

Kant's Commitment to Metaphysics of Morals

L. Nandi Theunissen

Abstract: A definitive feature of Kant's moral philosophy is its rationalism. Kant insists that moral theory, at least at its foundation, cannot take account of empirical facts about human beings and their circumstances in the world. This is the core of Kant's commitment to 'metaphysics of morals', and it is what he sees as his greatest contribution to moral philosophy. The paper clarifies what it means to be committed to metaphysics of morals, why Kant is committed to it, and where he thinks empirical considerations may enter moral theory. The paper examines recent work of contemporary Kantians (Barbara Herman, Allen Wood, and Christine Korsgaard) who argue that there is a central role for empirical considerations in Kant's moral theory. Either these theorists interpret Kant himself as permitting empirical considerations to enter, or they propose to extend Kant's theory so as to allow them to enter. With some qualifications, I argue that these interpretive trends are not supported by the texts, and that the proposed extensions are not plausibly Kantian. Kant's insistence on the exclusion of empirical considerations from the foundations of moral theory is not an incidental feature of his thought which might be modified while the rest remains unchanged. Rather, it is the very centre of his endeavours in moral philosophy. If we disagree with it, I argue, we have grounds for moving to a distinctly different theoretical framework.

1. Introduction

Kant has long had an image as an austere and aprioristic moralist.¹ This is perhaps of little wonder given that his avowed methodological commitment in moral philosophy, as he tells us in the preface to the *Groundwork*, is sharply to separate its rational from its empirical parts, and to work out the foundations of morality with the rational part alone. Since a notion of obligation lies at the heart of Kant's moral theory, Kant will put this as a commitment about the ground or basis of obligation. The ground of obligation is not to be sought in the nature of human beings or their circumstances in the world; the ground of obligation is to be sought a priori, in reason (GW 4:389).² Since obligation has a master principle, the Categorical Imperative, Kant will derive the master principle of morality a priori. So much is the purview of moral philosophy proper, of what Kant calls 'metaphysics of morals', and Kant is doing metaphysics in each of his canonical practical works.³ The remit of empirical ethics, what Kant calls 'practical anthropology', is to account for the material conditions which stand in the way of fulfilling obligations. But anthropology is separate from moral philosophy proper, and is given no systematic treatment by Kant himself.⁴ This anyway is the orthodoxy. But the orthodox image of Kant has been shifting over the past

thirty years or so, and a friendlier, a more humanistic, an altogether more empirically informed, Kant has grown up in its stead.

The shift is broadly coincident with emergent interest in Kant's late work, *The Metaphysics of Morals*.⁵ While it goes over some of the pure foundations from the *Groundwork*, this book is more concretely interested in the form morality takes in human life. For example, where the *Groundwork* had given just a handful of applications of the principle of morality, *The Metaphysics of Morals* exploits very general facts about human beings to provide a fuller picture of our moral duties to ourselves and others. In the book's first part, The Doctrine of Right, the principle of morality is shown to yield externally enforceable laws for public life.⁶ And in The Doctrine of Virtue there is discussion of various topics in moral psychology.⁷ Included there, too, are casuistical asides on concrete matters such as the morality of serving wine at dinner parties (MM 6:428), or of signing a letter 'Your Obedient Servant' when the custom of politeness strains sincerity (MM 6:431). There is practical advice on the moral instruction of young people,⁸ and practical concern for making morality effectual in the life of grownups.⁹ *The Metaphysics of Morals* raises questions, in this way, about the status of applied and empirical considerations in Kant's late work. It has led some commentators to urge that Kant wholly reconfigures the *Groundwork* division between metaphysics and anthropology in his mature writings, so that there is development across the practical oeuvre.¹⁰

This developmental story finds new emphasis among commentators interested in the status of virtue ethical considerations in Kant.¹¹ These commentators argue that *The Metaphysics of Morals* makes room for notions of moral development, virtue, character, and moral perception. This has led to a re-evaluation of the core doctrines standardly associated, for better or worse, with Kantian ethics. To objections against the Kantian model of deliberation—that it is rule-bound and inflexible; or against the 'thinness' of the Kantian person—that it is empirically insubstantial, the appeal to virtue in Kant has looked to yield ready lines of reply.¹² Indeed, the incorporation of virtue ethical considerations has been so pervasive that it has been possible to argue that there is nothing distinctive about virtue ethical, as against Kantian, theories in normative ethics.¹³

At the same time, if from another direction, efforts have been made to fill out Kant's moral anthropology, the empirical part of moral philosophy to which he never gave systematic attention, and commentators have turned to *Anthropology from a Pragmatic Point of View*, the essays on history, and *Religion within the Limits of Reason Alone*.¹⁴ On the basis of these texts some have claimed that far from being a subsidiary part of Kant's moral theory, empirical considerations about human nature actually lie at its foundation.¹⁵ To the question, 'What is Kant's ethics really about?', Allen Wood (1991) has urged that where others would say freedom, or the nature of practical reason, 'I say that Kantian ethics is most fundamentally about human social conflict, its psychological roots and historical meaning' (p. 345). This has led to an air of sea change, even revolution, in Kant interpretation, with commentators urging their peers to give up the aprioristic image of Kant as stereotypical and interpretively unsound.¹⁶

My aim here is to push back against these recent trends. Kant's commitment to metaphysics of morals is underplayed in recent literature. This may be because the moral philosophy is treated independently from the metaphysics generally, with insufficient attention to their relationship.¹⁷ But I will urge that Kant's commitment to metaphysics of morals sets real limitations on interpretation. The commitment is so central that to give up on it is to give up on the Kantian project. That is, I will argue that it is *definitive* of a Kantian position that it excludes empirical considerations from moral foundations.

My motivations here are both interpretive and philosophical (if you will allow the distinction). I would resist the idea that by softening Kant's rationalism, Kant's moral philosophy is made more attractive (qua Kantian moral philosophy).¹⁸ Recent Kantians have sought to offset the idea that Kant's moral philosophy is abstract,¹⁹ ahistorical,²⁰ or unduly concerned with the motive of duty,²¹ where the thought would be that a more concrete, historically situated, and sympathetically motivated theory would be preferable. But a Kantian might make the general point that simplicity and abstractness are significant philosophical virtues, perhaps especially in light of the complexity of moral phenomena. A Kantian might enlist Kant's own observations about the importance of principles in light of our tendencies to rationalise moral demands away when they cease to be convenient (GW 4:390 ff.). A Kantian might say, too, that the abstract form of the moral law is a consequence of the robustness Kant accords moral normativity, in making moral principles akin to principles in logic or mathematics. To conceive of moral demands as robust is a way of taking the domain of value very seriously, and that might be thought a virtue of the Kantian position. Again, that Kant's theory is on the face of it ahistorical might be important. Where more historically situated theories must answer worries about conservatism and relativism, Kant's theory seems fit to stand in critique of local practices and prejudices, including those to which Kant himself gives expression in the writings on history and anthropology.²² Moreover, while it is surely true that Kant is not committed to an opposition between duty and inclination, the extent to which he can and should accommodate motives of sympathy and love should not be overstated.²³ For Kant's characterisation of morality as categorical, as being such that moral demands must be heeded irrespective of inclination, including sympathetic inclination, is a distinctive claim that many moral theories look to recover.²⁴ So while I recognise that there have been fruitful exchanges between, for example, Kantians and virtue ethicists, I would urge caution about accommodating the diverse theoretical commitments of other positions. It is important to appreciate that the most interesting features of the Kantian position are due to its distinctiveness.

I say this at the same time as I am not a Kantian in ethics, and for the reason that Kant is committed to metaphysics of morals.²⁵ In his famous critique of utilitarianism, Bernard Williams wrote that 'forms of utilitarianism which help themselves too liberally to the resources of indirectness lose their utilitarian rationale and end up as vanishingly forms of utilitarianism at all' (Williams 1973: 81). While Williams was no utilitarian, he saw that heavy qualification of the

utilitarian position comes at the cost of losing what is distinctive and interesting about it. I am making the analogous argument about Kantianism. Kant's insistence on the exclusion of empirical considerations from the foundations of moral theory is not an incidental feature of his thought which might be modified while the rest remains unchanged. Rather, it is the very centre of his endeavours in moral philosophy. If we disagree with it we have grounds for moving to a distinctly different theoretical framework.

The structure of the paper is as follows. In section 2, I locate metaphysics of morals within Kant's broader philosophical system. In section 3, I lay out Kant's argument for undertaking metaphysics of morals. In section 4, I clarify the positive role for empirical considerations in Kant's moral philosophy. The first three sections form the basis for evaluating the interpretive trends at issue for the remainder of the paper. Section 5 considers the scope for incorporating virtue-ethical considerations in Kant's moral philosophy, focusing on the work of Barbara Herman. I argue that from the point of view of Kant's systematic commitments, Herman's are possible Kantian moves, though they ought to be resisted on other Kantian grounds. Section 6 examines Allen Wood's attempt to read the anthropology into Kant's core theorising; while section 7 examines Christine Korsgaard's attempt to read empirical considerations into Kant's account of the bindingness of morality. I argue that Wood's and Korsgaard's proposals, whatever their independent philosophical interest, are not plausibly Kantian proposals. I conclude, in section 8, by showing why this is important to notice.

2. Metaphysics of Morals

'Metaphysics' is Kant's term for a branch of philosophy.²⁶ It is a branch of pure, as opposed to empirical, philosophy, and it is the branch of pure philosophy concerned with synthetic a priori truths—substantive truths that are knowable on the basis of reason alone. The task of metaphysics is twofold. It is critically to investigate the capacity of reason to have substantive a priori knowledge, and it is to lay out the substantive truths arising out of pure reason. Call these, for sake of clarity, metaphysics and metaphysics*.²⁷ There is metaphysics of nature and metaphysics of morals, and in both cases, metaphysics must prepare the way for metaphysics*.

The *Groundwork* is primarily a work of metaphysics of morals. Its task is to prepare the ground, or lay the foundations, for the system of duties—the metaphysics of morals*—promised to come later.²⁸ We are told that the task of moral foundations involves 'nothing more than the identification and establishment of the supreme principle of morality' (GW 4:392).²⁹ 'Identification' and 'establishment' of the principle of morality are ways of describing the section divisions of the *Groundwork*. Sections I and II 'identify' the principle of morality, and its 'source', by analysing the deliverances of common human reason on the nature of morality and moral worth. From these deliverances, the Categorical

Imperative is derived as the supreme principle and its 'source' is taken to lie in reason, and more particularly, in autonomy of the will. But it remains to be 'established' that there really is a principle of morality with the source and features analysed from common human reason. So Section III provides a 'deduction' of morality, and the burden of the deduction is to explain how morality can have its source in something radically independent of all human interests, namely, in autonomy of the will. What needs to be shown is that morality can be binding on agents though it does not rest on their antecedent desires or interests. We heed moral commands because they are commands of reason.

Both the analytic task of 'identification' and the synthetic task of 'establishment', are the work of moral foundations, or what comes to the same, metaphysics of morals. At least, that is Kant's wider and for our purposes more important use of the expression. As an aside, I note that Kant sometimes uses 'metaphysics' more narrowly, for example, in the section titles of the *Groundwork*. There 'metaphysics' denotes the identification of autonomy as the 'source' of morality. It is then the work of 'critique' to establish that and how this is so, and to give the limits of the deduction (4:444–45). But this way of dividing up the terminology is of no real concern to us. For the important point is that no part of the foundational project in moral philosophy permits empirical considerations. Kant's distinction between the analytic and synthetic parts of his project does not find a ready analogue in contemporary discussions. For our purposes we can say that Kant is a rationalist in the sense (i) that the principle of morality is a principle of reason, and in the sense (ii) that the principle is binding on us by reason alone.

It is helpful to consider a contrast class—to consider the kind of theory Kant's rationalism excludes. One of the stalking horses of the *Groundwork* is moral theory which takes the task of reason to be that of figuring out what human beings need in order to live well (GW 4:395). This form of Aristotelianism puts an 'empirically conditioned' conception of reason at its basis (CPrR 5:16), empirically conditioned because reason is put in the service of the natural human desire for happiness. One of the tasks of critique in the practical philosophy is to prevent empirically conditioned reason from alone presuming to set the terms for what we should be doing. It is to prevent empirically conditioned reason from claiming 'absolute rule', and from 'expressing itself in demands and commands that go quite beyond its sphere' (CPrR 5:16).³⁰ Empirically conditioned reason is the target of practical critique just as speculative metaphysics is the target of theoretical critique; the difference is that reason shows opposite tendencies in the two cases. In the theoretical case, reason tends to stray beyond experience and claim knowledge of the super-sensible, while in the moral case, reason underestimates the practical reach of pure reason and tends towards empiricism. The reversal is not lost on Kant (CPrR 5:16).³¹

A we might put it, then, Kant's commitment to metaphysics of morals is a commitment to a substantive position in meta-ethics regarding the ground of obligation, or more neutrally, the starting point for ethics. This commitment

excludes theories which incorporate empirical considerations into the starting point, for example, theories which begin with the good for human beings. I have not yet said *why* Kant is committed to metaphysics of morals, just *that* he is and what that commitment comes to. But as I hope the discussion makes clear, the commitment is Kant's big move in moral philosophy. It is time now to say why he thinks it is needed.

3. Kant's Argument for Metaphysics of Morals

Why does Kant think that a metaphysics of morals is needed? Why must moral foundations proceed entirely a priori? Here is the characteristic answer Kant gives us:

[T]hat there must be such a philosophy [metaphysics of morals] is clear of itself from the common idea of duty and of moral laws. Everyone must grant that a law, if it is to hold morally, that is, as a ground of an obligation, must carry with it absolute necessity; that, for example, the command 'thou shalt not lie' does not hold only for human beings, as if other rational beings did not have to heed it, and so with all other moral laws properly so called. (4:389)

The passage begins with the claim that we have an ordinary notion of moral law, and the suggestion is that because we think in terms of law, morality cannot have an empirical foundation—that, as the passage continues:

therefore, the ground of obligation here 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 simply in concepts of pure reason.

The passage gives us the characteristics of moral laws that make for the exclusion of empirical considerations. Kant implies, since he says moral laws hold for all rational beings, that moral laws are universal, and Kant tells us outright that moral laws are necessary. What does necessity amount to in the present context? In general, for Kant, something is necessary if it could not be otherwise—if there are no possible worlds in which it is otherwise for beings with minds like ours.³² While Kant presumably thinks that the principle of morality is the principle of morality in all conceivable worlds that are also moral worlds, the passage seems to give us a different, a further, way of thinking about moral necessity. Here Kant seems to treat necessity as a feature of the law's grip on us; Kant says the law must be *heeded*. Moral requirements are practically necessary in the sense that they are rationally *binding* on agents.³³ Given Kant's background assumptions about the connection between necessity, universality, and the a priori, and between the a priori and reason, it follows that if morality is necessary and universal, it must have its source in reason. In brief, Kant thinks that necessity and universality are hallmarks of the a priori. As Kant puts it in the first *Critique*, universality and necessity are 'marks by means of which we can

securely distinguish a pure cognition from an empirical one' (B3). Since Kant thinks we inquire a priori only into what reason itself contributes (Bxviii), that we have aprioricity shows that the source of that cognition is reason. So Kant's argument for excluding empirical considerations from moral foundations—for undertaking metaphysics—turns on the nature of morality: that morality is such that it is binding on agents (is practically necessary), and holds for all dependent rational beings (with universal scope).³⁴

4. The Positive Role for Empirical Considerations

I have been discussing metaphysics of morals as the preparatory or foundational part of moral philosophy. I have outlined the kinds of questions it enquires after, as well as the constraints that, as a division of pure philosophy, are imposed on answers to those questions. We have seen that it is Kant's conception of morality that drives the exclusion of empirical considerations from its foundations. The question I turn to now is: what is the positive role for empirical considerations in Kant's theory? I ask the question because we need an interpretive basis with which to adjudicate the claims of recent interpreters to the effect that Kant is not the aprioristic philosopher we have taken him to be.

Now it is uncontroversial that empirical considerations enter in 'anthropology'. According to the divisions of philosophy set out in the preface to the *Groundwork* (GW 4:388), anthropology is the empirical counterpart to metaphysics of morals. But, for Kant, anthropology is not strictly speaking a branch of moral philosophy at all. This is presumably because anthropology is not normative but, as we might put it, strategic and psychological. For example, a feasible task for anthropology is to institute what Kant calls an 'aesthetic of morals': the vivid arousal of feelings of disgust and horror in the face of transgressions of the moral law which serve to fortify commitment to duty (MM 6:406).³⁵

So far the picture is this: we have moral foundations which admit nothing empirical, and moral anthropology which depends on empirical facts about human beings, but is non-normative. Does anthropology exhaust the role of empirical considerations in Kant's theory? Indeed, can it? It seems that more would need to be said. For while Kant insists that pure reason must determine the requirements of morality on its own, those requirements must be determinate, ultimately, for human beings. Consider that the principle of morality is to serve a deliberative function. On a standard reading, we are to formulate the maxim of a prospective course of action, and test whether the maxim can be universalised; if not, we ought not will to act.³⁶ Now in specifying our maxim (what we propose to do for some reason), we must specify the circumstances of our action and our motive. That specification would seem to involve empirical content. In this way, the application of the principle of morality incorporates empirical considerations. Unlike anthropology, however, its results are normative: some such particular action is permissible, or must, or must not, be done.

The application of the principle of morality would appear to lie somewhere between foundations and anthropology.³⁷

So indeed, in *The Metaphysics of Morals*, Kant complicates the *Groundwork* divisions by envisaging a part of moral philosophy that is neither metaphysics nor anthropology, but includes empirical content to form an addendum to metaphysics. In fact, as I would put it, there seem to be several layers of addenda to metaphysics, moving from the more abstract to the more concrete, as more empirical content is incorporated. It is in the addendum to metaphysics that the application of the moral principle presumably belongs, and with it, the determination of specific duties. And it is in the addendum to metaphysics that ethics falls into 'casuistry' on matters such as being a responsible host, and being thrifty, on abstaining from flattery, and complaint. We are told that casuistry is not part of systematic philosophy, but is 'woven into ethics in a fragmentary way' (MM 6:411). Kant marks this out, as it were, textually, by placing casuistical remarks in sections that are separate from the body of the text.³⁸ In this intermediate philosophical domain lies the task, also, of working out the duties that govern treatment of human beings of particular rank, age, and sex, of particular states of learnedness, cultivation and health (MM 6:468–9). Again, Kant tells us that this application of the moral law 'cannot be presented as sections of ethics and members of the *division* of a system (which must proceed *a priori* from a rational concept), but can only be appended to the system'. Kant adds that '[y]et even this application belongs to the complete presentation of the system' (MM 6:469). The qualification is presumably meant to signal that the duties to particular individuals, while they cannot be part of metaphysics in incorporating empirical content, still issue in necessary connections between a person's will and the treatment of someone. They have to that extent the kind of normative content that makes them unsuitable for anthropology.³⁹

What emerges is this. There is a sharp distinction between the foundations of morality, on the one hand, and the various dimensions of applying the principle of morality, on the other. The principle of morality is purely rational even if its application may draw on experience.⁴⁰ What the principle commands on a given occasion is pure in the sense that the will is determined, not by inclination, but by the law. However, working out what one's duty is involves applying the principle, and with that, empirical data (which for Kant is neither here nor there morally). In sum, the role for empirical considerations in Kant's moral philosophy is as follows. Empirical considerations enter indisputably outside of metaphysics, for example, in the application of the moral law in concrete situations, and in anthropology. But empirical considerations in no way enter into the identification or establishment of the principle of morality and its 'source'.⁴¹

Finally, recall that the *Groundwork* was to lay the foundations for a metaphysics of morals* to come later. Are we to understand that *The Metaphysics of Morals* is a work of metaphysics of morals*? What would that mean? When Kant introduces the examples of duties in the *Groundwork*, he says they will be systematically set forth in a future metaphysics of morals*. The future task is said to be not so much the elaboration of the examples, but the elaboration of the

divisions to which the examples belong. Kant says, 'It must be noted here that I reserve *the division of duties* [i.e. between perfect and imperfect] entirely for a future *Metaphysics of Morals*, so that the division here stands only as one adopted at my discretion (for the sake of arranging my examples)' (GW 4:421 fn; italics mine). In this way, the task of metaphysics of morals* is not to give an account of particular moral principles for human beings—that would involve empirical content, and metaphysics of morals* is a branch of *metaphysics*—but to make systematic divisions into which particular moral principles fall—divisions between duties to oneself and others, between perfect and imperfect duties, and between duties of 'right' and duties of 'virtue'. *The Metaphysics of Morals* indeed contains metaphysics* so understood, but it is a strangely mixed kind of book. It is a mix of metaphysics as critique, of metaphysics*, and of the various divisions of practical philosophy which lie *outside* of the system of metaphysics strictly speaking, and to which any mention of specifically *human* duties belongs.

5. Virtue-Ethical Lessons from *The Metaphysics of Morals*

Clearly, that empirical considerations enter in the extra-systematic reaches of moral philosophy bears on certain ways of thinking about Kant as an aprioristic moral philosopher, and some Kantians have fought against the charges of apriorism from precisely this point of view. Barbara Herman is a case in point, and I turn now to Herman as issuing the first of three interpretive challenges to Kant's apriorism. To look ahead, I take Herman's intervention to be consistent with the letter of Kant's commitment to metaphysics of morals, though I close by saying why I think it is nevertheless against the spirit of a Kantian proposal.

Herman is interested in the *philosophy of applied* Kantian ethics, or in what Herman (1993b) calls 'middle theory'. Middle theory is so called because 'it lies between the high theory of value and the low theory of applications'; middle theory 'effects the translation of the basic conception of value in the principles of practical rationality into principles that fit the circumstances of human action, judgment and deliberation' (233, 236).⁴² Since Kant's own work in this area is schematic, middle theory 'provides the missing link in a reconstruction of Kantian ethics' (233). It does so by supplying the necessary, supporting conditions for the Categorical Imperative to work as a test for maxims (Herman 1998: 143). Herman has sought to substantiate Kant's claim that the application of the Categorical Imperative requires judgment in light of experience (MM 6:411; GW 4:389). Her proposal is roughly as follows. In order to make use of the a priori principle of morality an agent must accurately formulate the maxim of his action. If he is not to conceive of his situation idiosyncratically, he must, among other things, have some understanding of the moral bearing of his situation. In Herman's terms, he must be able to single out what in his situation calls for 'moral attention'—the 'morally salient' marks and features of his action and circumstances. For Herman that presupposes a capacity for 'moral perception' (what she also calls a 'moral sensitivity'), as well as the possession of moral concepts, and these will need to be acquired as part of a normal upbringing.⁴³

Herman's work in middle theory in this way appeals to notions traditionally of interest to the virtue ethicist.⁴⁴ Middle theory, Herman says, makes room for a notion of character in Kant.⁴⁵ Let me note two qualifications about this move. Firstly, it is important to be clear that there are two ways a notion of character may be invoked. On the one hand, the idea might be that while what matters, morally speaking, is the exercise of practical reason, the exercise of practical reason presupposes the right formation of our desiderative natures.⁴⁶ On the other hand, the idea might be that what matters, morally speaking, is character, where character is understood as an assemblage of qualities or traits (the virtues of courage, temperance, and so on), where these are defined, roughly, as separate dispositions to behave appropriately in relevant situations.⁴⁷ Whether Kant's theory can in principle accommodate a notion of character depends on whether it is invoked in the first or second way. Since the first gives no foundational role to character, it is at least consistent with Kant's systematic commitments. But the second can in no way feature in Kant's theory; in putting an empirical notion at the foundation of moral theory, it is ruled out, not least, by Kant's systematic commitments. ('Not least' because in failing to make moral action action from duty, it cannot be part of a Kantian proposal).⁴⁸

The second qualification is that the notion of character, whether it is employed in the first or second way, is *not* Kant's own. To *have* character, or virtue, for Kant, is just to be the kind of person who commits himself to action from duty; it is to be someone who determines himself to act by thought of the moral law (MM 6:404).⁴⁹ In fact, the idea of virtue in *The Metaphysics of Morals* is explicitly contrasted with the more Aristotelian idea of virtue as 'an *aptitude* and [...] long-standing *habit* of morally good actions acquired by practice' (MM 6:383). A similar contrast is drawn in *Anthropology from a Pragmatic Point of View*. In one sense, character is an assemblage of particular traits, and in another, it is the mark of a principled person—a person who acts from duty. Of the distinction, Kant tells us that 'the first is the distinguishing mark of the human being as a sensible or natural being; the second is the distinguishing mark of the human being as a rational being endowed with freedom. The man of principles, from whom one knows what to expect, not from his instinct, for example, but from his will, has a character' (A 285).⁵⁰ To the extent Kant is interested in character, he is not interested in Aristotelian character.⁵¹

There are philosophical reasons for this. For the Aristotelian style virtue ethicist, a person must have a good upbringing to get it right ethically speaking. But it is important for Kant that this not be so. It is part of Kant's Enlightenment appeal that no matter what one's training and background, everyone is capable of knowing what their duty is, and of doing their duty. This is borne out in Kant's conception of common human reason. The Categorical Imperative is a formalisation of the implicit ideas that reason simply has—that it uses in deliberation and appraisal. Kant does not envisage that while common human reason grasps the principle of morality, it stumbles in applying the principle in concrete situations. Kant does not give voice to these epistemic concerns. On the contrary, he emphasises common human reason's facility with application. 'Here

it would be easy to show how common human reason, with this compass in hand, *knows very well how to distinguish in every case that comes up what is good and what is evil, what is in conformity with duty or contrary to duty . . .*. He goes on: 'We might have assumed in advance that cognizance of what it is incumbent upon everyone to do, and so also to know, would be the affair of every human being, even the most common' (GW 4:404; italics added). It is on account of his conception of common human reason that Kant takes no great theoretical interest in the formulation of maxims. To see an issue here, as Herman sees an issue, involves giving up that conception, and with it, a core Enlightenment commitment. Relatedly, since the formulation of maxims poses no deep difficulty for Kant, the need for Aristotelian style cultivation does not make itself felt here. Where Herman suggests that our formation of practical reason requires the right formation of our desiderative natures, to the extent Kant would see the scope for this kind of formation, he would likely see matters the other way. The thought would be that we must first settle on the right rational scheme—on the right principle—and cultivate our natures in terms of it. So much is suggested by his remarks about the educative function of anthropology (quoted above).

In sum, I would argue that Herman's work in 'middle theory', while it invokes notions that require some care, amends Kant's view in ways that are consistent with the letter of his system. In this Herman may be taken to draw out an important lesson, namely, that Kant permits empirical considerations to enter in the application of the moral law, that is, outside of his system. Still, while the rapprochement between Kantianism and Aristotelianism has the potential to address long-standing objections to the Kantian view, I would express caution about finding too easy a confluence of ideas.

6. A Foundational Role for Human Nature?

In the next two sections I turn to proposals which make what I take to be more controversial amendments. They are controversial in allowing empirical considerations to enter the 'identification' of the principle of morality and its source, on the one hand, and its 'establishment', on the other. It is to the first of these that I now turn.

In an influential paper, Allen Wood (1991: 326–27) claims that ethics is concerned with the conduct of human beings, and accordingly, must be 'based' on knowledge of human nature. Wood asks whether Kant's moral philosophy presents an exception to ethics so defined. It would appear to present an exception, indeed an objection, Wood says, given Kant's commitment to metaphysics of morals. Against appearances, Wood argues that Kant's moral philosophy is 'founded' on an empirical account of human nature, and, as such, presents no exception to his definition of ethics. 'Even in a theory whose fundamental principle is a priori', Wood tells us, 'that account will be based on an empirical account of human nature, on an anthropology' (326; see also 337). Wood takes this finding to license some radical interpretive conclusions: 'To see

Kantian ethics my way is to focus on very different issues from those which have usually occupied Kantian moral philosophers' (345). He goes on:

My aim here is not to decide the issues I have just been raising, but only to legitimize them. They are issues about human nature and its historical destiny, falling entirely outside the scope of a Kantian 'metaphysics of morals', belonging instead to a critical examination of the anthropological foundations of Kantian moral theory. We cannot begin to evaluate these foundations, however, until we have first admitted their existence, and that requires us to overcome some traditional ways of looking at Kantian ethics. (347)

The claims are repeated in Wood (1999), where the interpretation is worked into a book-length study.⁵² Since the claims are radical, and roundly influential, we must look into the arguments. There are two main lines of argument.

The first is as follows. Kant incorporates empirical considerations about human beings in his very formulations of the moral law. Wood says we see this already in the fact that the moral law takes the form of an imperative. In this Kant makes use of a contingent anthropological thesis to the effect that human beings only act rationally when their wills are under constraint (337). More than that, Wood urges that each of the specific formulations of the Categorical Imperative are devised with a view to the constitution of human beings, in particular human tendencies towards conceit, competition and antagonism towards others (337–38 and ff.). So the Formula of Universal Law finds its point given our tendency to make an exception of ourselves; so the Formula of Humanity speaks to our tendency to claim special dignity for ourselves alone; and so on (338). Wood substantiates these claims by appealing to the theory of human nature laid out in his writings on history. The writings on history are contemporaneous with the *Groundwork*, and we should read the intellectual projects as continuous (329).

Now Wood is right that Kant is concerned with the applicability of the law to human beings. And in part Wood makes a point we have met with before (section 4). The moral law is laid down strictly and a priori for human beings, but empirical considerations enter in supplying what is required to translate its requirements to our case. A version of this thought is encapsulated in the general introduction to the *Metaphysics of Morals*:

we shall often have to take as our object the particular *nature* of human beings, which is cognized only by experience, in order to *show* in it what can be inferred from universal moral principles. But this will in no way detract from the purity of these principles or cast doubt on their *a priori* source.—That is to say, in effect, that a metaphysics of morals cannot be based upon anthropology but can still be applied to it (MM 6:217; Cf. also GW 4:389).

What is new is Wood's claim that the formulation of the law as the Categorical Imperative is already an application involving anthropology—it is an application

of the pure form of the law whose content is simply universal validity (338). The issue here is classificatory: does the Categorical Imperative belong in pure foundations, or does it belong in what I have called the extra-systematic addenda to metaphysics? (Note that it would *not* belong in anthropology, for it is clearly normative). I argue that the law formulated as the Categorical Imperative belongs in pure foundations, and in particular, in metaphysics of morals*. The Categorical Imperative is the 'synthetic a priori proposition', as Kant repeatedly calls it (GW 4:420 and ff.), which emerges from the metaphysical investigation and requires deduction; and as we saw, synthetic a priori truths are the currency of metaphysics*.⁵³ As we should expect, Kant does not emphasise anthropological considerations when he formulates the law as an imperative. In the sections Wood refers to, Kant frames the discussion in terms of two ways reason might be related to the will: reason infallibly determines the will or it does not. Reason does not infallibly determine the will in the case of 'dependent rational beings'. The concept of a dependent rational being is the concept of a sensible being whose will is not always in accordance with reason but, because it inclines towards self-interest, must be commanded (GW 4:397; 4:413). As it turns out, this is the kind of will human beings have, but the account is not presented as a piece of anthropology. It is presented as conceptual analysis.⁵⁴

I add that while Wood uses the rich detail of Kant's conception of human nature from the writings on history to illuminate the canonical formulations of the Categorical Imperative, it is striking that Kant does not play up considerations of human nature here. And this is not accidental. For the more we emphasise Kant's interest in human nature—in our tendencies towards exceptionalism, towards claiming dignity for ourselves, towards discordance with one another—the more we are apt to suspect that the pure idea of universal validity is doing no real work for him. The real work of the theory would seem to be that human beings are so constituted that we are prone to self-love, and should not be! Then Kant's rationalism looks like window-dressing for a more basic theory of human nature. This is not a conclusion a Kantian will or should want to accept. It is important that Kant is self-conscious about what belongs within his system and what does not. For example, as we saw, Kant will permit himself casuistical asides in the *Metaphysics of Morals*, but he is explicit that these are not part of the system. An analogous claim should be made of the writings on history.

I turn now to Wood's second line of argument. If we attend carefully to Kant's argument for metaphysics of morals, we will see that it is a practical argument, and rests on empirical claims about the nature of human beings. Kant argues, for example, that human beings are beset by many unruly inclinations, and that only the thought of pure reason, unmixed with empirical incentives and feelings, has an influence on the human heart sufficient to master them. By contrast, a theory which grounds morality in a mix of inclinations and feelings as well as rational concepts causes the human mind to waver about what is required of it (GW 4:410–11). Human beings are also incapable of following examples of virtuous conduct, so they need an a priori principle of morality to live by (GW 4:406).

Moreover, the motive to morality must be the a priori motive of duty because only then can good actions be *reliably* produced (GW 4:390; 4:411). The purity of the moral law (and motive) serves these practical functions, and since the practical functions depend on empirical claims about human beings, Wood concludes: (i) that anthropology drives the need for purity in moral foundations, so that (ii) the basis of morality turns out to be human nature, after all (Wood 1991: 326–7).

What are we to make of (i) and (ii)? In a clear sense, even if (i) is true, (ii) in no way follows. That is, even if we need metaphysics of morals for psychological-cum-motivational reasons, it does not follow that metaphysics of morals is itself anthropological. The fact that we have an instrumental justification to pursue the pure justification of morality does not make the pure justification of morality itself instrumental. So it seems that Wood must be using the notions of ‘foundation’ and ‘basis’ here in a special sense. Indeed, Wood tells us that he means ‘foundations’ in ‘a somewhat larger sense’ than is usually intended, so that in speaking of the foundations of morality he is speaking of the significance of morality and its role in human life (326). Now of course philosophers can define terms as they see fit, but since the notions of ‘foundation’, ‘grounding’, and ‘basis’ are technical terms in Kantian ethics, doing so here is at best highly misleading—all the more so given that Wood explicitly sets himself against the orthodoxy.

Moreover, (i) needs to be qualified. Wood is of course quite right that Kant thinks there are practical reasons for undertaking metaphysics of morals, and right that these rest on claims about the nature of human beings. One might ask how strong the practical arguments are taken on their own terms. As they are given in the *Groundwork*, the claims are more stated than argued for, and invite quick rejoinders. Where Kant says that only the rational purity of the law subdues uncooperative inclinations, one thinks that a theological grounding of morality would do this too. In both cases, success would depend on one’s credence in the foundations. Where Kant says we need an a priori principle because examples of virtuous conduct are too hard to follow and too few, one thinks that this is not an exclusive disjunction. That the motive of duty alone reliably produces good conduct is, as Wood points out, a highly controversial sort of claim (327). Wood helpfully supplies the broader context from ‘Conjectural Beginning of Human History’, and to evaluate the claims in context we would need to think carefully about the philosophical status of the genealogical story given there.

But whatever the merits of these arguments, the argument for metaphysics of morals *cannot* rest on the practical arguments, and it is not meant to. It *cannot* rest on the practical argument because the practical argument depends on facts about the peculiar constitution of human beings. Human beings may have been constituted so that we had no psychological need for pure foundations—as presumably God has no need—but foundations would still have been pure for that. As we saw, morality must hold for all rational beings. Further, the argument for metaphysics of morals is not meant to rest on the practical argument, but on

an argument to which Wood here makes no mention, viz., the argument from the character of the moral law.⁵⁵ In confining his attention to Kant's practical reasons for undertaking metaphysics of morals, Wood leaves out of the account what I take to be the real philosophical impetus behind Kant's project. As I argued above, Kant is not just concerned to undertake metaphysics of morals for anthropological reasons, Kant is concerned to undertake metaphysics of morals because of what he thinks morality *is*.

7. Bindingness without Metaphysics?

I turn now to a final attempt to incorporate empirical content into Kant's account of the foundations of morality, this time into Kant's account of bindingness. While there have been a number of efforts to do this on the part of neo-Kantians, I will be concerned with the efforts, in particular, of Christine Korsgaard. Korsgaard's view has been articulated in a number of important publications over many years.⁵⁶ It is not my concern here to give a detailed account of Korsgaard's view, nor to trace out the relationship between the different versions.⁵⁷ Rather, I want to examine the first big move in Korsgaard's response to what she calls 'the normative question', a response Korsgaard marshals from lines of thought in Kant; and I want to show how Korsgaard's response fails as a Kantian proposal in violating Kant's basic commitment to metaphysics of morals.

The normative question is the question of how moral requirements are binding on us, where that may be understood as the question of how moral requirements are both authoritative and motivating independently of what we may happen to want. This is the question Kant asks in the third part of the *Groundwork*, and in many ways Korsgaard's response is an extended reconstruction of Kant's first moves. Kant argues (i) that in willing to act we cannot but take ourselves to be free agents, able to act on our own reasons; and (ii) that, when fully understood, this entails an ambitious notion of freedom, that of giving the law to ourselves, so that freedom really is autonomy. Kant has already argued for an equivalence between autonomy and morality.⁵⁸ So from a would-be datum about our agency—that we must take ourselves to be free if we are to act all—morality is taken to follow.⁵⁹

For Korsgaard, the normative question makes itself felt from the first person point of view of deliberating over what to do, and in particular, deliberating over whether to satisfy the demands of morality when doing so would come at some cost to the agent. What is wanted is an explanation of our reasons to act in accordance with morality which makes those reasons salient from the agent's point of view. It must be clear how the demands of morality have normative force *for the agent*. Korsgaard begins with our agency because she takes a concern with agency to be given for human beings; our agency is something for which we cannot reasonably ask the question, But why should I be an agent? Since we can guarantee the normative force of our agency, to the extent moral principles

can be shown to be a function of agency, we will guarantee the normative force of moral principles.⁶⁰

In what sense is agency a given concern for human beings? Sometimes Korsgaard writes as if agency is a given concern because agency is inescapable or 'practically necessary' for us. That is to return to the Kantian thought that we cannot act except under the idea of freedom. As Korsgaard (2009) puts it, 'Human beings are *condemned* to choice and action. Maybe you think you can avoid it, by resolutely standing still, refusing to act, refusing to move. But it's no use, for that will be something you have chosen to do, and then you will have acted after all. Choosing not to act makes not acting a kind of action, makes it something that you do' (1.1.1).⁶¹ The problem with this response is that the fact that something is inescapable for us does not guarantee that it has normative significance for us. We may be forced to be agents and yet regard our agency dimly, so that our sense of normative significance is quite disconnected from our situation as actors in the world.⁶² To answer the normative question, Korsgaard needs agency to be an inescapable *normative* concern for us.⁶³

Given this problem, a more promising response may be found in Korsgaard's remarks about practical identity and identification.⁶⁴ Here the thought is not so much that we are constitutively bound to be agents, but that we are constitutively bound to take a normative interest in our agency. We must not only act, but see our actions as extensions of ourselves, and not, moreover, just any old self, but ourselves considered under the auspices of a practical identity that we care about. As Korsgaard takes over themes from Harry Frankfurt (1971), we must act on reasons with which we identify.⁶⁵ And now the thought is that since we care about our reasons, care about being agents, if it can be shown that this care commits us to acting morally, then it can be shown that we really are bound to act as morality requires. Taken this way, Korsgaard's solution to the normative problem is to posit an essential connection between morality and caring, not just any care, but a care that is constitutive of our identity as human beings.

As will be clear, Korsgaard means her proposal to be an extension of Kant's own solution to the normative problem. And there is a way of reading Part III of the *Groundwork* that has Korsgaard's solution coming close to Kant's own. It can look as if Kant is saying something like the following there. We are bound by morality because we must take ourselves to be free, where in doing so we must take ourselves to be members of a noumenal world whose laws are the laws of morality. Since we are mixed beings—since we also have a phenomenal nature—the question is how the laws of the noumenal world are authoritative for the phenomenal world. And then the thought might seem to be that the laws of the noumenal world are regarded as authoritative because as phenomenal beings we *aspire* to being noumenal beings. This reading is encouraged by Kant's confirmation of the deduction of morality in common human reason (GW 4:454). Kant pictures 'even the most hardened scoundrel' as aspiring to be the better person he knows himself capable of being:

This better person, however, he believes himself to be when he transfers himself to the standpoint of a member of the world of understanding, as the idea of freedom, that is, of independence from *determining* causes of the world of sense, constrains him involuntarily to do; and from that standpoint he is conscious of a good will that, by his own acknowledgements constitutes the law for his evil will as a member of the world of sense—a law of whose authority he is cognizant even while he transgresses it. (GW 4:454–5)

Since Kant's deduction in *Groundwork* III is difficult, it is natural to turn to this more commonsensical formulation, so that Kant's solution to the normative problem turns out to be that morality is binding on us because we are motivated by a self-ideal.⁶⁶ That is, morality gets its grip on us in virtue of the desire of our phenomenal self to live up to our better self—to our noumenal nature.

It should be clear from the foregoing, however, that this cannot be Kant's view. It cannot be Kant's view because it rests the explanation of morality on something empirical, namely, the desires of our phenomenal self. For the same reason, whatever the independent merits of Korsgaard's solution to the normative problem, it cannot be Kant's solution. As we saw, Korsgaard thinks there is a practical identity—our identity as human beings—with which we must identify, and from which morality follows. Why *must* we so identify? As we saw, Korsgaard's answer is that our human identity, our identity as moral agents, is a constitutive normative concern for us. As Korsgaard spells out that constitutive normative concern, it is grounded in a basic *human need*. As she puts it in one of several places, 'a human being is an animal who needs a practical conception of her own identity, a conception of who she is which is normative for her' (3.4.9 and ff). Here she takes the communitarian conception of human beings as having a basic need to live in communities as a model (3.4.6–7). It is this which allows her to say that her account of obligation is in a sense naturalistic, for 'it grounds normativity in certain natural—that is, psychological and biological—facts' (4.4.1).⁶⁷ In this she sees her account as of a piece with Kant's: 'Kant, like Hume and Williams, thinks that morality is grounded in human nature, and that moral properties are projections of human dispositions' (3.1.1). It has been my concern in this paper to show that and why this cannot be so.

In the end Kant may not give us a satisfying explanation of categorical bindingness.⁶⁸ In the *Groundwork* he takes an explanation of the interest we take in morality to be beyond the limits of moral inquiry (4:461–2); and in the second *Critique* he faces up to the limits of explanation by telling us that categorical bindingness is simply *a fact of reason*. It is just a given that we take ourselves to be bound by morality. However unsatisfying, unlike the route taken by Korsgaard in positing basic human needs or categorical cares, and unlike the route which explains bindingness in terms of motivation by a self-ideal, to appeal to a fact of reason at least has the advantage of being consistent with Kant's commitment to metaphysics of morals.

8. Conclusion

Let me close by saying that my criticism of Wood's and Korsgaard's proposals as not tenable *Kantian* proposals is no mere matter of terminology. It has been the concern of this paper to argue that Kant's commitment to metaphysics of morals is basic to the kind of project he is engaged in moral philosophy. His commitment to metaphysics of morals follows from the kind of thing he takes morality to be. Now Kant is at his most fervent in those parts of his writing where he insists on a methodological advance over his predecessors—over those who, despite their quibbles with one another, are united by a shared oversight or blunder. In the critical philosophy, that advance consists in limiting the aspirations of pure reason to know by examining the preconditions of knowledge itself. The result is a thorough dressing down of traditional metaphysics—that great site of 'mock conduct' and 'groping among mere concepts' (Bxv)—and, against empiricism, the provision of a role for reason in working out the necessary conditions for experience. In the practical philosophy Kant also dismisses the fantastically ambitious optimism of some forms of rationalism. This shows itself in the fact that morality is in the end an idea of reason, and resistant to proof. But Kant's main target here is not reason's propensity to *transcend* experience, but as we might put it, its *failure* to do so. Kant's particular target are moral theories which, in blending rational and empirical considerations, 'substitute for morality a bastard patched up from limbs of quite diverse ancestry' (GW 4:426). The result is a thorough rebuke to the moral credentials of empirically conditioned practical reason, and a steadfast commitment to the rational purity of moral foundations. In incorporating empirical considerations into the foundations of Kantian ethics, contemporary interpreters violate the very move that Kant takes to be his singular contribution to moral philosophy.⁶⁹

Nandi Theunissen
 Department of Philosophy
 Johns Hopkins University
 USA
 nandi.theunissen@jhu.edu

NOTES

¹ Bernard Williams has been one of the more insistent on this conception of Kant in contemporary discussions. See, for example, the opening of Williams 1976.

² References to all of Kant's works, except the *Critique of Pure Reason*, are to the volume and page numbers of the Preussische Akademie edition, and will be given in the body of the text. References to the *Critique of Pure Reason* will follow the standard method of citing page numbers in the first and second editions (A and B).

³ By the canonical works I mean the Groundwork, the Critique of Practical Reason, and The Metaphysics of Morals.

⁴ This is notwithstanding the publication of *Anthropology from a Pragmatic Point of View*, a record of Kant's notes for his paid lectures on anthropology in the form of a textbook. For an account of the provenance of the work, (see Kuehn 2006: ix and ff.).

⁵ At least in the anglophone world, that interest has coincided with Mary Gregor's translation, the first complete translation since the original English language publication in 1799. For notes on the history of the translation, see the 'Further Reading' in Kant 1996a. For praise of Gregor's translation and a sense of its importance, see Dahlstrom 1993.

⁶ As Sullivan notes in his Introduction to the Cambridge Edition, this is an important datum against the thought that the domain of Kantian morality is individual, personal life (Sullivan 1996: xii). It may also bear on efforts to come to terms with the charge that Kant is socially disinterested, or a moral 'individualist'. Baier puts one version of that charge this way: 'What is of interest for my purposes is his [Kant's] taking of social cooperation and group membership as not of primary moral importance' (Baier 1994: 249).

⁷ These include the relationship between pleasure, desire, and the interests of reason (MM 6:211–14; 6:378), moral emotion (MM 6:399 ff.), and self-governance (MM 6:408 ff.).

⁸ See Sections I and II of the 'Doctrine of the Methods of Ethics'.

⁹ See MM 6:397, 6:408, 6:411, 6:485.

¹⁰ See for example the remarks of Wood 1996: xxx.

¹¹ Among them are Herman 1993a, 2007; O'Neill 1984; Sherman 1997; and Baron 1995.

¹² For an example of a Kantian response to the first kind of charge, see O'Neill 1984. For a forceful statement of the second kind of charge, see Williams 1976.

¹³ This is a central thesis of Nussbaum 1999. Nussbaum argues that to treat virtue ethics as a distinct position from Kantianism (and utilitarianism) involves a category mistake, for all of these theories have substantive conceptions of virtue. Encapsulating many of the ideas in the above paragraph, Nussbaum writes:

In one way, this increasingly popular way of talking [of virtue ethics as a distinct position in ethics] is an obvious category mistake. Immanuel Kant has a theory of virtue, and devotes a good deal of attention to its exposition. Although *The Doctrine of Virtue* was at one time a relatively neglected part of Kant's moral philosophy, read only by specialists, it is now widely discussed, and widely recognised as central. Nobody can any longer think of Kant's view as obsessed with duty and principle to the exclusion of character-formation and the passions. [...] Moreover, the rediscovery of Kant's theory of virtue has also led to serious re-evaluation of the substantive positions of his other ethical writings, as scholars depict a Kant who is less rigorist and more flexible, less concerned with abstract principle and more concerned with the exercise of moral judgment, than the Kant of previous generations (165).

See also Hursthouse's claim regarding the various positions in normative ethics: 'Let us by all means stop caring about how we distinguish ourselves and welcome our agreements' (1999: 7).

¹⁴ See for example the essays on Kant's anthropology edited by Jacobs and Kain 2003, and the monograph by Loudon 2000.

¹⁵ For example, Wood 1991 writes of 'the anthropological foundations' of Kantian moral theory (347), and of Kantian theory as 'deriv[ing] its intellectual power mainly from its anthropological insights' (345).

¹⁶ On giving up stereotypes, see Wood 1999: 10; on the revolutionary gain in so doing, see Wood's description of the aim of his book: to 'help to transform our conception of our

own history and of ourselves as heirs of the Enlightenment. The aspiration of this book is to contribute in some small way to that revolution' (1999: 14).

¹⁷ This has been remarked by Neiman 1994, who writes: 'The excellent and growing body of recent work on Kant's moral philosophy has not devoted sufficient attention to the question why Kant holds that only reason can function as the source of moral principles' (3). Neiman hypothesises that the lack of attention is due to treating Kant's moral philosophy separately from the critical philosophy. Foremost among commentators who have given systematic attention to Kant's notion of metaphysics of morals is O'Neill 1989. See also Gregor 1963: ch. 1. Note that while Gregor's work is cited as an authority by others—by, for example, Loudon 2000 whose opening chapter is on the topic—it is unfortunately out of print. Another good resource on the topic is Beck 1960.

¹⁸ I am grateful to Katja Vogt for discussion on this point.

¹⁹ So Frierson 2003 writes in the Preface to his monograph:

From my earliest exposure to Kant's moral theory, I was drawn to his emphasis on the centrality of freedom but bothered by the apparent abstractness of the moral law. Thus when I first approached the *Anthropology* many years ago, I did so with excitement. I hoped that Kant would incorporate all the rich details of human life that I found lacking in his *Grounding* [...] At first the *Anthropology* seemed more amusing than philosophically satisfying. However, as I came to appreciate the details of Kant's *Anthropology* and as more neo-Kantians incorporated anthropological insights into moral theory, I saw that Kant could provide as rich and concrete a moral theory as anyone (ix).

Frierson's book is an attempt to reconcile the anthropology with Kant's transcendental freedom.

²⁰ So Wood 1999 sees Kant as anticipating the historical materialism of Marx (14, 225, 245). This is not an issue I will pursue in this paper, but for relevant criticism see Bielefeldt 2001: 450.

²¹ This is the position taken in chapter 1 of Wood 1999. For further discussion see Baron 1995, who ultimately defends duty against the many charges against it.

²² It is common for Kantians to wield Kant's theory against his own views about marriage, race or sexuality. See for example some of the moves in Section I of Korsgaard 1992.

²³ The claim that Kant is committed to a problematic opposition between duty and love is memorably encapsulated in Friedrich Schiller's satire of Kant on helping friends in need: 'You must seek to despise them/And do with repugnance what duty bids you' (Schiller's satire is cited in Wood 1999: 28. A contemporary version of the objection may be found in Stocker 1976: 462). For an early defence of Kant against the objection, see Herman 1993a; for a pithy clarification of the issue, see Korsgaard 1997: xiii, fn. 6, and Loudon 1986: 487–8. I agree with Bielefeldt 2001: 449, that Wood overstates the reply to the objection when he claims that for Kant, 'philanthropic love is *an indispensable ground of morality*' (1999: 39; italics in original).

²⁴ Scheffler 1994: 61–72 has a helpful discussion of Kant's challenge to what Scheffler calls 'motivational naturalism'.

²⁵ This is not the place to give my own views in normative theory, but suffice it to say that I share Frankena's contention that 'Morality is made for man, not man for morality' (Frankena 1973: 116). I would reject Kant's argument for metaphysics of morals (see Section 3 below) by denying that morality has the character of *law*—that it holds of necessity and with universal scope. More positively, I would argue that normative theory

should begin by taking a more naturalistic perspective on human beings and our circumstances in the world. (I use the term 'naturalistic' with a view to Railton's discussion of the question 'why be moral?' Railton 1984: Section IX). I take the good for human beings, and other beings, as the starting point for ethics. See my remarks in Section 2 below.

²⁶ His division of the areas of philosophy is prefigured in the first *Critique*, at A841/B869. Kant writes, 'metaphysics' is the name given 'to all of pure philosophy including the critique, in order to comprehend the investigation of everything that can ever be cognised a priori as well as the presentation of that which constitutes a system of pure philosophical cognitions of this kind, but in distinction from all empirical as well as mathematical use of reason'. The division is taken up in the preface to the *Groundwork*.

²⁷ Kant does not always clearly distinguish between them. What I am calling *metaphysics* 'investigates the faculty of reason in regard to all pure a priori cognition', while what I am calling *metaphysics**, lays out the 'philosophical cognition of pure reason in systematic interconnection' (A841/B869). My concern in this paper is with metaphysics of morals, though I remark on metaphysics of morals* in section 4. Guyer 1998: xiv notes that the phrase is used with a 'dual sense'; and Gregor 1963: 4 recognises a more inclusive and a narrower use of the term.

²⁸ On the significance of the metaphors of 'grounding' and 'foundations' for the title, see Guyer 1998: xiii-iv. On this way of reading the title of the *Groundwork*, see Korsgaard 1997: x.

²⁹ I have adapted Gregor's translation here in light of discussion with Jens Timmermann. 'Identify' is better than Gregor's 'search for', because 'Aufsuchung' is a success term.

³⁰ Kant gives a provisional argument against empirically conditioned reason to the effect that reason is inadequate to the task of securing happiness; lacking insight into our needs and what fulfills them, reason actually proliferates our needs and makes us ever more unsatisfied (GW 4:395). I take it that this is a relatively weak argument against empirically conditioned reason, but that the case against it is not meant to rest on it alone. Kant gives a more forceful argument when he characterises morality as necessary and universal, an argument we turn to shortly.

³¹ That pure reason has so positive a role in the practical philosophy may come as a surprise given that it is criticised so severely in the theoretical philosophy. Neiman 1994 thinks that to be surprised about this is to be confused about the nature of reason for Kant. In particular, it is to confuse the aim of reason with the aim of knowledge. Against this Neiman urges that, for Kant, reason is thoroughly practical. Taken this way, Neiman encourages us to think of the first *Critique* as correcting improper demands placed on reason—that it give theoretical answers to speculative questions—so that it may be given full sway, given primacy, in the practical domain.

³² That is, for beings whose intuition is sensible.

³³ Note that it is for beings like us, beings with a mixed rational-sensible nature, for whom the law has the character of bindingness. This is because, unlike angels, and unlike God, subjective conditions may stand in the way of our acting in conformity with reason (GW 4:412). Kant says: 'All imperatives are expressed by an *ought* and indicate by this the relation of an objective law of reason to a will that by its subjective constitution is not necessarily determined by it (a necessitation)' (GW 4:413) See also the preceding passage in the *Groundwork*. God and angels are necessitated to act morally, but in a different way: they cannot but act in accordance with the moral law.

³⁴ Kant expands the Preface argument for metaphysics of morals at GW 4:408; 4:412; 4:414.

³⁵ See also MM 6:217 where anthropology is said to concern ‘the development, spreading, and strengthening of moral principles (in education in schools and in popular instruction)’.

³⁶ There is controversy over whether deliberation is indeed the purpose, or the sole purpose, of the Categorical Imperative. Plausibly, Kant does not intend us to test every maxim of action, though he does intend us to test some. This is the view taken by Louden 1986, but for a different view, to which Louden’s paper is a reply, see O’Neill 1984.

³⁷ For the idea that it is in maxims of action that empirical (situational) content enters Kant’s theory, see O’Neill 1985.

³⁸ On the importance of observing the section divisions in *The Metaphysics of Morals*, the divisions between what is part of systematic philosophy and what is not, see Vogt 2008: 236, 239 and ff.

³⁹ On this point see Gregor 1963: 8.

⁴⁰ It follows that it is not the consequences of a given action that rule out a maxim, but the contradiction that an agent would incur were he to act on it.

⁴¹ I am grateful to Jens Timmermann for discussion on these points.

⁴² Page references are to the reprinted (collected) editions of Herman’s papers.

⁴³ See Herman 1985: 74–8, and Herman 1996: 1 and ff.

⁴⁴ Or at least, to virtue ethicists working in a certain tradition. The tradition is Aristotelian, but in the interpretive vein of Wiggins 1980 and McDowell 1998.

⁴⁵ Herman 1993a tells us that: ‘Kant’s notions of virtue and character are in no way peripheral to the understanding of moral judgment and action. We are able to consider the nature of a Kantian moral agent—what motives, feelings, thoughts, and commitments guide her deliberations and actions. There is then room to develop an account of moral personality that places moral activity within the ongoing practical commitments of a good life’ (x). This position is developed in Herman 1983, but see also Herman 1996.

⁴⁶ See Bk. VI, Ch. 13 of the *Nicomachean Ethics* (Aristotle 2000). McDowell has emphasised the primacy of practical reason, and of virtue as necessary for practical reason, in contemporary discussions.

⁴⁷ The second is the more common conception of the role of character in contemporary virtue ethics. For influential formulations of the position along these lines, see Watson 1990 and Hursthouse 1999: ch. 1. Watson explicitly contrasts this approach with that of McDowell at (466, fn. 7). These theories tend to make little mention of practical reason, or where practical reason is mentioned, it can be difficult to see how it is integrated into the overall picture. This may be because it is hard to see what (fundamental) role is left for character once reason enters the account.

⁴⁸ It is sometimes complained that Kant’s moral philosophy suffers for lack of a role for character of the second kind. See, for example, the references to MacIntyre and Foot in Louden 1986: 473–4. But we might side with O’Neill 1984 in thinking that the point should rather go the other way; that virtue theories which give a role to character of the second kind need to be corrected by the Kantian (and indeed the Aristotelian) emphasis on practical reason.

⁴⁹ As Kant puts it, 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 as an authority *executing* the law’ (MM 6:405).

⁵⁰ I am grateful to Katja Vogt for drawing my attention to this passage.

⁵¹ The distinctiveness of Kant's conception of virtue and character is noted in some of the essays in Betzler 2008. See also Baxley 2010.

⁵² 'Kant's position is grounded on a distinctive theory of human nature and history' (xiii); 'Kant never meant to deny the essential place in ethics of an empirical study of human nature' (10). Despite the continued presence of these claims, the monograph seems to be more circumspect than the early paper. For one, in accordance with the systematic divisions outlined above, the book is divided into two parts, Metaphysical Foundations and Anthropological Applications.

⁵³ As Kant explains in the note to GW 4:420, the Categorical Imperative is synthetic in the sense that it is 'a practical proposition that does not derive the volition of an action analytically from another volition already presupposed (for we have no such perfect will), but connects it immediately with the concept of the will of a rational being as something that is not contained in it'. So unlike a hypothetical imperative, where the imperative is entailed by the end one wills, and so is analytic, for a categorical imperative the will must be connected with the imperative as something *new*. There is clearly more to say about what it means for the Categorical Imperative to be a synthetic a priori proposition.

⁵⁴ Even when Kant goes on to tell us that happiness is an end that may be presupposed in the case of all human beings (and more generally dependent rational beings) he does not take himself to be making an empirical claim. That we seek happiness 'can be presupposed surely and *a priori in the case of every human being, because it belongs to his essence*' (GW 4:415–6 *italics mine*). Presumably this last observation, while a priori is not pure, for presumably the concept of a human beings is drawn from experience. But that does not make it anthropology. On the distinction between pure and a priori propositions, see B3.

⁵⁵ In the monograph, in the course of laying out Kant's foundations properly so called, Wood (1999: ch. 2) describes the argument from the character of morality. That argument is not brought into contact with the earlier account of Kant's practical argument, and it is unclear how Wood now understands their relationship.

⁵⁶ See the essays collected in Korsgaard 1996a, the Tanner lectures published as Korsgaard 1996c, and more recently the Locke lectures published as Korsgaard 2009.

⁵⁷ This has been ably carried out elsewhere. See especially Fitzpatrick 2005, and O'Hagan 2004. In what follows I am mainly concerned with Korsgaard's position as it is articulated in her writings up to and including the *Sources of Normativity*. Korsgaard's later view in *Self-Constitution: Agency, Identity, and Integrity* may deserve separate treatment. But I leave discussion of it largely to one side here.

⁵⁸ This is not intended to be a close textual analysis of the opening of *Groundwork III* (GW 4:446–4:449). For Korsgaard's reconstruction, see Lecture 3 of Korsgaard 1996c.

⁵⁹ Or at least, morality is taken to follow 'analytically'. Kant thinks it still needs to be proven 'synthetically', and that is the concern of the remainder of the third section.

⁶⁰ See Fitzpatrick 2005: section 1.

⁶¹ This idea is given extended treatment in Korsgaard 1996b.

⁶² Enoch 2006 develops this criticism of Korsgaard at some length. As he puts it, Korsgaard's interlocutor may say: ' "Perhaps [. . .] I cannot opt out of the game of agency, but I can certainly play it half-heartedly, indeed under protest, without accepting the aims purportedly constitutive of it as mine" ' (188).

⁶³ See Enoch 2006: 'The kind of necessity the game of agency has to enjoy in order to solve the problem we are now in is *normative necessity*' (188).

⁶⁴ See Lecture 3 of Korsgaard 1996c.

⁶⁵ See Korsgaard 1996c: 99, fn. 8.

⁶⁶ I take this to be the reading of Kant favoured by Velleman 2006.

⁶⁷ For a criticism of Korsgaard on grounds that she gives, like Frankfurt, a metaphysically deflationary account of Kant's notion of the will, see Herman 2002: section III and ff. For critical discussion of Korsgaard's appeal to human nature, see Cohen 1996.

⁶⁸ But see, for example, O'Neill's suggestive postscript in 1989: 64–5. See also Franks 2005: chapter 5.

⁶⁹ This paper grew out of Katja Vogt's graduate seminar on Kant and his critics in the Spring of 2008. I am indebted to Katja for emphasising Kant's commitment to metaphysics of morals in that seminar and in subsequent conversations. I am grateful to Brad Weslake for ongoing discussion and helpful suggestions on an earlier draft. I am indebted to Wolfgang Mann and Jens Timmermann for helpful feedback on an earlier draft, and for discussion on several points. For written feedback I am grateful to David Velleman, Joseph Raz, and for discussion to Pat Kitcher, Jed Lewinsohn, Andreja Novakovic and Katie Gasdaglis.

REFERENCES

- Aristotle (2000), *Nicomachean Ethics*, Second Edition ed. T. Irwin, Indianapolis and Cambridge: Hackett Publishing.
- Baier, A. (1994), *Moral Prejudices: Essays on Ethics*. Cambridge MA: Harvard University Press.
- Baron, M. (1995), *Kantian Ethics Almost Without Apology*. Ithaca and London: Cornell University Press.
- Baxley, A. (2010), *Kant's Theory of Virtue: The Value of Autocracy*. Cambridge: Cambridge University Press.
- Beck, L. W. (1960), *A Commentary on Kant's Critique of Practical Reason*. Chicago: The University of Chicago Press.
- Betzler, M. (2008), *Kant's Ethics of Virtue*. Berlin and New York: Walter de Gruyter.
- Bielefeldt, H. (2001), 'Hegelianizing Kant? Allen W. Wood's Interpretation of Kant's Ethical Thought', *Canadian Journal of Philosophy*, 31: 445–51.
- Cohen, G. A. (1996), 'Reason, Humanity, and the Moral Law', in Korsgaard 1996c: 167–88.
- Dahlstrom, D. (1993), 'The Metaphysics of Morals', *The Review of Metaphysics*, 46: 850–3.
- Enoch, D. (2006), 'Agency Shmagency: Why Normativity Won't Come From What is Constitutive of Action', *The Philosophical Review*, 115: 169–98.
- Fitzpatrick, W. J. (2005), 'The Practical Turn in Ethical Theory: Korsgaard's Constructivism, Realism and the Nature of Normativity', *Ethics*, 115: 651–91.
- Frankena, W. K. (1973), *Ethics*, 2nd edn. Englewood Cliffs: Prentice-Hall.
- Frankfurt, H. (1971), 'Freedom of the Will and the Concept of a Person', *The Journal of Philosophy*, 68: 5–20.
- Franks, P. (2005), *All or Nothing: Systematicity, Transcendental Arguments, and Skepticism in German Idealism*. Cambridge MA: Harvard University Press.
- Frierson, P. (2003), *Freedom and Anthropology in Kant's Moral Philosophy*. Cambridge: Cambridge University Press.
- Gregor, M. J. (1963), *Laws of Freedom: A Study of Kant's Method of Applying the Categorical Imperative in the Metaphysik der Sitten*. New York: Basil Blackwell.

- Guyer, P. (1998), 'Introduction', in P. Guyer (ed.) *Kant's Groundwork of the Metaphysics of Morals: Critical Essays*. Lanham: Rowman and Littlefield.
- Herman, B. (1983), 'Integrity and impartiality', *The Monist*, 66: 233–50. Reprinted in Herman 1993a: 23–44.
- (1985), 'The Practice of Moral Judgment', *Journal of Philosophy*, 82: 414–36. Reprinted in Herman 1993a: 73–93.
- (1993a), *The Practice of Moral Judgment*. Cambridge MA: Harvard University Press.
- (1993b), 'Leaving Deontology Behind', in Herman 1993a: 208–42.
- (1996), 'Making Room for Character', in S. Engstrom and J. Whiting (eds) *Aristotle, Kant, and the Stoics*. Cambridge: Cambridge University Press.
- (1998), 'Training to Autonomy: Kant and the Question of Moral Education', in A. O. Rorty (ed.) *Philosophers on Education: New Historical Perspectives*. London: Routledge. Reprinted in Herman 2007: 130–53.
- (2002), 'Bootstrapping', in S. Buss and L. Overton (eds) *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge MA: MIT Press. Reprinted in Herman 2007: 154–75.
- (2007), *Moral Literacy*. Cambridge MA: Harvard University Press.
- Hursthouse, R. (1999), *On Virtue Ethics*. Oxford and New York: Oxford University Press.
- Jacobs, B. and Kain, P. (2003), *Essays on Kant's Anthropology*. Cambridge: Cambridge University Press.
- Kant, I. (1996a), *The Metaphysics of Morals, Cambridge Texts in the History of Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press.
- (1996b), *Practical Philosophy, Cambridge Edition of the Works of Immanuel Kant*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press.
- (1997a), *Groundwork of the Metaphysics of Morals, Cambridge Texts in the History of Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press.
- (1997b), *Critique of Practical Reason, Cambridge Texts in the History of Philosophy*, trans. and ed. M. J. Gregor. Cambridge: Cambridge University Press.
- (1998), *Critique of Pure Reason, The Cambridge Edition of the Works of Immanuel Kant*, trans. and ed. P. Guyer and A. Wood. Cambridge: Cambridge University Press.
- (2006), *Anthropology from a Pragmatic Point of View, Cambridge Texts in the History of Philosophy*, trans. and ed. R. B. Louden. Cambridge: Cambridge University Press.
- Korsgaard, C. (1992), 'Creating the Kingdom of Ends: Reciprocity and Responsibility in Human Relations', *Philosophical Perspectives*, 6: 305–32. Reprinted in Korsgaard 1996a: 188–224.
- (1996a), *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press.
- (1996b), 'Morality as Freedom', in Korsgaard 1996a: 159–87.
- (1996c), *The Sources of Normativity*. Cambridge: Cambridge University Press.
- (1997), 'Introduction', in I. Kant 1997a: vii–xxx.
- (2009), *Self-Constitution: Agency, Identity and Integrity*. Cambridge: Cambridge University Press.
- Kuehn, M. (2006), 'Introduction', in I. Kant 2006: vii–xxix.
- Louden, R. B. (1986), 'Kant's Virtue Ethics', *Philosophy*, 61: 473–89.
- (2000), *Kant's Impure Ethics: From Rational Beings to Human Beings*. Oxford and New York: Oxford University Press.
- McDowell, J. (1998), *Mind, Value and Reality*. Cambridge MA: Harvard University Press.
- Neiman, S. (1994), *The Unity of Reason*. New York and Oxford: Oxford University Press.

- Nussbaum, M. (1999), 'Virtue Ethics: A Misleading Category?', *The Journal of Ethics*, 3: 163–201.
- O'Hagan, E. (2004), 'Practical Identity and the Constitution of Agency', *The Journal of Value Inquiry*, 38: 49–59.
- O'Neill, O. (1984), 'Kant After Virtue', *Inquiry*, 26: 387–405. Reprinted in O'Neill 1989: 145–62.
- (1985), 'Between Consenting Adults', *Philosophy and Public Affairs*, 14: 252–77.
- (1989), *Constructions of Reason*. Cambridge: Cambridge University Press.
- Railton, P. (1984), 'Alienation, Consequentialism and the Demands of Morality', *Philosophy and Public Affairs*, 13: 134–71.
- Scheffler, S. (1994), *Human Morality*. Oxford and New York: Oxford University Press.
- Sherman, N. (1997), *Making a Necessity of Virtue: Aristotle and Kant on Virtue*. Cambridge: Cambridge University Press.
- Stocker, M. (1976), 'The Schizophrenia of Modern Ethical Theories', *The Journal of Philosophy*, 73: 453–66.
- Sullivan, R. (1996), 'Introduction', in I. Kant in Kant 1996a: vii–xxvi.
- Velleman, D. (2006), 'A Brief Introduction to Kantian Ethics', *Self to Self*. Cambridge: Cambridge University Press.
- Vogt, K. M. (2008), 'Duties to Others: Demands and Limits', in M. Betzler (ed.) *Kant's Ethics of Virtue*. Berlin and New York: Walter de Gruyter, 219–43.
- Watson, G. (1990), 'On the Primacy of Character', in O. Flanagan and A. O. Rorty (eds) *Identity, Character and Morality*. Cambridge MA: MIT Press.
- Wiggins, D. (1980), 'Deliberation and Practical Reason', in A. O. Rorty (ed.) *Essays on Aristotle's Ethics*. Berkeley: University of California Press.
- Williams, B. (1973), 'A Critique of Utilitarianism', in J. J. C. Smart and B. Williams (eds) *Utilitarianism For and Against*. Cambridge: Cambridge University Press.
- (1976), 'Persons, Character, and Morality', in A. O. Rorty (ed.) *The Identities of Persons*. Berkeley and Los Angeles: University of California Press. Reprinted in Williams, B. (1981), *Moral Luck*. Cambridge: Cambridge University Press.
- Wood, A. W. (1991), 'Unsociable Sociability: the Anthropological Basis of Kantian Ethics', *Philosophical Topics*, 19: 325–51.
- (1996), 'General Introduction', in I. Kant (ed.) *Practical Philosophy*. Cambridge: Cambridge University Press, xiii–xxxiii.
- (1999), *Kant's Ethical Thought*. Cambridge: Cambridge University Press.