (Wood, *KET* 329). Perhaps the most telling proof of the latter is that Kantian ethics honors commitments to self-love as long as these are in agreement with the moral law, in which case they fall under the heading of *rational* self-love: “What it cannot do is honor *unconditional attachments*” (Herman, *PMJ* 39). If the claims of self-love assumed the “categorical” status that Williams would like to ascribe to them, this would amount to Kant’s definition of self-conceit: it would make self-love legislative.

Williams’s second claim is the least convincing of the three, since, as Herman points out, his “argument involves a confusion”: it does *not* follow from the fact that we value Gauguin’s paintings today that we must value the means that lead to their production (Herman, *PMJ* 40). Similarly, we might marvel at the Great Pyramid of Giza, and value its existence as an architectural wonder, without endorsing the institution of slavery that enabled its construction.

Finally, Williams argues that Gauguin could not have regretted his choice, given his artistic success. But Kantian ethics does *not* deny “the possibility of deeply satisfying lives that have been built on morally impermissible actions” (Herman, *PMJ* 41). In fact, self-conceit, as we have seen, is built exactly on such grounds: in defiance of the moral law, it lends its endorsement to self-love “mak[ing] the incentives of self-love and their inclinations the condition of compliance with the moral law” (*R* 6:36). The nature of this reversal is *delusory* precisely because it caters to the agent’s happiness. In short, Gauguin may not regret the moral cost of his actions at all, but the moral spectator need not agree.

⁵⁶ In fact, Tolstoy’s moralism is even more pronounced, since “the main idea, the one he struggled with most bitterly...was that Anna’s suicide was the punishment for her adultery. It was from this struggle with himself that he made the poetry of his heroine” (Pevear, Introduction xvi).

The underlying mistake in each of Williams's arguments is that he views the ends or projects of the agent in *isolation*, without regard to the *deepest ends* or *projects of others*. This is a move that FKE explicitly forbids, since these ends are not ones that can be shared by others and involve the frustration of their own ground projects.

The key, once again, lies in how we understand Kant's moral imperative. In Williams's reading, the individual's central ground projects are pitted against a formalistic (i.e. vacuous) moral law. Here the impartial demands of morality seem not only unreasonable, but a form of rule-worship that robs the agent of meaning. With FH and FKE, however, we see that what the law cannot reasonably sanction are projects which *fail to treat others as ends in themselves* or *involve ends that cannot be shared by them*.

Upholding the laws of morality when they conflict with the claims of self-love is not an "omission of character" (Williams, *ML* 14), for Kant, but the *condition* of having one at all. Kantian ethics "merely *infringes upon* self-love, inasmuch as it restricts it, as natural and active in us even prior to the moral law, to the ***condition of agreement with the law***, and then it is called *rational self-love*" (*KpV* 5:73, emphasis added). What it cannot do is honor *self-conceit* which lends its unqualified endorsement to self-love. This stance is essential to Kant's ethics, since his emphasis on virtue as "the moral strength of a *human being's* will in fulfilling his *duty*" makes the notion of character (both in striving to make duty the sufficient motive for one's actions and cultivating a joyous disposition), central to his project (*MS* 6:405). Nor can "categorical desires" ever be reasonable *if* they leave in their wake destitute families, or battlefields strewn with corpses, or desolate lovers (as we shall see next)—in the name of self-love.

Desolation

In her eloquent essay, "Duty and Desolation," Rae Langton charts the 1790s correspondence between Kant and Maria von Herbert, a young Austrian student of Kant's philosophy whom he had never met, and who, distraught, called on him "for help, for comfort, or for counsel to prepare me for death" (qtd. in Langton, *SS* 198).

Herbert had lost a person's love, who, in her eyes, "encompassed...all that is worthwhile" to "a long drawn out lie" (about a former lover who misused her trust) which she had finally disclosed to him. This recent desertion, on the heels of the first, had plunged her into despair: "I've had read the metaphysics of morals, and the categorical imperative, and it doesn't help a bit. My reason abandons me just when I need it. Answer me, I implore you—or you won't be acting in accordance with your own imperative" (qtd. in Langton, *SS* 198).

Impressed by her "deeply felt letter," Kant's carefully-worded reply outlined the difference between reticence and lying, and addressed Herbert's disconsolateness in the form of a short sermon. In her second letter, however, despair had given way to desolation:

My vision is clear now. I feel that a vast emptiness extends inside me, and all around me—so that I find myself to be superfluous, unnecessary. Nothing attracts me...the demands of morality are too easy for me. I would eagerly do twice as they command....I comfort myself with the thought that, since the practice of morality is so bound up with sensuality, it can only count for this world. I can hope that the afterlife won't be yet another life ruled by these few, easy demands of morality, another empty and vegetating life....I beg you to give me something that will get this intolerable emptiness out of my soul. (qtd. in Langton, *SS* 210).

Herbert ends her letter by begging Kant for permission to visit him in Königsberg, so that she can ask him in person "what kind of life your philosophy has led you to—whether it never seemed to you to be worth the bother to marry, or to give your heart to anyone" (qtd. in Langton, *SS* 210).

Kant did not reply to the letter. Instead, he enquired about Herbert's predicament through a mutual friend (J. B. Erhard), who informed him that Herbert "has capsized on the reef of romantic love. In order to realize an idealistic love, she gave herself to man who misused her trust. And then, trying to achieve such love with another, she told the new lover about the previous one" (qtd. in Langton, *SS* 216). Kant then passed on Herbert's letters as well as Erhard's explanation to Elisabeth Motherby (the daughter of one of his English merchant friends) as "an example of warning, to guard against the

wonderings of a sublimated fantasy” referring to Herbert as an “ecstatic little lady”—*die kleine Schwärmerin* [the little dreamer] with a “curious mental derangement” (qtd. in Langton, SS 216-7).

Roughly ten years later, Maria von Herbert committed suicide by drowning herself in the Drau River.

Langton argues for three main claims. The first takes issue with Kant’s *treatment* of Herbert. According to Langton, Kant fails Herbert on two fronts: (a) he ignores the core request of her letters, which is to demonstrate to her, in her desolation, that life has value, and (b) by bundling up her private correspondence and sending them to a third party, “Kant is doing something with her as one does a tool: Herbert cannot share the end of his action” (Langton, SS 218)—he is not only violating her trust, he is treating her as a *thing*. This is reflected by Kant’s shift in language: “my dear friend” is now *die kleine Schwärmerin*. Herbert is no longer a person, one who possesses “a heart created for the sake of virtue,” but a case study in “mental derangement.” With all these charges, we are wholly in agreement.

But Langton also makes two further claims: she argues that, unbeknownst to Kant and Herbert, Herbert is the “Kantian saint” which we saw in Wolf’s criticism of sainthood (see Introduction, § II), and that Herbert may have had a *duty to lie* to preserve humanity in her own person and the promise of an ideal kingdom of ends.

Langton’s first point is predicated on the view that in Kantian ethics “all inclinations are to be abjured as ‘blind and slavish’” (Langton, SS 214). *If* this familiar conception is true, and Herbert’s desolation has left her *without passions*, clinging to the husk of an unalloyed morality, it follows that Herbert represents the Kantian ideal: “She has no inclinations left to master. She respects the moral law, and obeys it. But she does not battle her passions to do so. She has no passions. She is empty....She is a Kantian saint” (Langton, SS 214). Langton argues that this is a profound indictment of Kantian ethics since it realizes the ideal Kantian subject, bereft of inclination, only to show her to be desolate, floundering in despair. The triumph of the categorical imperative is the vitiation of life.

Hopefully, at this stage, we have marshalled enough evidence to dispel this impoverished view of Kantian ethics which serves as Langton's basic premise. Kantian ethics is not hostile towards the inclinations, let alone built on their abjuration—a stance that puts Kant, by his own admission, at odds with the Stoics (*MS* 6:452; *R* 6:34-5, 57-8).

Langton comes close to recognizing this when she distinguishes between what she calls the “sane Kant”—“a warm and kind Kant, a Kant who thinks well of spontaneous natural sentiments, and thinks we should cultivate them, a Kant who shares much common ground with Aristotle” (Langton, *SS* 201)—and the “severe” Kant, describe by Blackburn as the “embodiment of pure practical reason, detached from all desires...a nightmare—of pure, authentic self-control....Context-free, non-natural, and a complete stickler for duty” (Blackburn, *Ruling Passions* 246-8). The problem is that, in her estimation, “the sane Kant is partly a reconstruction [by Kant's sympathizers]” (Langton, *SS* 201). What Kant requires, according to Langton, is a *synthesis* in the ethical sphere between the inclinations and morality: “We need both” (Langton, *SS* 215).

Is the sane Kant merely a construction? Does Kant truly believe, as Langton contends, that the “inclinations, all inclinations, are to be abjured as ‘blind and slavish’” (Langton, *SS* 214)? Here is Kant's emphatic answer:

[W]e must distinguish the *ethica morosa* [killjoy ethic]...from the *ethica rigida*. The latter demands that all duties determine strictly (*stricte*) and absolutely, yet not with the stifling of all sensory inclinations, insofar as they may give us pleasure with the approval of our intellectual nature. The *ethica morosa*, on the other hand, demands in addition, that a man repress all the joys in life, crucify himself and his flesh, renounce himself, and thus must be his own self-torturer. (*VE* 27:662-3)

The man, Kant argues, who lives by an *ethica morosa* violates humanity in his own person because he “must constantly punish and demean himself...as a criminal” and “must actually regard himself in a supersensible light, since he sacrifices all wished-for enjoyment of happiness” (*VE* 27:663).

The *ethica morosa* “renounces all claim to happiness”—this renunciation, according to Kant, “distinguishes it from morality, it is from that viewpoint unnatural” (VE 27:304).

The conscience which corresponds to an *ethica morosa* is a “morbid conscience” which “seeks to impute evil” to all our actions (VE 27:356). Echoing his condemnation of fantastic virtue, Kant insists that “conscience should not be a tyrant in us” (VE 27:356), and that “we ought not to deprive ourselves of all amenities and pleasures and enjoy none whatever; that would be a monk’s virtue to forgo everything that is *proper to human life*” (VE 27:393, emphasis added). Herbert, insofar as she is bereft of passion, lacks the pleasures and joys that are “proper to human life.” This is the verdict of Kantian ethics.

Far from Langton’s *severe* Kant, is the Kant who says, “the man who violates neither his duties to himself, nor those he has to others, may enjoy as much pleasure as he has the ability and taste for. In doing so, he remains always well disposed, and fulfils the purpose of his creation” (VE 27:395). Accordingly, Langton’s prescription that Kantian ethics requires both the inclinations and morality is superfluous since “the *praxis* of moral philosophy,” according to Kant, “consists in that formation of the inclinations and taste which makes us capable of uniting the actions that lead to our gratification with moral principles” (qtd. in Guyer, “Moral Feelings” 150). When Langton says that “we need both” the inclinations and morality what she does not realize is that her and Kantian ethics are already in accord.

Relying on Korsgaard’s essay, “The Right to Lie,”⁵⁷ Langton also argues that Herbert has a duty to lie for the sake of the kingdom of ends. Her argument begins with the Rawlsian distinction between ideal and non-ideal theory: *ideal theory* functions under the assumption that everyone will act justly (and that there is no history of oppression or injustice, etc.), but when circumstances are far from ideal, *non-ideal theory* tells us that “we must think of the Kingdom as an ideal to work towards, not to live by” (Langton, SS 219).

⁵⁷ It is important to note that Korsgaard does not endorse Langton’s conclusion (that Herbert has a duty to lie) in her essay, since Korsgaard reserves “evil” for those circumstances involving violence or death.

Now, Herbert exists in a patriarchal society in which women are on an unequal footing with men. Having achieved a relationship of mutual love and respect, she has only held one truth back. If she reveals it, she becomes a thing, a “second-hand good” in a sexual marketplace in which virginity fetches the highest price, and her friendship as well as the kingdom of ends will vanish (Langton, SS 221-2). If, on the other hand, she lies, she preserves both her friendship and the kingdom of ends as *goals*. Moreover, Langton believes that FH affords Herbert a duty to lie, because it instructs her to respect her own inalienable dignity and reject servility in all its guises.

These proposed revisions, although couched in Kantian terms, are decidedly *un-Kantian* in spirit. Langton acknowledges that, if this is the strategy for the kingdom’s sake, “Kant would not allow it,” but she hopes to “remain faithful to the Kantian spirit” (Langton, SS 218-22). Where does she go wrong?

First, her *strategy for the kingdom’s sake* endorses a form of consequentialism: it permits otherwise immoral actions, such as lying, for the sake of achieving a just state of affairs. Kant is clear, however, that even when a person “intend[s] to achieve a really good end by it,” lying is “contrary to his duty to himself” as a moral being (MS 6:430). We have a strict or perfect duty *to ourselves* against lying (see Ch. 3, § VI). This takes us to the second point.

In light of the former, Langton’s appeal to FH is surprising for two reasons: (a) by lying, Herbert does not preserve her self-respect, according to Kant, but subverts it: “By a lie a human being throws away and...annihilates his dignity as a human being” since he treats humanity in his person as a mere means (MS 6:429). Herbert recognizes this when she writes, “[when] finally I had the strength and revealed the truth to my friend...the stone in my heart was gone” (qtd. in Langton, SS 209). And (b) by Langton’s own admission, if Herbert’s friend was deceived,

and it seems he was, she forced him to perform actions he had no chance to choose. To that extent, she made him thing-like. She prevented him from being the initiator of his own action....Herbert failed to treat

him as an end, in the Kantian sense: as a being who must be able ‘to share the end of the action’.

(Langton, SS 207-8)

Of course, the action that Herbert thwarted, and which her friend thought he was performing, was “courting a beautiful young virgin” (Langton, SS 208). As objectionable as this formulation may be, FH forbids us from regulating the agent’s choice by robbing him of it.

This is why Kant’s words of “comfort” are both of little consolation to Herbert *and* indicative of a Kantian position that goes unacknowledged by Langton. He writes: “only time will be needed to...transform his [Herbert’s friend’s] coldness into a more firmly grounded love. If this doesn’t happen, then the earlier warmth of his affection was more physical than moral, and would have disappeared anyway...” (qtd. in Langton, SS 199). Having told the truth, there is nothing more Herbert can do. The decision, and thus the moral onus, now rests with her friend: if he acts *morally*, according to Kant, his initial anger will give way to a “firmly grounded love”; if, on the other hand, he rescinds his affections, then they were suspect all along. This suggests that Kant believes that forsaking Herbert, even in light of her deception, is *vicious*.⁵⁸

Finally, as James Edwin Mahon points out, Langton’s argument “proves too much”: if sexism and patriarchy count as circumstances of evil (which they certainly are) that *permit lying*—then “lying is widely, if not universally, permissible for women” since sexism and patriarchy are “sadly, near-universal, if not universal” (Mahon, “Kant and Maria” 443). Far from securing the kingdom of ends, such a strategy would banish it from the horizon altogether.

From these three points, we can conclude that the kingdom of ends is not merely “a long-term political and moral goal for humanity” (Korsgaard, CKE 153). We must act, according to FKE, as if we are *already* inhabitants of the former: “every rational being must so act as if through its maxims it were *at all times* a legislating member of the kingdom of ends” (G 4:438, emphasis added). Kant recognized fully

⁵⁸ Compare Langton’s interpretation of Kant’s response: “Kant’s reply suggests that love is deservedly lost, that misery is the appropriate response to one’s moral failure...” (Langton, SS 200).

the evil that exists in civilized society and frustrates our efforts at every turn, but he modeled the kingdom of ends on what Augustine and Leibnitz called the “kingdom of grace” (VE 29:610-1; 629), and it retains a similar revolutionary element insofar as it

presupposes in all this that there is still a germ of goodness...a germ that cannot be extirpated or corrupted....[which requires] a *revolution* in the disposition of the human being (a transition to the maxim of holiness of disposition). And so a “new man” [who] can come about only through a kind of rebirth, as it were a new creation...a change of heart. (R 6:45-7)

Kant reconciles this idea of a “rebirth” with the endless striving of virtue by arguing that “revolution is necessary in the mode of thought” while “incessant laboring and becoming” is necessary in the mode of sense (R 6:47-8). Combining the two, we get the following: an unalterable decision (which reverses the supreme ground of corrupt maxims) makes the human being receptive to the good *before* she is good.

This means that the kingdom of ends is both “a theoretical idea for explaining what exists” and “a practical idea for the sake of bringing about—in conformity with precisely this idea—what does not exist but can become actual *by means of our behavior*” (G 4:436n, emphasis added).

The kingdom is here now, in the muck and mire, guided by the unrelenting prescription of humanity as an end in itself: it is the condition of character which is necessary for virtue since the latter “must be coupled with...rules of prudence, calling for a contented outlook, peace of mind, and freedom from reproach” (VE 27:465). It is the social realization of an ethical community of equals who actively participate in each other’s joys and sufferings, and pursue their ends in common, abolishing competition, rivalry and self-conceit. It is our recourse from the Hobbesian distrust inherent in our unsociable sociability which prevents us from revealing ourselves to others and communing with them. It is the duty which delivers us from desolation.

Conclusion

The “aretaic turn” in moral philosophy ushered in a host of criticisms of Kantian ethics centered of Kant’s moral “rigorism” and “empty formalism.” Although these concerns had been around since Kant’s time, virtue ethicists found in them a specific critique of Kantian ethics. They charged Kant with failing to provide an adequate account of the virtues, and with it, the attendant desiderata of sociality, character and the emotions, which they saw as central to a meaningful life. The Kantian agent was a “moral saint” (Wolf, “Moral Saints” 90), severely restricted in both his activities and traits, “detached from all desires” (Blackburn, *Ruling Passions* 246), sacrificing a life of substance to the form a contentless, impartial law (Williams, *ML* 18). “To be genuinely moral,” according to Kant’s critics, the Kantian agent had to

tear himself from his inclinations a loving human being, drown the sympathetic promptings of his heart, scorn any fruits of his efforts, think last of all of the feelings, needs, desires, and inclinations either of himself or his fellows and, perhaps detesting what he has to do, do it anyway—solely from respect for the Law. (Taylor, *Good and Evil* 113)

In response, we argued that this impoverished view of Kantian ethics stems from an overemphasis on the *Groundwork* and FUL as the key elements in his moral philosophy. Once we turn to FH and the Doctrine of Virtue in the *Metaphysics of Morals* (as Kant clearly intended), a rich and novel ethical theory emerges, and many of the criticisms from the field of virtue ethics are subsequently disarmed. To stay true to Kant’s scholastic presentation of the moral law, the chapters closely followed its progression from *form*, to *matter*, to its *complete determination* (G 4:436).

Chapter 1 began by offering a deflationary account of the good will and acting from duty in the *Groundwork*. Kant’s notion of “unsociable sociability” was introduced as one of the key reasons for his rejection of the inclinations as moral motives. This was followed by a discussion of the place of happiness in Kant’s hierarchy of imperatives which, we claimed, was merely an expression of the idea that moral considerations should be overriding. Kant refers to an inversion of this hierarchy whereby

prudential imperatives assume categorical form as “self-conceit.” Finally, owing to their formal, provisional and negative nature, FUL and its variant, FLN, which represent the *form* of the moral law, were found to suffer from congenital defects that fail to secure the conclusions that Kant’s four examples are meant to deliver.

In Chapter 2 we turned to the *matter* or ground of the moral law, which we identified as rational nature as an end in itself. Rational nature (a) fits Kant’s three value conceptions as an existent end-in-itself which has absolute worth, and (b) belying charges of “empty formalism,” acts as the fundamental value underpinning Kantian ethics. This establishes FH as the primary formula in Kant’s ethical thought. Before turning to his derivation of the formula, we located the radical egalitarianism of FH in the “predisposition of humanity” which Kant defines as the general capacity to set ends. Of the two arguments in support of FH, we found Kant’s “argument from elimination” unconvincing, while defending a modified version of Korsgaard’s “regress on conditions” which involves an inference from the objective value we place on the ends we set, to the unconditional value of the rational nature that set them. After responding to five possible objections to FH, we revisited Kant’s four examples through the lens of FH to show that the structural problems that attended FUL/FLN were now resolved.

Having established the basic value in Kantian ethics as humanity as an end in itself, in Chapter 3 we argued for a new “expressive ethics” based on FH and defended the *Metaphysics of Morals* as the culmination of Kant’s ethical thought. After defining virtue as an endless progress towards holiness, we turned to the positive role of FH in furnishing ends that are also duties. Kant identifies two such ends: our own perfection and the happiness of others. The two ends, in turn, generated fourteen duties of virtue. Here we saw that, for Kant, (a) “conceptions of the ‘moral point of view’ and the appropriate demandingness of morality, cannot be understood without a conception of the relevant virtues” (Swanton, *Virtue Ethics* 5), and (b) the formula grounding these virtues, or morally prescribed ends is FH. Of special importance were Kant’s duties of love which, we argued, required agents to cultivate certain

affective dispositions *constitutive* of the relevant end. Finally, we addressed the charge of “rigorism” by pointing to the wide latitude permitted by imperfect duties and rejecting any rigorist interpretation of Kant’s ethics in light of his condemnation of “fantastic virtue” which threatens to “turn the government of virtue into tyranny” (*MS* 6:409).

In Chapter 4 we looked at the place of the emotions in Kant’s ethics, mounting an argument from the inclinations to the *complete determination* of the moral law in the form of the kingdom of ends, and its corresponding formula, FKE. We began by surveying Kant’s account of desire, the affects, passions, and delusions in the *Anthropology*. In the social passions we found the anthropological underpinnings of Kant’s idea of unsociable sociability: the manias for honor, dominance, and possession. In response to Williams’s charge that Kantian ethics denies the emotions a substantive role in morality, we argued for the importance of: (a) “moral feelings” which involve *rational* desires and serve as “aesthetic preconditions” of receptivity to duty, without which moral conduct would be impossible, and (b) empirical feelings which not only attend a joyous, and hence genuine, virtuous disposition, but are also necessary for a moral agent to enjoy his existence, so that he “remains always well disposed, and fulfils the purpose of his creation” (*VE* 27:395). Contrasting Aristotle’s conception of virtue with Kant’s, we defended Kant from the criticism that Kantian virtue is an inferior species of virtue, aligned with Aristotelian self-control.

In the second half of the chapter, we argued that the enemy of virtue resides not in the inclinations, but in reason, which lends its endorsement in the form of a socially conditioned self-conceit to the claims of self-love over those of the moral law. Because the nature of self-conceit is social, Kant believes it requires a social remedy: this is the kingdom of ends in which “self-love is swallowed up in the idea of generous mutual love” (*VE* 27:423). Having finally arrived at the *complete determination* of the moral law, we were now in position to address the desiderata of sociality, character and emotions. The first of these finds its full realization in the “friendship of disposition and sentiment” which

prefigures the kingdom of ends. Just as true friendship delivers us from the distrust and reticence which prevents us from revealing ourselves to others and being wholly companionate, the kingdom of ends delivers us from the unsociable sociability inherent in civilized society: competition, enmity and social antagonism give way to mutual cooperation, amity, and the pursuit of common ends. When we turned to Williams's "ground projects," far from an impartial Kantian morality built on the omission of character, what we found was the very *condition* of character, without which "categorical desires" would supplant the moral law, time and time again, without regard for other rational beings or their ends. Finally, we argued that the kingdom of ends is what ensures that we are free from the desolation at the heart of our unsociable sociability, since FKE directs us to be openhearted, benevolent, joyful, and to participate in the joys and sufferings of others in the form of a free communion of sentiment.

Rae Langton argues that there are two Kants: the "sane Kant" which we have been describing and the "severe Kant" of FUL and the *Groundwork*, found in the pages of so many scholarly journals. For Langton, the sane Kant "is partly a reconstruction, but a reconstruction well worth performing" (Langton, SS 201). In the final analysis, however, she finds the sane Kant difficult to reconcile with his teachings (Langton, SS 214). If we have been successful in our arguments, in the process of responding to the criticisms of Kant's moral philosophy from the field of virtue ethics, we have shown that once we shift the focus from FUL and the *Groundwork* to FH and the Doctrine of Virtue in the *Metaphysics of Morals*, a new ethics employing a rich taxonomy of virtues emerges, and with it, a new Kant.

Bibliography

- Adorno, Theodor W. *Problems of Moral Philosophy*. Trans. Rodney Livingstone. Ed. Thomas Schröder. Stanford: Stanford University Press, 2001.
- Allison, Henry E. *Kant's Theory of Freedom*. New York: Cambridge University Press, 1990.
- Anscombe, G. E. M. "Modern Moral Philosophy." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997. 26-44.
- Aristotle. *Nicomachean Ethics*. Trans. Christopher Rowe. Introduction and Commentary. Sarah Broadie. New York: Oxford University Press, 2002.
- . *Nicomachean Ethics*. 2nd ed. Trans. Terence Irwin. Indianapolis: Hackett Publishing Company, 1999.
- Badiou, Alain. *In Praise of Love*. Trans. Peter Bush. New York: The New Press, 2012.
- Baron, Marcia M. *Kantian Ethics (Almost) Without Apology*. New York: Cornell University Press, 1995.
- Baron, Marcia and Melissa Seymour Fahmy. "Beneficence and Other Duties of Love in *The Metaphysics of Morals*." *The Blackwell Guide to Kant's Ethics*. Ed. Hill, Thomas E. Jr. Oxford: Wiley-Blackwell Publishing, 2009. 211-28.
- Beck, Lewis White. *A Commentary on Kant's Critique of Practical Reason*. Chicago: The University of Chicago Press, 1960.
- Blackburn, Simon. *Ruling Passions*. New York: Clarendon Press, 1998.
- Broadie, Sarah. Philosophical Introduction. *Nicomachean Ethics*. By Aristotle. Trans. Christopher Rowe. Commentary. Sarah Broadie. New York: Oxford University Press, 2002. 9-91.
- Cioran, E. M. *A Short History of Decay*. Trans. Richard Howard. New York: Arcade Publishing, 1975.
- Colman, Andrew M. "Relative Deprivation." *Oxford Dictionary of Psychology*. 3rd ed. New York: Oxford University Press, 2009. 648-9.
- Cottingham, John. "Partiality and the Virtues." *How Should One Live? Essays on the Virtues*. Ed. Roger Crisp. New York: Oxford University Press, 1996.
- Cohen, G. A. "Reason, Humanity, and the Moral Law." *The Sources of Normativity*. Ed. Onora O'Neill. Commentary. G. A. Cohen, Raymond Guess, Thomas Nagel and Bernard Williams. Cambridge: Cambridge University Press, 1996. 167-88.
- Dean, Richard. "The Formula of Humanity as an End in Itself." *The Blackwell Guide to Kant's Ethics*. Ed. Hill, Thomas E. Jr. Oxford: Wiley-Blackwell Publishing, 2009. 83-101.
- Dent, N. J. H. *The Moral Psychology of the Virtues*. New York: Cambridge University Press, 1984.

- Donagan, Alan. *The Theory of Morality*. Chicago: University of Chicago Press, 1979.
- Epictetus. *Discourses and Selected Writings*. Ed. and trans. Robert Dobbin. New York: Penguin Books, 2008.
- Foot, Philippa. "Virtues and Vices." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997. 163-77.
- Grenberg, Jeanine. "What is the Enemy of Virtue?" *Kant's Metaphysics of Morals: A Critical Guide*. Ed. Lara Denis. New York: Cambridge University Press, 2010. 152-69.
- Guyer, Paul. *Kant on Freedom, Law and Happiness*. New York: Cambridge University Press, 2000.
- . "Moral Feelings in the Metaphysics of Morals." *Kant's Metaphysics of Morals: A Critical Guide*. Ed. Lara Denis. New York: Cambridge University Press, 2010. 130-51.
- Herman, Barbara. "Could It Be Worth Thinking About Kant on Sex and Marriage?" *A Mind of One's Own*. 2nd ed. Ed. Louise M. Antony and Charlotte E. Witt. Colorado: Westview Press, 2002.
- . "The Difference That Ends Make." *Perfecting Virtue: New Essays on Kantian Ethics and Virtue*. Ed. Lawrence Jost and Julian Wuerth. New York: Cambridge University Press, 2011. 92-115.
- . *The Practice of Moral Judgement*. Cambridge: Harvard University Press, 1993.
- Hill Jr., Thomas E. *Human Welfare and Moral Worth: Kantian Perspectives*. New York: Clarendon Press, 2002.
- Hursthouse, Rosalind. *On Virtue Ethics*. New York: Oxford University Press, 1999.
- Irwin, T. H. "Kant's Criticisms of Eudaemonism." *Aristotle, Kant and the Stoics: Rethinking Happiness and Duty*. Ed. Stephen Engstrom and Jennifer Whiting. Cambridge: Cambridge University Press, 1996. 63-101.
- Kant, Immanuel. *Anthropology from a Pragmatic Point of View*. Ed. and trans. Robert B. Louden. Introduction. Manfred Kuehn. New York: Cambridge University Press, 2006.
- . *Critique of Practical Reason*. Ed. and trans. Mary Gregor. Introduction. Andrews Reath. New York: Cambridge University Press, 1997.
- . *Groundwork of the Metaphysics of Morals*. Revised Edition. Ed. and trans. Mary Gregor and Jens Timmermann. Introduction. Christine M. Korsgaard. New York: Cambridge University Press, 2012.
- . *Idea of a Universal History with a Cosmopolitan Aim*. Trans. Allen W. Wood. *Anthropology, History, and Education*. Ed. Günter Zöller and Robert B. Louden. New York: Cambridge University Press, 2007.
- . *Lectures on Ethics*. Trans. Peter Heath. Ed. Peter Heath and J. B. Schneewind. New York: Cambridge

- University Press, 1997.
- . *The Metaphysics of Morals*. Ed. and trans. Mary Gregor. Introduction. Roger J. Sullivan. New York: Cambridge University Press, 1996.
- . *Practical Philosophy*. Ed. and trans. Mary J. Gregor. Introduction. Allen Wood. New York: Cambridge University Press, 1996.
- . *Religion within the Boundaries of Mere Reason*. Ed. and trans. Allen Wood and George di Giovanni. Introduction. Robert Merrihew Adams. New York: Cambridge University Press, 1998.
- . *Religion and Rational Theology*. Ed. and trans. Allen Wood and George di Giovanni. New York: Cambridge University Press, 1996.
- Kerstein, Samuel. "Deriving the Formula of Humanity (GMS II, 427-437)." *Groundwork for the Metaphysics of Morals*. Ed. Christoph Horn and Dieter Schönecker. Berlin: Walter de Gruyter, 2006. 200-21.
- Korsgaard, Christine M. *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- . "From Duty and for the Sake of the Noble: Kant and Aristotle on Morally Good Action." *Aristotle, Kant and the Stoics: Rethinking Happiness and Duty*. Ed. Stephen Engstrom and Jennifer Whiting. Cambridge: Cambridge University Press, 1996. 203-36.
- Kuehn, Manfred. "Kant's *Metaphysics of Morals*: The History and Significance of its Deferral." *Kant's Metaphysics of Morals: A Critical Guide*. Ed. Lara Denis. New York: Cambridge University Press, 2010. 9-27.
- Langton, Rae. *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*. New York: Oxford University Press, 2009.
- Louden, Robert B. "Evil Everywhere: The Ordinarity of Kantian Radical Evil." *Kant's Anatomy of Evil*. Ed. Sharon Anderson-Gold and Pablo Muchnik. New York: Cambridge University Press, 2010. 93-115.
- Louden, Robert B. *Kant's Impure Ethics: From Rational Beings to Human Beings*. New York: Oxford University Press, 2000.
- MacIntyre, Alasdair. *A Short History of Ethics: A History of Philosophy from the Homeric Age to the Twentieth Century*. Cornwall: Routledge Classics, 1966.
- . *After Virtue: A Study in Moral Theory*. 3rd ed. Indiana: University of Notre Dame Press, 1981.
- . "The Nature of the Virtues." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997. 118-40.
- Mahon, James Edwin. "Kant and Maria von Herbert: Reticence vs. Deception." *Philosophy* 81 (2006):

417-44.

Marx, Karl. *Economic and Philosophical Manuscripts of 1844*. Ed. and trans. Martin Milligan. New York: Dover Publications, 2007.

---. "On the Jewish Question." *Selected Writings*. Ed. Lawrence H. Simon. Indianapolis: Hackett Publishing Company, 1994.

---. *Wage-Labour and Capital & Value, Price and Profit*. New York: International Publishers, 1976.

Murdoch, Iris. *The Sovereignty of Good*. New York: Routledge Classics, 1971.

McDowell, John. *Mind and World*. Cambridge: Harvard University Press, 1996.

Papadaki, Evangelia (Lina). "Feminist Perspectives on Objectification." *The Stanford Encyclopedia of Philosophy* (Winter 2012 Edition). Ed. Edward N. Zalta. URL = <http://plato.stanford.edu/archives/win2012/entries/feminism-objectification/>.

Paton, H. J. *The Categorical Imperative: A Study in Kant's Moral Philosophy*. Philadelphia, Pennsylvania: University of Pennsylvania Press, 1947.

Pevear, Richard. Introduction. *Anna Karenina*. By Leo Tolstoy. Trans. Richard Pevear and Larissa Volokhonsky. New York: Penguin Books, 2000. vii-xvii.

Pippin, Robert B. "Kant's Theory of Value: On Allen Wood's *Kant's Ethical Thought*." *Inquiry* 43 (2000): 239-66.

Potter, Nelson. "Duties to Oneself, Motivational Internalism, and Self-Deception in Kant's Ethics." *Kant's Metaphysics of Morals: Interpretative Essays*. Ed. Timmons, Mark. New York: Oxford University Press, 2002. 371-90.

---. "The Argument of Kant's *Groundwork*, Chapter 1." *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Ed. Paul Guyer. New York: Rowman & Littlefield Publishers, 1997. 29-50.

Regan, Tom. *The Case for Animal Rights*. Los Angeles: University of California Press, 1983.

Rousseau, Jean-Jacques. *The Discourses and Other Early Political Writings*. Ed. and trans. Victor Gourevitch. New York: Cambridge University Press, 1997.

Schneewind, J. B. Introduction. *Lectures on Ethics*. By Immanuel Kant. Trans. Peter Heath. New York: Cambridge University Press, 1997. xiii-xxvii.

Scruton, Roger. *Kant: A Very Short Introduction*. New York: Oxford University Press, 2001.

Sedgwick, Sally. *Kant's Groundwork of the Metaphysics of Morals: An Introduction*. New York: Cambridge University Press, 2008.

Seymour, Melissa M. *Duties of Love and Kant's Doctrine of Obligatory Ends*. Diss. Indiana University, 2007. Ann Arbor: UMI 3277970.

- Sherman, Nancy. *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press, 1997.
- Slote, Michael. *From Morality to Virtue*. New York: Oxford University Press, 1992.
- Sneddon, Andrew. "A New Kantian Response to Maxim-Fiddling." *Kantian Review* 16 (2011): 67-88.
- Stocker, Michael. "The Schizophrenia of Modern Ethical Theories." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997. 66-78.
- Swanton, Christine. *Virtue Ethics: A Pluralistic View*. New York: Oxford University Press, 2003.
- Taylor, Richard. *Good and Evil. A New Direction: A Forceful Attack on the Rationalistic Tradition in Ethics*. New York: Prometheus Books, 1984.
- Velleman, J. David. "Love as a Moral Emotion." *Self to Self: Selected Essays*. New York: Cambridge University Press, 2006. 70-109.
- Whiting, Jennifer. "Self-Love and Authoritative Virtue: Prolegomenon to a Kantian Reading of Eudemian Ethics viii 3." *Aristotle, Kant and the Stoics: Rethinking Happiness and Duty*. Ed. Stephen Engstrom and Jennifer Whiting. Cambridge: Cambridge University Press, 1996. 162-99.
- Williams, Bernard. *Moral Luck: Philosophical Papers 1973-1980*. Cambridge: Cambridge University Press, 1981.
- . "Morality and the Emotions." *Problems of the Self*. New York: Cambridge University Press, 1973. 207-229.
- Wolf, Susan. "Moral Saints." *Virtue Ethics*. Ed. Roger Crisp and Michael Slote. New York: Oxford University Press, 1997. 79-98.
- Wood, Allen. "Duties to Oneself, Duties of Respect to Others." *The Blackwell Guide to Kant's Ethics*. Ed. Hill, Thomas E. Jr. Oxford: Wiley-Blackwell Publishing, 2009. 229-51.
- . General Introduction. *Religion and Rational Theology*. By Immanuel Kant. Ed. and trans. Allen Wood and George di Giovanni. New York: Cambridge University Press, 1996. xi-xxiv.
- . "Humanity as End in Itself." *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Ed. Paul Guyer. New York: Rowman & Littlefield Publishers, 1997. 165-188.
- . "Kant and Agent-Oriented Ethics." *Perfecting Virtue: New Essays on Kantian Ethics and Virtue*. Ed. Lawrence Jost and Julian Wuerth. New York: Cambridge University Press, 2011. 58-91.
- . *Kant's Ethical Thought*. New York: Cambridge University Press, 1999.
- . *Kantian Ethics*. New York: Cambridge University Press, 2008.
- . "The Final Form of Kant's Practical Philosophy." *Kant's Metaphysics of Morals: Interpretative Essays*. Ed. Timmons, Mark. New York: Oxford University Press, 2002. 1-22.