

KANT'S  
*Metaphysics of Morals*  
A Critical Guide

EDITED BY  
LARA DENIS  
*Agnes Scott College*



## Contents

<i>Notes on contributors</i>	page vii
<i>Acknowledgements</i>	ix
<i>List of translations and abbreviations</i>	x
Introduction	I
LARA DENIS	
1 Kant's <i>Metaphysics of Morals</i> : the history and significance of its deferral	9
MANFRED KUEHN	
2 Reason, desire, and the will	28
STEPHEN ENGSTROM	
3 Justice without virtue	51
KATRIN FLIKSCHUH	
4 Kant's innate right as a rational criterion for human rights	71
OTFRIED HÖFFE	
5 Intelligible possession of objects of choice	93
B. SHARON BYRD	
6 Punishment, retribution, and the coercive enforcement of right	111
ALLEN W. WOOD	
7 Moral feelings in the <i>Metaphysics of Morals</i>	130
PAUL GUYER	
8 What is the enemy of virtue?	152
JEANINE GRENBERG	

9	Freedom, primacy, and perfect duties to oneself LARA DENIS	170
10	Duties to and regarding others ROBERT N. JOHNSON	192
11	Duties regarding animals PATRICK KAIN	210
12	Kant's <i>Tugendlehre</i> as normative ethics THOMAS E. HILL, JR.	234
	<i>Bibliography</i>	256
	<i>Index</i>	264

# *Introduction*

*Lara Denis*

The *Metaphysics of Morals*, Kant's final major work in moral philosophy, is wide-ranging, complex, and often provocative. Its focus is not rational beings as such, but human beings in particular, and the duties, rights, and morally practical relations that obtain among us. As a whole, the *Metaphysics of Morals* deepens Kant's accounts of morality, moral psychology, and the spheres of right and virtue. Its *Doctrine of Right* sets forth not only fundamental, coercively enforceable principles of interpersonal conduct, but also a rational idea of a peaceful community of all nations. Its *Doctrine of Virtue* develops Kant's conception of virtue and accounts of particular ethical duties. Aspects of the *Metaphysics of Morals*, such as its overall lack of abstraction relative to the *Groundwork* or second *Critique* and its richness of anthropological detail, recall Kant's lectures on ethics. Other features, such as its preliminary discussion of the relation of the faculty of desire to practical principles, clearly exhibit the commitments of his critical moral philosophy.

Some of the positions that Kant articulates here – for example, about the relationship between the fundamental principles of right and the categorical imperative, or the role of feeling in moral motivation – have long been sources of dispute among Kant's interpreters. Some of his arguments – for instance, those concerning capital punishment, or duties regarding non-human animals – have ignited debate among ethicists more broadly. The *Metaphysics of Morals* invites questions about where and how Kant augments, refines, revises, withdraws, or supplants arguments and positions set forth in earlier works, as well as about how best to understand the arguments and positions he provides here, and how plausible, defensible, or distinctive they are.

The dozen essays collected here address questions both interpretive and philosophical. They focus on passages, positions, or arguments in the *Metaphysics of Morals* that strike us as particularly interesting and important – and that we hope will engage the interest of colleagues specializing

in ethics, political philosophy, Kant interpretation, and the history of philosophy. Many of the essays in this collection draw heavily on other (often less-studied) works by Kant, thereby enriching our understanding of their topic and the relevant arguments in the *Metaphysics of Morals*. There is much here both for readers interested in the development of Kant's own thought or the history of ethics and for those curious about how best to construe and evaluate various facets of Kant's mature moral and political philosophy.

A natural place to begin our exploration is by asking about the nature of Kant's project in the *Metaphysics of Morals*. How did he conceive of this work? Why did it take so long for him to write it? How does it relate to his more overtly "critical" works in ethics? In the opening chapter, Manfred Kuehn situates the *Metaphysics of Morals* within the broad contexts of Kant's philosophical development and ethical thought. He argues that Kant's *Metaphysics of Morals* should not be disqualified as precritical, for although it contains precritical elements, they are not what is most important about it. Furthermore, it would be wrong simply to judge the *Groundwork* and the second *Critique* from the point of view of "the final form of Kant's practical philosophy," for the concerns of the *Metaphysics of Morals* are different from those of his critical philosophy. In order to establish the significance and place of Kant's late work on moral philosophy in relation to his philosophy as a whole, Kuehn first explores the history – or prehistory – of its composition, which is primarily a history of postponements. The projected *Metaphysics of Morals* radically changed several times in these deferrals, and the ways in which the project changed are not insignificant for understanding the place it ultimately assumed. Second, Kuehn briefly indicates what Kant's lectures reveal about his project of a "Metaphysics of Morals," suggesting that the relevance of these lectures has been largely underestimated in discussions of the "final form" of Kant's practical philosophy. Third, Kuehn shows how the *Metaphysics of Morals* is related to the *Groundwork* and the second *Critique*. He argues that the later work both responds to concerns different from those of the earlier two, and yet also reveals an important, new perspective on issues central to those earlier works, such as the categorical imperative.

The rest of the essays in the first half of the book focus on questions raised primarily by Part I of the *Metaphysics of Morals*, that is, the Introduction to the *Metaphysics of Morals* and the *Doctrine of Right* proper. In chapter 2, Stephen Engstrom turns our attention to a topic within the Introduction to the *Metaphysics of Morals* that goes to the heart of Kant's moral philosophy: Kant's conception of the will as it relates to reason and

desire. While much attention has been devoted to Kant's doctrine of the will's autonomy, Engstrom points out that Kant advances another striking proposition about the will: it is nothing but practical reason. Engstrom contends that this less-investigated idea is vitally important in its own right as well as for the light it throws on other aspects of his ethical thought, including autonomy. According to Engstrom, Kant *transforms* the traditional understanding of the will as rational desire. Kant's way of combining the notions of the will, reason, and desire in his conception of a practical or efficacious employment of the cognitive capacity of reason is responsible for much of what is distinctive in his moral philosophy. After outlining Kant's conception of reason as a cognitive capacity, Engstrom takes up Kant's conception of desire as a form of causality peculiar to living beings, which provides the broad heading under which he situates the will and relates it to the power of choice. He then draws on the preceding accounts of reason and desire to elucidate Kant's account of the will. Engstrom's interpretation places the will in the desiderative economy of human life, while underscoring the cognitivist character of Kant's conception of the will and of his practical philosophy more generally.

Chapter 3 offers a critical discussion of ongoing debates concerning the moral status of Kant's philosophy of right. Katrin Flikschuh defends an account of right as a public morality and, as such, as systematically distinct from the personal morality of Kant's ethics. It follows that the principle of autonomy, as a principle of *ethical* self-obligation, has no place within the philosophy of right which, as public, concerns the morality of *external* legislation. From the irrelevance of the principle of autonomy the non-moral status of right does not, however, follow. Flikschuh employs Kant's *Wille/Willkür* distinction to show that, within the domain of right, the a priori idea of the general united will replaces the principle of autonomy as the ground of moral obligation. As *Wille* in its juridical conception, the idea of the general united will locates the grounds of juridical moral obligation outside the subjective will of the individual agent, ensuring conformity of action of the subjective *Willkür* of each with universal laws of right. The externally free agent is, as such, *non-autonomously* free. One implication of this view is that the presumed centrality of the principle of autonomy to Kant's practical philosophy in general must be reassessed; the principle of autonomy is derivative of the philosophically deeper idea of freedom itself. Flikschuh resists attempts to close the gap between ethical and political judgment. Instead of seeking to align political with ethical judgment, we should, she says, acknowledge the political as a distinct mode of public moral judgment.

Chapter 4 considers questions concerning Kant's conception of human rights (or of the *one* human right) and how it operates within Kant's political philosophy – especially according to the *Rechtslehre*. Here, Otfried Höffe explicates the innate right to freedom, which, Kant says, belongs to every human being “by virtue of his humanity.” He aims to show how this single innate right functions as a rational criterion for multiple human rights. To this end, Höffe clarifies Kant's distinction between moral (natural) right and positive (empirical) right; the relation between the moral concept of right, the universal principle of Right, and the universal law of Right; and Kant's justification of coercive enforcement of legal rights. Provocatively, Höffe argues that Kant views self-recognition – specifically, the practical recognition of humanity in one's own person by upholding one's rightful honor and refusing to submit to legal degradation – as a primary condition for establishing oneself as a legal entity. Höffe then explores the derivation of the four human rights Kant regards as implicit within the innate right to freedom. Finally, he suggests that while the rights to own property and to live in a public legal order are not, for Kant, human rights in the strict sense, they are grounded in such a way as to be considered “quasi-human rights.”

Chapter 5 addresses a question fundamental to Kant's doctrine of private right: how to have something external as one's own. Sharon Byrd traces Kant's arguments and shows that they turn on his concepts of possession. These concepts move from an *empirical concept* of possession as having something in one's hand to an *intelligible concept* of possession as having something as one's own based on a duty all others have not to interfere with what one intelligibly possesses. His arguments depend on the postulate of practical reason. This postulate has been interpreted to provide a justification. A justification, however, suggests that what would otherwise be wrongful or prohibited conduct is rightful conduct because of the situation. Byrd's position is that there is nothing wrongful about taking something external to oneself and calling it one's own. The taker thus does not need any justification for doing so. Byrd here relies on an alternative interpretation of the postulate as a *power-conferring norm*. On this reading, the postulate empowers us to have external objects of our choice as our own. Although we may unilaterally impose an obligation on all others to respect what we have declared to be ours, this power flows from our will's compatibility with the universal united will. Nothing in Kant's arguments for individual rights to have objects of choice as our own depends on the existence of a state. Indeed Kant notes that without a right to property and other objects of our choice there would be no duty

to move to the civil social order. Property rights therefore are rights we have in the state of nature. They do not depend on social approval any more than our right to freedom of choice in general depends on social approval and recognition. The sole purpose of the state for Kant is securing rights we already have before leaving the state of nature and moving to the civil state. The state secures our right to freedom and our rights to external objects of our choice.

Chapter 6 tackles questions of the substance and justification of Kant's theory of punishment. Regarding issues about legal punishment, Kant is best known as a defender of an extreme retributivist position on the justification of punishment. Allen Wood argues, however, that the deeper truth about his views on this topic is far more complex and even troubling. According to Wood, although Kant is undoubtedly a retributivist, the justification of punishment Kant provides that is best rooted in his theory of right is *not* a retributivist one. Furthermore, Kant's retributivism is apparently *inconsistent* with some fundamental tenets of his practical philosophy.

The remaining chapters focus primarily on the *Doctrine of Virtue*. In chapter 7, Paul Guyer considers the relation of feelings to moral motivation, and traces the development of Kant's view of this matter. Kant is commonly supposed to have excluded all feeling from the incentives of morally worthy action, and accepted only the determination of the will by the moral law itself as a morally worthy motivation. Guyer shows that this view ignores Kant's increasingly sophisticated moral psychology, which reaches its zenith in the *Metaphysics of Morals*. In the *Groundwork*, Kant recognizes a feeling of respect as the *effect* of the moral law, but does not assign it any clear role in the etiology of moral action. By the *Critique of Practical Reason*, however, Kant clearly holds that the feeling of respect plays a *causal role* in the production of morally worthy action at the phenomenal level, even though he is unclear what this role is. Finally, in the Introduction to the *Doctrine of Virtue*, Kant refines this recognition into a sophisticated theory of the "aesthetic preconditions" of receptivity to duty, or complex causes of moral action at the phenomenal level, and argues that the cultivation of these predispositions is a fundamental feature of what he called, much earlier, "moral praxis."

Chapter 8 confronts a fundamental question for readers of the *Doctrine of Virtue*, namely, what is Kant's conception of virtue? Jeanine Grenberg seeks to understand Kantian virtue indirectly, by asking: what is the *enemy* of virtue? What explains the empirically undeniable fact that becoming virtuous is a struggle, something accomplished in the face of some



opposing force? As Grenberg sees it, if we do not understand what virtue has overcome, we do not really understand the state that results from the struggle. Kant, however, appears contradictory, or at least ambiguous, in identifying the enemy of virtue. He sometimes suggests that we must struggle against our inclinations; yet at other times he suggests that it is a corruption of reason itself that is the true enemy. Grenberg investigates both lines of thought, eventually showing that Kant's apparently contradictory claims in the *Metaphysics of Morals* and *Religion within the Limits of Mere Reason* can be reconciled. Ultimately, she argues, the central connection Kant makes between virtue as strength and inner freedom in the *Metaphysics of Morals* can make sense only if we reject any natural opponents of virtue and admit that the battle for virtue takes place on the territory of reason and freedom. Finally, Grenberg argues that by appeal to Kant's notion of an internal, rational enemy of virtue, we can more clearly distinguish Kantian and Aristotelian virtue.

Chapter 9 considers the primacy of perfect duties to oneself within Kant's moral theory. Kant makes a variety of striking pronouncements about the significance of perfect duties to oneself. But what exactly is the nature of their primacy, and why do they have it? To answer these questions, I explore the *Doctrine of Virtue* account of these duties (as concerned with one's moral self-preservation and moral health), along with two of Kant's earlier accounts from his lectures. In the Collins lecture notes, Kant explicates perfect duties to oneself as prohibiting acting against the necessary conditions of one's greatest, self-consistent use of freedom ("the essential ends of humanity"); in the Vigilantius lecture notes, as duties to which we are directly constrained by humanity in one's person (and thus as immediately grounded in "the right of humanity in our own person"). I show that on all three accounts, perfect duties to oneself bear especially fundamental, vital, and direct relations to freedom, and that these relations generate multiple, interrelated sorts of primacy for these duties.

In chapter 10, Robert Johnson raises and resolves an apparent puzzle about the duty to adopt others' happiness as our end. Because this is a wide and imperfect obligation, no one has a claim on our assistance in advancing her happiness in particular. However, in general, that we have an obligation *to* someone, as opposed to merely *regarding* her, is best understood as her having some claim on us. This apparently generates a puzzle: if we have a duty to others to adopt their end, then it seems that others have a claim on our so doing; but if our duty is wide and imperfect, then no one has a claim on our having her happiness as our end. Johnson shows that the puzzle arises only if we assume that there cannot be a collective

right – the right of collective “others” – and argues that Kant’s followers should not assume this. Johnson begins by exploring Kant’s views on the nature of beings to whom we can owe duties and about what it means to have a duty to a person rather than merely regarding her. Johnson then turns to the relationship between our having duties to a person and the claims she may have on us as a result of such duties. Johnson argues that duties and rights are reciprocal, that there are “collective” rights, and that the reciprocity between rights and duties allows us to distinguish those *to whom* we have duties from those only *with regard to whom* we have duties in the context of Kant’s duties to others. We are then in a position to understand how the wide, imperfect duty to promote others’ happiness is nevertheless genuinely a duty *to* others.

The theme of duties *to* and duties *regarding* others continues in chapter II, which sheds new light on Kant’s provocative insistence that we have duties *regarding*, but not *to*, non-human animals. Patrick Kain confronts questions both interpretive and philosophical about Kant’s account of moral status. Kain shows that a better appreciation of Kant’s commitments in a variety of disciplines reveals that Kant had a deeper understanding of human and non-human animals than commentators generally recognize. This sheds new light on Kant’s claims about the nature and scope of moral status, and helps to address, at least from Kant’s perspective, many of the familiar objections to his notorious account of “duties regarding animals.” According to Kain, Kant’s core principles about the nature of moral obligation structure his thoughts about the moral status of human beings and non-human animals. Through an examination of a broad array of little-studied sources, Kain shows that Kant’s commitments in biology, psychology, anthropology, and physical geography support his account of the nature of and distinction between humans and non-human animals. Kain argues that this account supports Kant’s judgment that we have duties to every human being and significant duties regarding non-human animals, duties which involve direct concern for animals because of their nature. Finally, by comparing Kant’s account with some recently proposed Kantian alternatives, Kain offers us additional perspective on some of the distinctive features, and strengths and weaknesses, of Kant’s approach.

The final chapter of the collection reviews, highlights, and raises questions about themes in Kant’s *Metaphysics of Morals*, especially the *Doctrine of Virtue*. In a wide-ranging discussion, Thomas E. Hill, Jr. comments briefly on how Kant’s normative ethics relates to science, metaphysics, metaethics, and philosophy of law and justice; the relation of

Kant's first principles to more specific moral principles and judgments; the contrast between contemporary and Kantian conceptions of duties to oneself; problems regarding moral negligence, self-deception, and weakness of will; and moral motivation. Hill emphasizes the constraints of law and justice on virtue, the moral (if not epistemological) priority of the first principles of ethics, and the irrelevance of many contemporary objections to Kant's conception of duties to oneself. Hill also highlights Kant's important second-order principles regarding due care in deliberation, self-scrutiny to expose excuses, and strength of will to resist temptations. The morally necessary motive of duty is interpreted, not as an extra duty added onto each particular duty, but as the basic choice to maintain a pervasive attitude that places moral responsibility before self-interest. Hill's exploration provides a fresh, broad perspective on Kant's mature normative ethics. This essay is a fitting one with which to conclude the collection. Partly this is because it revisits – from a different, illuminating angle – a variety of topics touched on in previous chapters. Equally, however, it is because it treats Kant's *Doctrine of Virtue* not simply as a rich, complicated work of practical philosophy, but also as a vibrant, even viable, normative ethics. By doing this, it encourages readers, whether primarily ethicists or Kant scholars, to plunge still more deeply into Kant's *Doctrine of Virtue*, *Metaphysics of Morals*, and moral philosophy as a whole, to discover all they have to offer.

## CHAPTER I

# *Kant's Metaphysics of Morals: the history and significance of its deferral*

*Manfred Kuehn*

### I INTRODUCTION

Kant's *Metaphysics of Morals* appeared in 1797.<sup>1</sup> It was one of Kant's last works. Only two other books appeared later: *The Dispute of the Faculties* and the *Anthropology from a Pragmatic Point of View*, both published in 1798. The *Logic* of 1801 and the *Physical Geography* of 1802 were edited by others, namely Benjamin Jäsche and Friedrich Theodor Rink, on the basis of Kant's lecture notes. It is tempting to view the *Metaphysics of Morals* and the *Anthropology from a Pragmatic Point of View* also as editions of lecture notes. The difference is just that Kant did the editing himself, though his age and the ever-increasing weakness of his mental faculties made this task very difficult. Some have argued that it might almost have been better if someone else had taken over this task for Kant in the case of the *Metaphysics of Morals* and *Anthropology* as well.<sup>2</sup>

Many of Kant's contemporaries felt this way, in any case. Friedrich Schleiermacher wrote a very negative review of the *Anthropology*, finding that "a summary of this book could not be much more than a collection of trivial matters. If, on the other hand, it were intended to give a sketch of the plan and its execution ... it would necessarily give a distinct picture of the most peculiar confusion."<sup>3</sup> Arthur Schopenhauer found that in the

<sup>1</sup> The title was *Die Metaphysik der Sitten in zwey Theilen. Abgefaßt von Immanuel Kant*. Königsberg, by Friedrich Nicolovius, 1797. "Erster Teil: Metaphysische Anfangsgründe der Rechtslehre ... Zweiter Teil: Metaphysische Anfangsgründe der Tugendlehre." One year later, it appeared again, with the title now indicating a second edition of the *Metaphysical First Principles of Right*, "with an Appendix of Explanatory Remarks and Additions." It also appeared separately as *Erläuternde Anmerkungen zu der Rechtslehre von Immanuel Kant* (Königsberg: Friedrich Nicolovius, 1798). A second edition of the *Metaphysics of Morals* was published in 1803.

<sup>2</sup> In a certain sense this has happened: Bernd Ludwig's edition of the *Metaphysics of Morals* (Hamburg: Meiner Verlag, 1986, 1990). Though Ludwig maintains that he is restoring Kant's original text, eliminating corruptions introduced by an incompetent copyist, there is absolutely no evidence that could establish that it was not Kant himself who introduced the mistakes.

<sup>3</sup> See Friedrich Schleiermacher, *Kritische Gesamtausgabe* (Berlin: Walter de Gruyter, 1984), vol. v.i, 365–69. All other quotations from Schleiermacher are from this review.

*Tugendlehre*, “this counterpart of his deplorable *Rechtslehre*, the effects of his weakness brought on by old age are predominant.”<sup>4</sup> Judgments like these could be multiplied. Thus it has been argued that Kant’s practical philosophy ultimately constituted a relapse into precritical dogmatism, which is not significantly different from the theories of his predecessors and contemporaries.<sup>5</sup> Nor are such verdicts without justification. The *Metaphysics of Morals*, like the *Anthropology*, *Logic*, and *Physical Geography* pales in comparison with the three *Critiques*, and it seems to be less critical than it should be. The centrality of the duty of one’s own perfection, for instance, seems to be a throwback to Wolffian ethics, just as the first legal duty of not harming anyone does not significantly depart from Pufendorf’s natural law theory.

<sup>4</sup> Arthur Schopenhauer, *Werke in fünf Bänden*, Ludger Lüdtkehaus (ed.) (Ulm: Haffmanns Verlag, 1988), vol. III, 475. In the “Critique of Kantian Philosophy” (“Kritik der kantischen Philosophie”) which appeared as an Appendix to the *World as Will and Representation*, he found that the *Rechtslehre*, “one of the latest works by Kant ... is so weak ... that it seems to be not the work of this great man, but the product of an ordinary human being and has to die of its own weakness” (*Werke*, vol. 1, 529–626, 607–08). It seemed to him in many places like a “satirical parody” of Kant.

I shall use “*Rechtslehre*” for the first part and “*Tugendlehre*” for the second part of the work. I will also use these terms to refer to Kant’s concerns with law and virtue throughout his various works. The main reason is that I consider the translation “Metaphysical First Principles of Right” (and, in general, the translation of “right” for “*Recht*”) as seriously misleading. The German word “*Recht*” does not mean what “right” means in English. “*Recht*” is much closer to “law” in English. “Natural law” in German means, for instance, “*Naturrecht*,” and a lawyer is a *Rechtsanwalt*, etc. Since the doctrine of rights is only a part of the doctrine of law in English, this way of translating “*Recht*” tends to identify Kant’s doctrine with only a part of law, and thus to confuse the reader. Mary Gregor argues in her “Translator’s Note on the Text of the *Metaphysics of Morals*,” in Immanuel Kant, *Practical Philosophy*, M.J. Gregor (ed. and trans.) (Cambridge University Press, 1996), 358, that “law” would obscure the conceptual ties of “*das Recht*” and “*ein Recht*.” I am not sure that there are any real or deep *conceptual* ties that go beyond the surface meaning of the German. But, however that may be, since these conceptual ties certainly do not exist in English, an English translation should not try to “preserve” them.

<sup>5</sup> This is most often argued with regard to the *Rechtslehre*, but it also concerns the *Tugendlehre*. Thus Christian Ritter, *Der Rechtsgedanke nach den frühen Quellen Kants* (Frankfurt am Main: Vittorio Klostermann, 1971) argued that Kant’s *Rechtslehre* remains essentially precritical and is not part of his transcendental philosophy. But see also Hariolf Oberer, “Zur Frühgeschichte der Kantischen Rechtslehre,” *Kant-Studien* 64 (1973), 88–102, and Werner Busch, *Die Entstehung der kritischen Rechtsphilosophie Kants, 1762–1780* (Berlin: Walter de Gruyter, 1979). Also relevant in this context are Josef Schmucker, *Die Ursprünge der Ethik Kants in seinen vorkritischen Schriften und Reflexionen* (Meisenheim am Glan: A. Hain, 1961), and Karl-Heinz Ilting, “Gibt es eine kritische Ethik und Rechtsphilosophie Kants?” *Archiv für Geschichte der Philosophie* 63 (1983), 325–45. But it is not always clear what the question amounts to, as Ilting, for instance, thinks there is no critical moral philosophy at all, and that even the *Groundwork* and the second *Critique* are non-critical, which seems to me absurd. But it appears to me that he makes a similar mistake when he tries to show that Kant’s categorical imperative depends upon the “principle of law,” as developed by Kant in his precritical work.

I have used the translations of the Cambridge Edition of the Works of Immanuel Kant (CE). In cases where I have found it necessary to change the translations, I have indicated the changes. Where there is no CE translation yet in print, translations are my own.

Still, there can be no doubt that the *Metaphysics of Morals* is the most important of these late works. There are at least three differences between it and the other late works. First, the others are much more closely related to the lectures on which they are based. The *Metaphysics of Morals* has more of a structure of its own. It follows a logic that is different from that of his lectures. While many of its particular doctrines can also be found in the lectures, the purpose is different. Second, it is a work that he had planned for a long time, not an idea that occurred to him relatively late. Third, it is a work that sums up one of the most important concerns of Kant's philosophical thinking as a whole, for morality or, as Kant would have put it, "the moral law within us," is concerned with what for Kant is most important about human beings; without morality we would be just like any other animal. 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 either.<sup>6</sup>

## 2 THE PROTO-HISTORY OF THE *METAPHYSICS OF MORALS*

As has often been noted, however, the project of a "Metaphysics of Morals" goes back to the early days of Kant's philosophical development. He had clearly been concerned with it ever since observing in 1764 in the "Inquiry Concerning the Distinctness of the Principles of Natural Theology and Morality" that "even the fundamental concept of obligation" is inadequately understood and still in need of being properly analyzed, and that "the fundamental principles of morality in their present state are not capable of all the certainty necessary to produce conviction" (AA 2:298).<sup>7</sup> It became Kant's goal to change this unsatisfactory condition.

As early as December 31, 1765 he wrote to Lambert that he would publish "little essays" on the "Metaphysical First Principles of Natural Philosophy" and the "Metaphysical First Principles of Practical Philosophy" before going on to work "on the proper method of metaphysics and thereby also the proper method for philosophy as a whole."

<sup>6</sup> Allen Wood, "The Final Form of Kant's Practical Philosophy," in Mark Timmons (ed.), *Kant's Metaphysics of Morals: Interpretative Essays* (Oxford University Press, 2002), 1–21. See also the part of General Introduction in Kant, *Practical Philosophy* that bears the same title, i.e. xxx–xxxiii.

<sup>7</sup> All references in the text are to Immanuel Kant, *Gesammelte Schriften*, vols. I–XXII, Preussische Akademie der Wissenschaften (ed.); vol. XXIII, Deutsche Akademie der Wissenschaften zu Berlin (ed.); from vol. XXIV, Akademie der Wissenschaften zu Göttingen (ed.) (Berlin: Walter de Gruyter, 1907–). Since the CE includes the pagination of the AA, they can also be checked in the English translation.

Furthermore, “their materials [*Stoff*] lie ready” before him (Br. 10:56).<sup>8</sup> If they had seen the light of day, we would probably have heard much of sentiments and moral taste, and their relation to the theory of Hutcheson and other British writers. We know that these essays did not see the light of day.<sup>9</sup>

Still, it appears that Kant kept pursuing these subjects. Hamann wrote on February 16, 1767 that Kant was working on a “metaphysics of morals.” On May 9, 1768 Kant told his former student Herder that he had finally succeeded in his quest of knowing “the actual nature and the limits of human capacities and inclinations” at least as far as morality is concerned and that he was now working on a “Metaphysics of Morals,” in which he would present the “evident and fruitful principles of conduct and the method that must be employed” in it (HN 19:24). Within a year, if his health permitted, he would be done – or so he thought. Yet he was wrong again.

On September 2, 1770, after writing and defending the *Inaugural Dissertation*, he wrote to Lambert that in the coming winter he would be busy bringing order into his moral philosophy and completing his “investigations of pure moral philosophy, in which no empirical principles are to be found, as it were the Metaphysics of Morals” (Br. 10:97).

Three years later, in a letter to Marcus Herz written toward the end of 1773, he confided that he would be glad when he finished with his “transcendental philosophy, which is actually a critique of pure reason,” as he then could “turn to metaphysics,” which “has only two parts, the metaphysics of nature and the metaphysics of morals.” And he made the further claim that the metaphysics of morals would appear first. He also told Herz that he was very much looking forward to the metaphysics of morals (Br. 10:145). It was almost as if working on this project would be a relief from the critical business. Perhaps it would have been good, but it was not to be. Instead, he found it necessary to keep working on the *Critique of Pure Reason* for another seven years, and other works concerned with transcendental philosophy or the critique of pure reason.

In the *Critique of Pure Reason* itself, he distinguishes between the “speculative and the practical use of pure reason,” and based on this between metaphysics of nature and metaphysics of morals (KrV A841/B869). The metaphysics of morals is to contain all the principles that

<sup>8</sup> I have translated “*Anfangsgründe*” in the two projected works and the three works that appeared consistently as “First Principles” to make clear that the titles are indeed very similar.

<sup>9</sup> The “Remarks in the Observations on the Beautiful and Sublime” (AA 20:3–192) may very well contain the materials for these essays.

“determine action and omission a priori and make them necessary.” It will be “pure morality, which is not grounded on any anthropology (no empirical condition)” (KrV A841/B869). He contrasts metaphysics in general very sharply with critique, which is merely propaedeutic or preparatory, and not really “the system of pure reason.” What he has in mind for the metaphysics of nature at this point probably also holds for the metaphysics of morals: it “will be not half so extensive but will be incomparably richer than this critique, which had first to display the sources and conditions of its possibility” (KrV Axxi), even though he “set[s] it aside” in the *Critique of Pure Reason* because it does not really pertain to its end at this stage (KrV A842/B870). This claim should not be confused with another claim that he makes in the Introduction, where he argues that “the supreme principle of morality and the fundamental concepts of it ... do not belong to transcendental philosophy, since the concept of pleasure ... of desires and inclinations, of choice, etc. which are all of empirical origin, must there be presupposed” (KrV A14/B29). He claims here that the fundamental concepts and principles of morality cannot form part of the propaedeutic or preparatory project or the critical project. In other words, he claims that there will be no *Critique of Pure Practical Reason*.<sup>10</sup>

Kant’s *Groundwork* clearly was not meant as the preparation of a second *Critique*, but rather as an antecedent of the *Metaphysics of Morals*. It was to establish “a completely isolated metaphysics of morals, mixed with no anthropology, theology, physics, or hyperphysics, and still less with occult qualities (which would be called hypophysical)” (G 4:410). In such a metaphysics of morals – Kant insists on using the name, even though he thinks it is “decried” – “moral principles are not based on what is peculiar to human nature but must be fixed a priori by themselves, while from such principles to must be possible to derive practical rules for every rational nature, and accordingly for human nature as well” (G 4:410n), and when he enumerates “a few duties in accordance with the usual division of them into duties to ourselves and to other human beings and into perfect and imperfect duties,” he notes that he reserves “the division of duties entirely for a future *Metaphysics of Morals*,” and that the division he gives in the present context is merely one of convenience (G 4:421). And it is clear that at this point Kant still does not envisage that a *Critique of Pure Practical Reason* would need to be written and that all the preparatory work for the

<sup>10</sup> See also Lewis White Beck, *A Commentary on Kant’s Critique of Practical Reason* (University of Chicago Press, 1960), 9. Beck argues that Kant widened his conception of transcendental philosophy. One might, however, also say that he narrowed his conception of the basic concepts and principles of moral philosophy by eliminating any references to desires, pleasures, etc.



*Metaphysics of Morals* has been done. Indeed, some passages suggest that the step “into metaphysics of morals” is already taken in some parts of the *Groundwork* (G 4:427).

But be that as it may, it is clear that Kant wanted to proceed to the “complete elaboration” of the metaphysics of morals immediately after finishing his *Metaphysical Foundations of Natural Science* in 1785 (Br. 10:407).<sup>11</sup> Yet, it is just as clear that he did not succeed in doing so. Other, more pressing, matters interfered. In the very letter in which he said that he would immediately undertake the completion of the metaphysics of morals, he also agrees to write a review of the second part of Herder’s *Ideen*. There was also the second edition of the first *Critique* (1787). But more importantly, he came to think that he needed to write a *Critique of Practical Reason* (1788).

None of this means that Kant had given up on the *Metaphysics of Morals*. In the second *Critique* itself, he held out hope for “the system of science” that would go beyond the “system of critique” (KpV 5:8). Thus, in a letter to Jung-Stilling of April 7, 1789, he promised that “around the end of the summer I shall begin to work on my ‘Metaphysics of Morals,’” which also means that at this time he still had actually not begun work – or so I would think (HN 23:495). On May 26, 1789, he complained that his health was becoming “progressively worse” in his sixty-sixth year, and of the burden that “the last part of the Critique, namely, that of judgment” constitutes while he is also “working out a *system of metaphysics*, of nature as well of morals, in conformity with those critical demands” (Br. 11:49).

Friends and acquaintances expected the work to appear as early as Easter of 1791. On December 21, 1792 he reported to Erhard of working on the *Metaphysics of Morals*, referring to his discussion under the “heading of *Duties to Oneself*,” which he said he was treating “in a manner quite different from what is customary” (Br. 11:309). What is interesting is that, even though the letter goes on to discuss the social contract and natural law, Kant makes no reference to any discussion of such matters in the book he is working on. This is consistent with the fact that Fichte, who had visited Kant in 1792 and written his *Critique of All Revelation* in Königsberg, referred in letters to Kant to a “*Metaphysics of Morals*.” In a letter of May 12, 1793, Kant tells Fichte that, if he were not already seventy years old and held back by his age, he would probably already be dealing with the problem of revelation in the “planned *Metaphysics of*

<sup>11</sup> Zweig translates “völlige Ausarbeitung” as “full composition” (CE *Correspondence*, 229).

Morals” (Br. II:434). It looks very much as if Kant’s project at this time did not include the *Rechtslehre*, or, at the very least, that he had not begun working on it.

We do not know when Kant actually began the writing of the *Metaphysics of Morals* in earnest. It was probably sometime around 1795. In this context it is of some interest that Kant offered his usual lecture course on metaphysics in the winter semester of 1793–94 for the first and only time under the title of “Metaphysics of Morals or Universal Practical Philosophy in accordance with Baumgarten.”<sup>12</sup> While this new title does not necessarily mean that he was at this point already actively writing the two parts that now make up the *Metaphysics of Morals*, it does show that he more intensively dealt with the matters that make up this work.

It appears to me that we may divide Kant’s deferment of the proposed *Metaphysics of Morals* into three periods, with the first one dating from 1762 to about 1770, the second one from 1770 to 1785, and the third period from 1785 to 1797. In each period there are the pressures that come from the perceived need to finish the most important part of his transcendental philosophy, but also other kinds of interferences. The work on the proper method of metaphysics turned out to be much more time-consuming and demanding than Kant anticipated in 1765. It would lead to the *Inaugural Dissertation, De mundi sensibilis atque intelligibilis forma et principiis* of 1770, the *Critique of Pure Reason* of 1781, and the *Prolegomena* of 1783. And after he had written the *Metaphysical First Principles of Natural Science*, it became unexpectedly necessary to engage in further preparatory critical investigations in moral philosophy, namely the *Groundwork of the Metaphysics of Morals* of 1785 and the *Critique of Practical Reason* of 1788. Other work on theoretical and other matters resulting from the first *Critique* also interfered. And then there was the *Critique of the Power of Judgment* of 1790 and the work connected with it, not to speak of the *Religion within Reason Alone* that made his life difficult after 1794. Between 1785 and 1793 the “complete elaboration of the metaphysics of morals” was prevented by other works that were made more urgent by developments having to do with the reception of his thought, and by reasons of health caused by his advancing age.

<sup>12</sup> Actually: “Metaphysicam morum, sive Philosophiam practicam universalem, una cum Ethica ad compendia Baumgartiana.” The usual title would be something like “Metaphysicam, duce Baumgarteni” or “Metaphysicam praeunte Baumgarten explicat.” See Michael Oberhausen and Riccardo Pozzo (eds.), *Vorlesungsverzeichnisse der Universität Königsberg (1720–1804)*, 2 vols. (Stuttgart-Bad Cannstatt: frommann-holzboog, 1999).

In 1794 the so-called censorship edict made the less than ideal working conditions even more difficult. Kant had to promise that he would not publish on religious topics again, but any *Metaphysics of Morals* had to include a discussion of the relations between morality and religion. Thus he wrote on November 24, 1794 to a publisher that he could not promise a firm date for delivering any specific work, and explained that anything he might write would be problematic in any case, since his subject was “really metaphysics in the widest sense, and as such includes theology, morality (and thus also religion) as well as natural law (including public law [*Staatsrecht*] and international law [*Völkerrecht*]), though only to the extent that reason can address these subjects, but the hand of the censor lies heavily on all of these topics and one cannot be sure that all one’s work in any of those fields will not be rendered futile by a stroke of the censor’s pen” (Br. 11:531). He was hoping for peace and clear new rules about what was and was not admissible in publishing about religious and political matters.<sup>13</sup>

There is every reason to suppose that the final work on morals that appeared in 1797 is as different from the projected work of 1765 as the *Metaphysical First Principles of Natural Science* of 1786 is different from what he thought it would be in 1765.<sup>14</sup> The “system of metaphysics ... of morals, in conformity with ... critical demands” had to be very different from what Kant thought it would be in 1765 or 1770. On the other hand, there can also be little doubt that the *Metaphysics of Morals* was the fulfillment of the early hopes and the intermediate promises. We might, therefore, ask what, if anything, remained the same, and thus would justify the claims of some scholars that the practical philosophy is non- or precritical, and what precisely it was that that changed, and whether it goes to justify the claims of other scholars that it presents the final form of Kant’s critical moral philosophy. I think that the answer to both questions is a qualified yes. What remained the same is the actual content of Kant’s *Rechts-* and *Tugendlehre*. What changed was the perspective from which this content must be viewed according to Kant. This gave rise to tensions in Kant’s system, but they are not tensions that are fatal to Kant’s view, at least not when the content and form of his moral philosophy are properly understood.

<sup>13</sup> Kant is, of course, referring to the French Revolutionary Wars and the First Coalition, which was defeated in 1797. However, the Prussian army took part in the war only between 1792 and 1795. So, “peace” was indeed near.

<sup>14</sup> This has not kept scholars from arguing that there is a serious disanalogy.

3 THE RELATION OF THE *METAPHYSICS OF MORALS*  
TO KANT'S LECTURES

Though Kant found he had to postpone the completion of this project again and again, he clearly thought about it during all that time, if only because he was regularly teaching moral subjects. The lectures that were most relevant for this were those on natural law, on ethics, and on anthropology. While the lectures on metaphysics are important as well, they are less so. Kant lectured on moral philosophy twenty-six times between 1755 and 1793/94, twelve times during his so-called precritical period, and fourteen times during his critical period (1770–1804).<sup>15</sup> While the titles of these lectures varied widely, he seems to have always used the same textbooks, namely Alexander Gottlieb Baumgarten's *Initia philosophiae practicae primae acroamaticae* (Halle 1760) for the first part of the course, and his *Ethica philosophica* (Halle 1751) for the second part.<sup>16</sup>

Kant did not slavishly follow these textbooks in his lectures. Rather, he used them in a fashion similar to the way others had used commonplace books since the sixteenth century. His textbook served mainly as an outline of the discipline to be taught. It suggested the order in which the materials would be discussed in class. Perhaps more importantly, it also provided the “heads of inquiry” of the science in question. That Kant used them in this way is shown by the fact that he employed interleaved copies of textbooks and used both the margins and the separate pages for his reflections on the topics they covered. Accordingly, they were tools both in the preparation of his lectures and in his thinking about the subject matter at hand. Even the briefest look at the Table of Contents of Baumgarten's textbooks shows that they were eminently suited for such purposes. Furthermore, given Baumgarten's laconic style, he did not get as much in the way of independent reflections on these concepts as some other authors would have. This shows why Kant preferred his textbooks to those of more prolix authors.<sup>17</sup>

He was almost inevitably influenced by the textbooks in his thinking about moral matters. Thus Baumgarten's broad division of duties towards

<sup>15</sup> See Emil Arnoldt, *Gesammelte Schriften*, 6 vols., Otto Schöndörffer (ed.) (Berlin: Bruno Cassirer, 1907–09), vol. v, 338, who claims “with certainty” that Kant lectured 28 times on moral philosophy “under various titles,” but had intended to lecture at least 30 times.

<sup>16</sup> The *Initia* is reprinted in AA 19:7–21; the *Ethica philosophica* can be found at AA 27:737–1028.

<sup>17</sup> Karl Vorländer, in his Introduction to Immanuel Kant, *Metaphysik der Sitten*, 2nd edn. (Leipzig: Verlag der Dürr'schen Buchhandlung, 1907), ix–li, xxiii, 28 says that he used it as the “skeleton” for his lectures and notes. Arnoldt, to whom he refers, speaks of the “*Fachwerk der Einteilung*” or the “pigeon-holes of divisions” in this context.

God, towards oneself, and towards others in the *Ethica* had a lasting effect on Kant. Even though he relatively early rejected the idea that we have any duties towards God, the idea that we have duties to ourselves or that there are *officia erga te ipsum* and duties towards others or *officia erga alia* still informs the Doctrine of the Elements of Ethics in the *Metaphysics of Morals*. Even the subdivisions of these two broad divisions are clearly indebted to the *Ethica*. That we must differentiate between duties we have towards ourselves as animal beings as well as moral beings or rational selves is just as much indebted to Baumgarten as are many of the discussions of particular duties that are discussed in this context, like suicide (§252), misuse of one's sexuality or *crimina carnis* (§272–75), lying or *falsiloquium* (§343), avarice (§287), etc.<sup>18</sup>

It is possible to show which parts of the lecture notes taken by Kant's students are about which parts of the *Ethica*. Werner Stark has done this for one of the lectures from the 1770s, for instance, but it would be desirable for all the lectures.<sup>19</sup> And it is also possible to show that there is a relatively large degree of similarity between these notes from different times during the 1770s, 1780s, and 1790s. To be sure, different matters are emphasized at different times and new ideas are introduced, but there is no radical change in the parts of the lectures that concern Baumgarten's *Ethica*. Put differently, the contents of the sections that correspond to what later became the Doctrine of Elements remains relatively static.

The passages of the lectures that correspond to Kant's discussion of the *Initia* are very different. They exhibit large, and often radical differences over time. Thus, in the Herder lectures from the beginning of the 1760s, the discussion of the moral sense or moral feeling is absolutely central. Kant claims that "the moral feeling is unanalyzable, a basic feeling [*Grundgefühl*], the basis of conscience." While he already argues that in morality we only ask "for the formality of what is perfect in free actions," it is "only the consideration of free actions with moral feeling that is intrinsic to conscience" (H 27:5). "My reason can err, my moral feeling only if I mistake what is habitual for what is natural ... my ultimate criterion remains moral feeling" (H 27:6). In the Collins lecture notes from the middle of the 1770s, in which Kant's recent discovery of the

<sup>18</sup> There are also interesting differences. Avarice and lying belong according to Kant among the duties towards oneself, whereas Baumgarten discusses them in the context of duties towards others, but the similarities cannot be overlooked either. Duties towards others are characterized by Baumgarten by the notion of love, while Kant argues they consist in making others happy.

<sup>19</sup> Immanuel Kant, *Vorlesungen zur Moralphilosophie*, Werner Stark (ed.), with Introduction by Manfred Kuehn (Berlin: Walter de Gruyter, 2004), 419–28.

importance of anthropology for the application of moral principles is very apparent, Kant argues by contrast that “there must be a single principle emanating from the ground of our will” (C 27:253).<sup>20</sup> And the principle is already identified with “the moral imperative” that “expresses the goodness of the action in and for itself” and “shows that moral necessitation is categorical and not hypothetical. Moral necessity consists in the absolute goodness of free actions and that is *bonitas moralis*” (C 27:255–6). In the Mrongovius lecture notes from about 1785, we find a discussion of the categorical imperative as “the rule of a will [that is] intrinsically good” (M II 29:607) and in the lectures that came later, we do get essentially the same view as is found in the *Groundwork*.

These shifts in the meaning of the basic principle of morality also have *some* influence on the way the specific virtues are discussed in the lecture notes, but it does not go as far as one might expect. The particular duties and virtues Kant espouses do not really change as a result of the radical changes in his view of the basic principle of morality. The contents of morality, or the actual duties that human beings are said to have, remain more or less the same. What changes is Kant’s critical discussion of them, and this is what makes some of the lectures part of the critical corpus.

Similar things are probably true of the lectures on jurisprudence. The lectures on natural law, always listed as “*jus naturae*,” were always based on Gottfried Achenwall’s *Jus naturae in usum auditorum* (first published in 1763). The only set of lecture notes that has survived follows the outline of Achenwall’s compendium fairly closely. It is from the winter semester of 1785. While Kant seems to have followed the textbook closely, he does often criticize the author on specific points. One thing that makes these lecture notes interesting is that he rejects Achenwall’s definition of law as independent of morality and consistently argues against the consequentialism and eudaimonism present in his work. Again, it is the Introduction that is very interesting because it contains an extensive discussion of the human will, freedom, the distinction between hypothetical and the categorical imperative, which is also called the “unconditional imperative of wisdom” (F 27:1324), the distinction between actions in accordance with duty and from duty, and clearly shows that Kant discussed in this lecture some of the matters that are most characteristic of the *Groundwork*. But some of the basic distinctions of the *Metaphysics of Morals* are already

<sup>20</sup> For his view on morality see C 27:244. See also Manfred Kuehn, Introduction, in Immanuel Kant, *Anthropology from a Pragmatic Point of View*, Robert B. Louden (ed.) (Cambridge University Press, 2006), vii–xxix.

present. Thus we find that “Ethics is practical philosophy of actions in regard to dispositions [*Gesinnungen*]. *Jus* is practical philosophy of actions without regard to disposition. All that possesses obligation belongs to ethics, thus all duties. The law concerns duties and actions that are in accordance with the law and can be coerced. An action is called right when it agrees with the law, virtuous when it is based on respect for the law” (F 27:1327).

This set of lecture notes seems to confirm the conclusion we have reached with regard to the lectures on morals, while the particular legal subject matter is largely dependent on the textbook author (and may therefore be called precritical), the way this content is embedded in a discussion of the first principles of practical philosophy is at the very least on the way to critical philosophy.<sup>21</sup>

It is very tempting to see the discussions of Baumgarten’s *Initia* in Kant’s lectures as closely connected with Kant’s critical project. Put differently, the *Groundwork* and the second *Critique* come out of these discussions. They are in some ways just as closely related to his discussion of the first principles of practical philosophy as the substantive parts of the *Metaphysics of Morals*, the Doctrine of Law and the Doctrine of the Elements of Ethics are much more closely tied to Achenwall’s *Jus naturae* and Baumgarten’s *Ethica*. However, there is clearly one fundamental difference: while the doctrinal parts of the *Metaphysics of Morals* did not radically depart from the textbooks, at least insofar as the content is concerned, the *Groundwork* and the second *Critique* have rather little in common with what Baumgarten taught in the *Initia*.

If this is correct, then we may agree with those German critics who have claimed that Kant’s moral and legal philosophy remained ultimately precritical at least insofar as we admit that the contents of the Doctrine of Law and the Doctrine of the Elements of Ethics are not essentially new with Kant. These parts are not what is characteristic about Kant’s critical ethics. On the other hand, this would also mean that the Introduction to the *Metaphysics of Morals* (MS 6:211–21) as well as the Preface and Introduction to the *Doctrine of Virtue* (MS 6:375–413), which are attempts to make clear how the substantive parts of this work fit in with the critical project as a whole, more or less unequivocally belong to the critical project.<sup>22</sup>

<sup>21</sup> These results could be further supported by a careful discussion of Kant’s reflections on moral and legal subjects, but this is not possible in this context.

<sup>22</sup> I do not want to claim, of course, that the other parts of the *Metaphysics of Morals* are entirely uncritical. It is just that they are much more indebted to the textbooks Kant used.

We may ask what, if any, consequences this has for the question as to whether the *Metaphysics of Morals* is the final form of Kant's practical philosophy. It appears to me that the answer to this question needs to be equally measured. Historically speaking, it is just true that it is the final form Kant gave to his moral philosophy. It is also true that the development of a *Metaphysics of Morals* was Kant's ultimate goal throughout most of his philosophical life. But it is far from clear that what Kant ultimately produced is representative of his best intentions and fits unproblematically with his critical moral philosophy as developed in the *Groundwork* and second *Critique*. I think we need to be careful especially when we evaluate its substantial moral doctrines, such as his views on servants (MS 6:283) or "on defiling oneself by lust" (MS 6:424). But be that as it may, it is clearly more important in this context to take a closer look at how this work is related to the *Groundwork* and the *Critique*.

#### 4 WHAT MAKES THE METAPHYSICS OF MORALS METAPHYSICAL?

In the *Critique of Pure Reason*, Kant defines "metaphysics" as the "system of pure a priori knowledge" which "exhibits in systematic fashion the whole body of philosophical knowledge arising out of pure reason" (KrV A841/B869), and divides it into the "speculative" and the "practical employment of pure reason" or the metaphysics of nature and the metaphysics of morals. By "critique" he understood investigations propaedeutic to both divisions of metaphysics (KrV Axxi). The metaphysics of morals is to contain *all* the principles which "determine action and commission a priori and make them necessary." It will be "pure morality, which is not grounded on any anthropology (or any empirical conditions)" (KrV A841/B869).

In the *Critique of Practical Reason*, however, he declines to give a complete classification of all practical sciences just because "the special determination of duties as human duties, with a view of classifying them, is possible only after the subject of this determination (the human being) is cognized as he is really constituted" (KpV 5:8). The *Critique of Practical Reason*, however, gives an account of the possibility of the principles of practical philosophy "without special reference to human nature" (KpV 5:8). He makes essentially the same claims in the *Groundwork*, in which he radically differentiates moral philosophy from anthropology, claiming that the metaphysics of morals must precede practical anthropology, must be completely a priori, and must therefore be "purified"



or “cleansed” of anything empirical, a posteriori, or belonging to mere anthropology (G 4:388).

As a result, “Virtue” or “*Tugend*” does not play a significant role in Kant’s *Groundwork* and the *Critique of Practical Reason*. Virtue has a special reference to human nature. Indeed, it is the greatest achievement for a specific kind of reason, namely human reason. Kant makes this very clear in the early lectures on moral philosophy, where he argues that any kind of doctrine of virtue cannot capture moral philosophy as a whole because “virtue entails not just *morally good* actions, but at the same time the possibility of the opposite, and thus incorporates an inner struggle, this is therefore too narrow a concept, since we can also ascribe *ethics*, but not virtue (properly speaking) to the angels and to god, for in them there is assuredly holiness but not virtue” (H 27:13). As Kant also says: constant progress toward virtue is “the utmost that finite practical reason can effect” and “virtue itself ... as a naturally acquired ability can never be completed” (KpV 5:33).<sup>23</sup>

For this reason, it is not a topic for pure philosophy, but only one for pure moral philosophy *applied* to human beings. This is why it must play the most central role in the *Metaphysics of Morals*, where it is defined as the strength of resisting “what opposes the moral disposition *within us*” (MS 6:380). What opposes the moral disposition within us are our “sensible inclinations,” which cannot be effectively opposed by other sensible inclinations according to Kant, and therefore require “a moral end ... that must therefore be given a priori, independently of inclinations” (MS 6:381). And it is for this reason, the argument quickly proceeds, that “ethics can also be defined as the system of the ends of pure practical reason” (ZcF 8:381). Indeed, these are ends that are themselves duties.

It is probably unnecessary (and certainly impossible) to enter here into a complete discussion of Kant’s conception of ends that are also duties and the role it plays in the *Metaphysics of Morals*. But it is important to understand that ends that are also duties are absolutely central in Kant’s account of morality. For they are what supposedly gives rise to the two fundamental duties human beings are said to have, namely the duties towards ourselves and duties to others as the ends of one’s own perfection and the happiness of others (MS 6:385). All other particular duties, such as those that he found in Baumgarten’s *Ethica* and other textbooks on moral philosophy, are species of these two basic genera.

<sup>23</sup> In the *Grundlegung* the word appears only ten times. In the second *Critique* Kant addresses the question of the relation of happiness and virtue.

What is even more important to understand is how these ends that are also duties are related to the categorical imperative, and more particularly, which of these is more basic. Given the way that Kant is usually read, i.e. from the point of view of the *Groundwork*, we might suppose that only the categorical imperative can establish the duty that we need to perfect ourselves and that we need to make others happy. And are these two duties not used as the third and fourth examples concerning imperfect duties in the *Groundwork* itself?<sup>24</sup> And does not Kant say there that these are some “actual duties, whose derivation from the one principle cited above [i.e. the categorical imperative] is clear” (G 4:424)?

Well, actually Kant does not say that. What he does say is: “These are a few of the many actual duties, or at least of what we take to be such, whose *division* is clear from the one principle cited above” (G 4:424). The translation is based on an emendation by Hartenstein, who without any argument whatsoever substituted “*Ableitung*” for “*Abteilung*” in his 1838 edition of the *Groundwork*, and the majority of editors have followed him without any good reasons. “*Abteilung aus*” would indeed be an odd expression, but Kant is not using this expression. The sentence must be parsed differently. Kant is actually saying only that the *Abteilung* is clear “from” the principle given before or “*aus dem einigen angegebenen Prinzip.*” Kant introduces the examples in accordance with the “usual division,” says that the real division will follow in the *Metaphysics of Morals*, and then makes a point about what the categorical imperative shows about the “usual division,” namely that there are different kinds of contradictions when perfect and imperfect duties are considered. If this were not enough, he picks up the topic of the “division” of duties in the Introduction of the *Doctrine of Virtue*, claiming that “all the *divisions* of ethics will only have to do with duties of virtue” (MS 6:410).

The emendation is also at odds with the *Metaphysics of Morals*, where Kant finds that the categorical imperative, in which “I abstract from all ends,” shows only that maxims qualify for a possible universal law, but does not actually show that these maxims are universal laws. It is “only a negative principle (not to come into conflict with a law as such),” and there needs to be a law for the maxims of actions that goes “beyond this principle” (MS 6:389). And this is, according to Kant, “only the concept of an *end* that is also a duty, a concept that belongs exclusively

<sup>24</sup> It should perhaps also be noted that the other two examples are actually examples of legal duties. It was (and is) against the law to break contracts, and it was (and in some places still is) illegal to commit suicide.

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).

If we take this seriously, and it is difficult to take it seriously given the weight of the traditional reading, the categorical imperative alone is not sufficient for moral action. We need also ends that are duties.

However, I would argue that there is less here than meets the eye. For it might be argued that Kant's discussion of the categorical imperative in the *Groundwork* in some sense already contains this notion of ends that are duties, for at least one of the three applications of the categorical imperative, which are based upon the categorical imperative as the formula of universalization, namely the formula of the "practical imperative," does already contain "ends" in the required sense of "ends as duties" (G 4:430). This should not be too surprising, since it is, after all, also called the "principle of humanity" (G 4:431). He also speaks of this as the "matter [of all maxims], namely as an end," and finds "that in this respect the formula says that a rational being, as an end by its nature and hence as an end in itself, must in every maxim serve as the limiting condition of all relative and arbitrary ends" (G 4:436). This matter is in the *Groundwork* determined only negatively as well.

The concept of ends that are also duties, as developed in the *Metaphysics of Morals*, is meant to extend this merely limiting condition and to transform it into a positive law. "Hence, if there is an end that is also a duty, the only condition that maxims of actions, as means to ends, must contain is that of qualifying for a possible giving of universal law" (MS 6:389).

Wood, in considering these very same matters, has argued that Kant "overwhelmingly prefers the Formula of Humanity as the formula in terms of which the moral law is to be applied."<sup>25</sup> But it appears to me that he is seriously misconstruing the situation. It is not that Kant justifies "no fewer than nine of the sixteen ethical duties by means of the formula of humanity," but rather that he argues these duties are duties that are at the same time ends. It is true that such arguments have a more or less "natural" connection with the formula of humanity, but they cannot be reduced to an application of it.

Indeed, the categorical imperative and its subformula play a rather subdued role in the entire book. One might get the impression that it is not really needed in the derivation of duties that are also essential ends of embodied rational beings. And one might ask why this is so. In the very

<sup>25</sup> Wood, "The Final Form," 13.

first footnote of the *Doctrine of Virtue* (MS 6:375n) Kant makes a point about the distinction between someone who is merely versed in practical philosophy and a practical philosopher that may easily appear gratuitous. He claims that “someone *versed in practical philosophy* is not for that reason a *practical philosopher* already. A practical philosopher is one who makes *the final end of reason* the principle of his actions and joins with this such knowledge as is necessary for this.” To know *what* it is one’s duty to do is one thing. To transform the duty into “the inner principle of the will” in accordance with the knowledge necessary for it is quite another. And Kant suggests that the first kind of knowledge is “easily stated” because it has to do with “the ends all human beings have by their nature.” Even someone merely acquainted with moral philosophy can know what should be done. The second kind of achievement is more important. Only someone who has attained it deserves the title of “moral philosopher.” The “inner principle of the will” is the realization that “the consciousness of this duty is also the *incentive* to actions” (MS 6:375n).

This way of dividing up the work of morality suggests that the determination of a “mere duty of virtue” does not require mental acrobatics. We all know already what such duties are, at least insofar as we know who we are and what our essential ends consist in. There is no significant philosophical or moral problem here, and Kant contrasts this apparent ease of the determination of what our particular duties are with the deeper and more important problem, namely the one that has to do with understanding “the inner principle of the will.” And only someone who has joined these two things is truly a practical philosopher, or so Kant says.

If we call the first requirement “the principle of the individuation of particular duties” and the second one “the principle of the inner nature of dutiful action,” then Kant says just about as clearly as one might wish that we individuate duties by our essential ends. And I would submit that is what he actually does. He does not explicitly say what the more important second requirement or the “the principle of the inner nature of dutiful action” amounts to, but it is obvious to me that it is another way of stating the categorical imperative. It is the “supreme principle of morality” of the *Groundwork* or the “principle of the will” that makes for the moral worth of the action (G 4:400).

That this reading is plausible can also be shown by the distinction Kant makes between *Tugendverpflichtung* or *obligatio ethica*, on the one hand, and *Tugendpflichten* or *officium ethicum s. virtutis*, on the other. The first term is difficult to translate, and it is rendered in the Cambridge edition as “obligation of virtue,” while the other one is translated as “duties of

virtue.” Perhaps “ethical obligation” or “obligation to be virtuous” would be marginally better. In any case, Kant claims that we have many duties of virtue. Indeed, there are just as many duties of virtue as there are ends that are also duties. But there is only one *Tugendverpflichtung* or obligation to be virtuous, and this obligation precedes any particular duty that we may have. It also precedes any conception of ends and is

the virtuous disposition [*Gesinnung*], the subjective determining ground to fulfill one’s duty, which extends to duties of right as well although they cannot, because of this, be called duties of virtue. – Hence all the *divisions* of ethics have to do only with duties of virtue. Viewed in terms of its formal principle, ethics is the science of how one is under obligation without any regard for any possible external lawgiving. (MS 6:411)

To say that the categorical imperative as the supreme principle of morality has only to do with the “inner principle of the will” or with the principle of dutiful action amounts to saying that the categorical imperative has to do only with *Tugendverpflichtung* and does not go very far in telling us what our actual duties are. It explains what it means to act or will morally or dutifully, leaving unaddressed the question as to what our duties actually are, because that is not really perceived as a problem by Kant. And that is what one would expect from a critical discussion of morality that is “without special reference to human nature” (KpV 5:8), which is what the *Critique of Practical Reason* and the first two parts of the *Groundwork* are purported to be.

## 5 CONCLUSION

There is, of course, a problem that looms large in all of this, and this is the question of Kant’s large and largely essentialist claims about the a priori “system of the ends of pure practical reason.” What justifies us in assuming these? Kant introduces this idea far too quickly and does not sufficiently justify them. It is clear that they are supposed to be very different from anthropological claims that would be based on empirical observations. They are meant to be philosophical or a priori anthropology, which is another subject that Kant hoped to develop, but never did. It would have dealt with the a priori constituents of human nature as necessary for morality and filled the space between pure moral philosophy and empirical anthropology. Whether such a discipline would have been possible in accordance with the principles Kant established in the first *Critique* is highly doubtful.

In other words, I agree with Wood that “it is a mistake to think that rights and juridical duties for Kant rest on the moral imperative ... or that Kantian ethics has no place for ends or virtues.”<sup>26</sup> But this is a very mixed blessing. The unabashed essentialism about human nature present in the *Metaphysics of Morals* may not be enough to qualify “the final form of Kant’s practical philosophy” as precritical, but it shows that the historical Kant cannot have an answer to Nietzsche and others on their own terms. It would make no sense reflectively (or otherwise) to endorse essential ends of human nature or pure practical reason that are also duties, for instance. But then again, such essential ends of pure practical reason themselves may make no sense either.

<sup>26</sup> Wood, “The Final Form,” 20.