

THE OXFORD HANDBOOK OF

THE HISTORY
OF ETHICS

Edited by
ROGER CRISP

OXFORD
UNIVERSITY PRESS

CHAPTER 21

KANT'S MORAL PHILOSOPHY

ANDREWS REATH

21.1 INTRODUCTION

KANT's moral philosophy is one of the most significant treatments of moral thought in the history of philosophy and one of the leading models for thinking systematically about morality to come out of the western tradition. It is prominent in contemporary moral philosophy in that many contemporary theorists take themselves to be developing Kant's basic ideas, and many philosophers believe that moral theory can be advanced by interpreting Kant's texts and working out the details of his arguments.

Kant develops a morality of universal principle with, he argues, a basis in reason, so that moral principles present requirements on action that are valid for any rational agent (regardless of their desires and ends). He purports to articulate the underlying principles and structure of ordinary moral thought and to confirm that the authority that it claims for itself has a rational basis. Although Kant's moral philosophy can be approached without taking on all of his transcendental idealism, it is an integral component of his critical philosophy. 'Critique', for Kant, is the critical self-examination of reason aimed at establishing the basic powers and limits of human cognition. In the *Critique of Pure Reason*, Kant undertakes both to establish the legitimacy of the *a priori* concepts and principles that structure theoretical cognition and to define the limits of theoretical knowledge. His project in moral philosophy is similar—to establish the authority of the concepts and principles that we use in ordinary moral thought, and the legitimacy of the conceptions of self and agency that it presupposes. As a component of his critical philosophy, his moral philosophy attempts to show that the moral law is the basic norm of pure reason applied to action.

Kant's moral theory is non-consequentialist, and is best characterized as a 'principle-based' approach to moral reasoning. It holds that certain ways of acting can be right or wrong in themselves independently of their consequences. In particular, the underlying

principle or ‘maxim’ of action must meet a condition of universal validity, and that is ascertained by determining whether the maxim can consistently be willed as universal law for agents with autonomy. This condition of universal validity leads both to inviolable limits on individual choice and the pursuit of good and to positive requirements to have certain ends and attitudes towards others. This general principle and the subsidiary requirements to which it leads express respect for the dignity of each person as a rational agent with autonomy. Good action or ‘good willing’ consists in conformity to principle so understood. Kant’s theory is standardly thought of as a form of deontology, but it differs from the deontology of rational intuitionism in important respects. Rational intuitionism takes certain basic principles of obligation to be self-evident upon reflection. Further, it holds that it is a mistake to seek a justification of the authority of a moral claim outside of the specifically moral form of reflection through which an action is apprehended as obligatory (e.g. an explanation of its obligatory character in terms of some good or value that is realized through an action). (Cf. Prichard 2002.) Kant’s theory supports an explanation of the rightness or wrongness of various ways of acting in value-based terms of respect for persons as rational agents. Furthermore, the basic condition of universal validity itself is justified through its connection with the nature of free agency.

Kant’s moral philosophy is developed principally in three major works: the *Groundwork of the Metaphysics of Morals*, the *Critique of Practical Reason*, and *The Metaphysics of Morals*. In addition, Book I of *Religion within the Boundaries of Mere Reason* is an important treatment of Kant’s moral psychology and theory of free agency, and aspects of his moral thought are filled out in his shorter writings on politics and history (such as *Idea for a Universal History with a Cosmopolitan Aim*, *On the Common Saying: That may be Correct in Theory but it is of No Use in Practice*, *Toward Perpetual Peace*). Further insights into his applied moral thought are provided by his *Lectures on Ethics*—compendia of student lecture notes from the 1770’s through the early 1790’s. Kant gives us both a foundational theory aimed at establishing the authority of ordinary morality and a normative theory—an account of the content of morality and a framework for moral reasoning and judgement. Though his foundational and normative theories are linked, they can be approached separately since one might accept his basic normative theory independently of the ultimate success of his foundational project. I shall provide first an overview of Kant’s foundational theory, and then turn, more briefly, to his normative theory. Since a chapter of this nature cannot discuss, much less resolve, all the issues raised by a theory as complex as Kant’s, I limit myself to an outline of his views, pointing to issues that require further discussion where possible.

21.2 KANT’S FOUNDATIONAL PROJECT

It may help to begin by highlighting the problem that drives Kant’s foundational project and his attempted resolution in the *Groundwork*. Kant describes that work as ‘nothing more than the search for and establishment of the *supreme principle of morality*’ (G 4: 392). This project involves, first, stating the fundamental principle that underlies common-sense

moral thought, and then establishing its authority in reason by connecting it to the nature of rational agency—a task that he here refers to as ‘a pure moral philosophy’ or ‘metaphysics of morals’. The need for such an endeavour becomes evident as soon as we see what goes into the ‘common idea of duty and of moral laws’ (G 4: 389). Ordinary moral thought (Kant believes) takes moral requirements to apply with ‘absolute necessity’ and to apply unconditionally to all rational beings. But requirements can apply with genuine necessity only if they have a basis in reason. Moreover, morality presents us with a set of substantive requirements whose authority does not depend on our actual desires and ends and that take priority over and limit the force of desire-based reasons. It is a set of categorical requirements that are not derived from anything else that we might happen (contingently) to care about or will.¹ But one might ask—as Kant explicitly does in the Second Section of the *Groundwork* (G 4: 419)—how are such categorical practical requirements possible? Are we really required and is it reasonable to give deliberative priority to moral requirements when they conflict with our interest in happiness? Furthermore, if morality presents us with requirements that take priority over desire-based reasons, it presupposes a motivational capacity to act independently of empirically given desires and interests. Do we have such practical and motivational capacities? Once we unpack the ordinary understanding of morality, we encounter the possibility that it presents us with unreasonable demands or is based on illusion. Perhaps there really are no duties or moral requirements as we understand these concepts.

Kant's foundational project is to show that the authority of morality is genuine by connecting the moral law to the nature of free rational volition—or as Kant says, to the ‘form of volition as such’ (G 4: 444). In more technical terms, Kant tries to establish the authority of the moral law by arguing that it is the ‘formal principle’ of free volition, and that as rational agents we necessarily take ourselves to have and identify with a capacity for free agency. To explain: for Kant, the ‘formal principle’ of a domain of cognition or rational activity is the internal constitutive norm that a subject must follow in order to engage in that activity—a principle that defines and makes it possible to engage in that activity. As we might say, it specifies the form of that activity.² For that reason it provides an authoritative norm that tacitly guides all instances of and necessarily governs the activity—one that cannot be rejected by any subject engaged in the activity. The *Groundwork* contains an extended argument for the claim that the moral law is the formal principle of free volition. Furthermore, Kant claims that it is a necessary feature of the self-consciousness of a rational agent that one is free: a rational agent necessarily ‘acts under the idea of freedom’ (G 4: 448). In exercising our wills, we understand ourselves to guide our choices by judgements of what we have good reason to do, and free agency is the capacity to originate action in such judgements—a capacity with which we identify as our ‘proper self’ (G 4: 457–8). In a word, Kant tries to establish the authority of morality by showing that the

¹ Kant thinks that the moral law is a synthetic *a priori* practical proposition: it sets out certain requirements on action, including the requirement to have certain ends and values, without deriving those requirements from any prior volition. Further, these requirements take deliberative priority over empirically given interests.

² For discussions of the idea of internal constitutive principles, see Korsgaard 2008: 7–10, and Reath 2010: 41–8.

moral law is the internal or constitutive principle of a necessary self-conception, of a way in which we necessarily understand ourselves and our practical capacities when we act.

One of the more striking features of Kant's theory is his thesis that, while moral requirements apply unconditionally, the moral law is a principle of autonomy. Moral requirements are not imposed on us by any external authority (such as society, the will of God, or even our own psychology), but are based on a principle that the rational will in some sense legislates for itself. Indeed, his thesis is that the unconditional authority of morality is genuine only if the fundamental principle is a principle of autonomy that the will gives to itself. Understanding the moral law as the formal principle of free agency helps explain why this is so. The formal principle of volition would be the internal norm that both describes and makes possible the exercise of that capacity and serves as its regulative standard, and its normative hold on us is based on our self-conception as free agents. In that sense it is a principle that the rational will gives to itself. And the formal principle of volition (assuming that we can make sense of this idea) is uniquely suited to govern volition unconditionally. Thus, only if the moral law is the formal principle of volition that the will gives to itself could moral requirements apply with genuine rational necessity.

This line of thought leads Kant to claim that previous moral philosophy has approached its task through the wrong methodology. Kant's predecessors in the tradition all base moral obligation in what he terms 'material principles of morality'. Roughly, they ground the normative force of morality in some good that human beings are presumed to desire or take pleasure in as a natural feature of our psychology. The Epicureans (among whom Kant would presumably include Hobbes) appeal to our interest in happiness, Hutcheson's moral sense theory (of which Kant regards Hume's as a variant) bases moral judgement in the immediate satisfaction that we take in observing virtuous conduct, the perfectionism of Leibniz and Wolff bases obligation in the pleasure that we take in a perfection realized in oneself or observed in others, and theological conceptions derive it from an interest in agreeing with the will of God. The common error of these accounts is that they are theories of 'heteronomy'. They ground the authority of morality in an object, or conception of good, external to the will in which we may have an interest, but without establishing that that interest is rationally necessary. By making the normative force of morality conditional on an interest that one can lack without obvious irrationality, they make moral requirements hypothetical rather than categorical imperatives. Kant's insight is that only by showing that the moral law is the formal principle of volition—a principle internally connected to rational volition—can one ground the necessity that is part of our common idea of moral requirement. Moral philosophy must begin by establishing the fundamental law of practical reason. If morality is to be based on any notion of the good, it must be one that is internally connected to rational volition.

A complete grasp of Kant's foundational project is complicated by the fact that the *Critique of Practical Reason* approaches these issues in a somewhat different way. However, the basic point remains: the moral law is the formal principle of free agency, and it is part of our self-consciousness as agents that we have the capacity for free agency. In the next section, I shall sketch the main lines of argument of *Groundwork* I and II. I then briefly discuss the argument of *Groundwork* III, the argument of the Analytic of

the second *Critique* (in particular Kant's claim that the moral law is given as a 'fact of reason'), and some of the apparent differences between these works.

21.2.1 The Argument of the *Groundwork* I and II

As already noted, the project of the *Groundwork* is to 'search for and establish' the fundamental principle of morality that underlies ordinary moral thought, and the argument has both an analytic and a synthetic component. Sections I and II articulate this principle first through analysis of certain judgements of moral value that are central to common-sense moral thought (Section I), then through a more abstract analysis of practical reason or rational agency (Section II). An analysis of a set of concepts and principles can spell out their content, but does not suffice to show that they are legitimately employed. Thus the analytic portion of the argument tells us roughly what goes into our shared conception of morality (e.g. that duties, if there are any, must be expressed as categorical imperatives) but without establishing its rational authority. That task is left to Section III, which is synthetic and provides what Kant calls a 'deduction' of the moral law. This is not a deductive argument in the modern sense, but an argument aimed at establishing the legitimacy and the authority of the principles of ordinary moral thought (similar to the transcendental deduction of the categories of the understanding in the *Critique of Pure Reason*).

Section I of the *Groundwork* is a 'transition from common rational to philosophic moral cognition': it moves from certain tenets of ordinary moral thought—from analysis of widely accepted judgements about the value of a good will—to a philosophical statement of its fundamental principle. A good will is a disposition to adhere to moral requirements for their own sake, and Kant opens Section I with the claim that only a good will is good absolutely and without limitation. A good will is the condition of the value of all other goods (including happiness) in the sense that they are truly worth having only when acquired or used in accordance with the principles of a good will. Further, a good will is good in all circumstances in the sense that it is never worth abandoning your good will for some other kind of good (Cf. Hill 2002a: ch. 2, and 2002b: 23–8). The absolute value of the good will thus amounts to a claim about the deliberative priority of moral considerations over other kinds of reasons. By itself this formal feature appears to tell us little about the fundamental principle of morality. But if one maintains one's good will by conforming to any objective practical principles that, in one's judgement, apply to one's circumstances and would apply to anyone in one's situation—and moreover, out of a direct recognition of their authority—then the absolute value of a good will points to the necessity of 'the conformity of actions as such with universal law' (G 4: 402).

Kant moves closer to a statement of the fundamental principle by developing the concept of good will in imperfectly rational beings subject to non-rational motives—that is, through the notion of duty. He constructs a set of contrasting examples designed to show that we judge an action to have moral worth—to display a good will and thus be

worthy of esteem—only when done from duty. The thesis that an action has moral worth only when done from duty is not intended as a general account of good willing or virtue, but rather is a step in the analysis that identifies the principle on which a good will acts. The examples show that the moral worth of action done from duty lies in its underlying principle—specifically in the fact that the agent is moved by the principle of respect for morality, of conforming to moral principle simply because it is required. Assuming that moral principles are universally valid, the formal principle through which one achieves a good will is then the principle of conforming to universal law as such. Kant thinks that this principle is equivalent to the Universal Law formulation of the Categorical Imperative—to act only from maxims that one can consistently will as universal law (*G* 4: 402). Commentators disagree whether this move on Kant's part is legitimate.³ But clearly what Kant is getting at is that one has a good will when one's choice satisfies a standard of universal validity (say, of rational acceptability to any agent) and one takes this standard to be authoritative. Kant thinks that one ascertains whether a choice meets this standard of universal validity by asking whether its underlying principle or maxim can be willed as universal law without inconsistency. To be universally valid, a maxim of action must be a possible law for all rational agents, and one that cannot consistently be willed as universal law does not meet this standard.

Before turning to Section II, we should note that Kant's analysis of good will has introduced the distinctly moral motive that he terms 'respect' for the morality. To be moved by respect is to be moved by one's immediate recognition of the authority of moral principle, and this motivational capacity is presupposed by ordinary moral thought.

Section I may well show that the so-called Formula of Universal Law (FUL) underlies common moral thought, but that does not mean that it has a basis in reason. Accordingly Section II derives a statement of this principle 'from the universal concept of a rational being as such' (*G* 4: 412)—a project that Kant terms a 'metaphysics of morals.' Indeed, by focusing on different aspects of rational agency, Kant lays out a sequence of alternative formulations of the Categorical Imperative that he takes to be equivalent—'so many formulae of the very same law' (*G* 4: 436). Practical reason or rational agency is the capacity to derive action from the representation of laws, or principles, and this aspect of rational agency leads to the FUL, which Kant regards as the basic statement of the Categorical Imperative—'act only on that maxim through which you can at the same time will that it become a universal law' (*G* 4: 421). But when the principle is stated in this very abstract way it is natural to ask why we ought to adhere. To make headway with this issue, Kant works through the different formulae of the Categorical Imperative, introducing the ideas of humanity as end in itself ('treat humanity, in your own person or in that of another, always as an end in itself and never merely as a means'), the autonomy of the moral agent (agents subject to moral requirements must be regarded as giving moral law to themselves), and the realm of ends (a union of rational beings under shared moral

³ Commentators who question this move on Kant's part include Hill 1992: 121–2 and Wood 1999: 47–9, 81–2. For further discussion and defences see Allison 1996: 143–55, Kerstein 2002, Reath 2006: 204–11, and Engstrom 2009, especially 5–10, 118–28.

principles which is possible through free agency). The sequence of formulas completes the 'search for' the fundamental principle by fully articulating this principle and its presuppositions (e.g. about persons). Like Section I, it is part of an analytic enterprise: it simply shows that the moral law may be expressed in these different ways or presupposes these ideas, while leaving its authority open. But the sequence of formulas also advances the 'establishment' of the moral law in two ways. First, it shows that respect for persons as ends, the autonomy of the individual, and the ideal of moral community exemplified by the 'realm of ends' are implicit in the FUL. Since these are attractive ideals that can command our allegiance, it brings this principle 'closer to intuition... and thereby to feeling' and 'gains access for the moral law' (*G* 4: 436, 437). Second, the sequence of formulas, especially the move from universal law to autonomy, sets up the argument at the opening of Section III that the Categorical Imperative is the formal principle of free volition. This thesis is one key component of Kant's 'deduction' of the moral law.

There is a kind of needle that Kant must thread here. In order to advance the argument in the first of these ways ('gain access' for the moral law), Kant must show that these more substantive values and ideals (of the person, or of moral community) are implicit in or follow from FUL. At the same time, since FUL is a formal principle—the formal principle of moral requirement—and since Kant takes the formulae to be equivalent, the subsequent formulae must likewise be formal principles or expressions of formal features of the moral law. Unless the subsequent formulae are restatements of FUL that draw out what is implicit in it, the sequence cannot advance the overall argument in the desired ways.

Let us now turn to the main formulations of the Categorical Imperative and the movement of the argument through section II.

Kant holds the FUL can be derived from 'the mere concept of a categorical imperative' or 'practical law' (*G* 4: 420). Presumably that concept emerges from his analysis of practical reason or rational agency, though there is some ambiguity as to how. One might think that the idea of practical reason, or of the complete determination of action through reason, leads to a distinction between practical principles that apply conditionally on something else that one wills and unconditional principles that serve as ultimate starting points of practical reasoning. Since the latter would be 'practical laws' or categorical imperatives, once that concept is in hand Kant can move to FUL. Alternatively, Kant could be giving an analysis of familiar forms of practical reasoning that shows that we take moral requirements to apply categorically. Again, once the concept of a categorical imperative is in hand, he can move to FUL. Since Kant's conception of rational agency (at *G* 4: 412) is followed by a discussion of imperatives and the distinction between hypothetical and categorical imperatives, the latter seems to be what he actually does. In the balance of Section II he is interested in what is presupposed by genuine moral requirements, which are taken to be an example of practical laws.

The application of FUL is illustrated through four examples that represent perfect and imperfect duties to self and other. Kant argues that suicide to avoid unhappiness and the deceptive promise cannot be conceived as universal law and are thus impermissible;

they violate perfect duties to self and other. The maxims of neglecting to develop one's natural talents and of indifference to the needs of others cannot be made universal without a contradiction in the will, and this reasoning leads to the imperfect duties of self-development and mutual aid. These examples are not meant to be the basis of a complete normative theory, but rather have the more limited aim of showing that the abstract principle just stated does support some familiar duties and maps onto the distinction between perfect and imperfect duties.

The Formula of Humanity (FH)—that rational nature is to be treated as an end in itself, and never merely as a means—is a principle of respect for persons as rational agents with autonomy. Rational nature or humanity is best understood as the full range of our rational capacities, including the capacity to act for reasons and to set ends, moral personality, and theoretical reason.⁴ An end in itself has an absolute worth that is the source of laws governing its proper treatment. Kant stresses that it is not an end to be produced, but an 'independently existing end' (G 4: 437) that sets inviolable limits on action and choice of subjective ends. The Formula of Humanity is illustrated through the same four examples. Suicide to avoid unhappiness and deceptive promising are inconsistent with respecting humanity as an end in itself—the latter because others must be able to 'contain in themselves the end of the very same action' (G 4:430), i.e. rationally endorse our actions, and the victim of deception cannot. In addition, respecting humanity as an end requires that one's actions 'harmonize' or 'positively agree' with humanity as an end, and this leads to the imperfect duties to make one's own perfection and the happiness of others one's ends—what Kant in the *Metaphysics of Morals* calls the 'duties of virtue'. Kant's claim that FUL and FH are equivalent can be partially explained as follows: to respect persons as ends in themselves is to act in ways that one can justify to the person (as rational agent with autonomy)— i.e. to act from principles that others can rationally endorse—and that one does by acting from maxims that are fully universalizable. That is, one respects humanity as an end in itself by limiting one's maxims of action by the condition of universal validity (G 4: 437–8. Cf. Rawls 2000: 190–5, O'Neill 1989: 126–44, and Korsgaard 1996: 126–8).

Although it is not difficult to interpret FUL and FH in ways that align their practical implications, questions about their equivalence persist in light of the fact that the Formula of Humanity appears to introduce a very different set of concepts not found earlier in the argument ('rational nature' or 'humanity' and 'an end in itself'). One might address these questions by noting that like FUL, FH is introduced through an analysis of practical reason, though one that focuses on a different aspect of practical reason—that it is a faculty of ends. This argument has two steps. First, Kant claims that practical laws presuppose an end of absolute value. Practical reason is the capacity to derive actions from laws—a faculty of principles—but it is also a faculty of ends (Cf. *KpV* 5: 58–9).

⁴ For discussion, see Hill 1992: 38–41, Rawls 2000: 187–90, Korsgaard 1996: 110–14, and Wood 1999: 118–22. Hill and Rawls argue that humanity or rational nature include the full range of our rational capacities, while Korsgaard and Wood focus on remarks where Kant identifies humanity with the capacity to set ends for oneself, including non-moral ends.

Since volition must be directed at some end that it aims to realize, a practical principle must specify an end in order to determine the will (*MdS* 6: 385, 395). So if there are practical laws, there must be an end of absolute value. Here Kant need not mean that the existence of such an end is a further condition that has to be satisfied in order for there to be practical laws. Rather, it is another way to spell out the idea of a practical law. 'End in itself' is the same basic concept as 'practical law', expressed through the lens of practical reason as a faculty of ends, or the need for rational volition to have an end.

The second step of the argument specifies what the end in itself is. An end in itself is a necessary end that limits all subjective or relative ends. But an end can be necessary only if it is internally related to practical reasoning—if it is an end that one has or values simply in so far as one engages in practical reasoning. The obvious candidate for such an end is rational nature: in so far as one engages in practical reasoning, one in some way values that capacity and conformity to its standards. There are different ways of understanding this component of Kant's argument. On one standard reading, it amounts to the claim that a commitment to valuing persons as rational agents is built into the nature of practical reasoning. In so far as one responds to reasons and to what one takes to be of value or sets ends for oneself, one values one's rational capacities, which are capacities essential to one's person. Moreover, one values these capacities on general grounds that commit one to valuing them wherever they are found. That is to say that one is committed to valuing persons as rational agents.⁵

Some commentators have recently proposed a thinner and more formal reading of the claim that rational nature is an end in itself because it enables us to understand *FUL* and *FH* to be equivalent in a deep sense—in effect, as intensionally equivalent. One necessary value that one has *qua* reasoner is that of correctly exercising one's reason. This would seem to be the formal aim of reasoning, in that someone who did not have this aim would not be engaged in practical reasoning. Moreover, this formal aim is authoritative for all exercises of practical reasoning. Accordingly the idea that rational nature is an end in itself can be read as the claim that rational nature has its own proper exercise as its formal end. If *FUL* is the basic norm of practical reason, then one values rational nature as an end in itself by treating this norm as authoritative and by subjecting all practical reasoning to the general condition of universal validity. In order to incorporate the substantive value of respect for persons as rational agents, this reading needs to hold that this value is already implicit in *FUL*—e.g. because it is the principle of acting only from maxims that can serve as universal law for agents with autonomy. It rules out maxims that if universalized undermine rational autonomy, and requires ends that are needed to support its effective exercise in individuals. In that sense, the proper exercise of practical reason will be guided by principles of respect for rational autonomy. According to this line of thought, the moral standing of persons is not a ground of *FUL* but one of its implications: persons have value because *FUL* prescribes principles that respect their rational capacities.⁶

⁵ See Korsgaard 1996: ch. 4; Hill 1992: 143–6; and Wood 1999: 124–31.

⁶ See Herman 2007: 251–2, Engstrom 2009: 167–78, Reath 2012 and 2013, and Sensen 2009 and 2010.

Following the discussion of humanity as an end in itself, Kant introduces the idea of autonomy through a third version of the Categorical Imperative, stated incompletely as 'the idea of the will of every rational being as a will giving universal law' (G 4: 431). The introduction of autonomy marks a turning point in the argument. Until this point Kant has stressed the deliberative priority of moral requirements and subjection to duty. He now claims that the human will is not just subject to duty, but 'subject to it in such a way that it must be viewed also as giving law to itself and just because of this as subject to the law (of which it can regard itself as author)' (G 4: 431). The basis of the authority of morality is that the fundamental principle of morality is a law that the rational will in some sense legislates for itself—which is to say that it is a principle that originates in the very nature of rational volition (G 4: 432–3, 440). Moral agents thus have autonomy in the sense that they are not bound to any external authority, but only by principles that originate in their own rational will, which 'in accordance with nature's end is a will giving universal law' (G 4: 432).

These claims about autonomy appear to follow from the necessity of moral requirements. Kant's idea is that the necessity and deliberative priority of morality is genuine only if based in a principle that the rational will in some sense gives to itself, or that originates *a priori* in the nature of rational volition. Earlier in this chapter we sketched one way to understand this idea—that the moral law is the 'formal principle' or internal constitutive norm of rational volition and is thus uniquely suited to govern volition unconditionally. The arguments of *Groundwork* II have shown that the moral law is a principle that is based in the nature of rational volition by deriving statements of the Categorical Imperative from an analysis of practical reason or rational agency, and these claims figure in the opening argument of *Groundwork* III. Conversely, these claims about autonomy imply the failure of heteronomy: foundational theories that base morality in some object external to the will present moral requirements as hypothetical, rather than categorical imperatives, because they make its authority conditional on having an interest in that object. Thus no such foundational account provides adequate grounding for the common idea of moral requirements.

If the moral subject is a kind of autonomous legislator, then persons interact with others possessing the same legislative capacities as themselves and the universal laws that they are presumed to will are to govern a community of such agents. In this way, the idea of autonomy leads to the 'very fruitful concept' of the 'realm of ends', which completes the (by Kant's count) third version of the Categorical Imperative—that one is to act from principles that can serve as law for a realm of ends (G 4: 433, 438ff.). The realm of ends is Kant's ideal of moral community based on universally valid principles and relations of mutual respect. Membership is determined by possession of the basic capacities of reason, which confers the capacity for a good will and a 'share... in the giving of universal law' (G 4: 435). Among other things, it would be a social world in which individuals pursue their own ends within the limits of respect for the rights of others, and in addition are moved by the ends of virtue. The realm of ends is an ideal of social relations based on mutual respect and mutual recognition of persons as agents with autonomy.

21.2.2 *Groundwork* III and the *Critique of Practical Reason*

The analytic enterprise of *Groundwork* II articulates the basic principle of common-sense morality and the conceptions of agency that it presupposes, but does not establish its rational authority for us. For example, it may be a conceptual truth that agents subject to moral requirements have autonomy and thus have certain kinds of rational capacities, but that does not mean that we are such agents. To complete the overall argument for the authority of morality, *Groundwork* III offers a 'deduction' of the moral law. Unfortunately, it is the most obscure section of the work, and we can only sketch the main points, ignoring many complications.⁷ Furthermore, the second *Critique* offers a somewhat different approach to the authority of the moral law.

Groundwork III opens with two preliminary analytical arguments (*G* 4: 446–448). The first claims that the moral law is the basic principle of a free will, the second that a rational being necessarily acts under the idea of freedom and is entitled to regard itself as free. The first argument turns on the idea that while a free will is a capacity to act independently of external influence, it is not for that reason lawless. Its activity must be governed by some principle, and moreover by a principle that the will gives to itself, independently of external influence. But *Groundwork* II has argued that the moral law is based in the nature of rational agency and in that sense is the law that the rational will gives to itself. Since it is the principle that originates in the nature of volition *a priori*, choice guided by this principle is fully self-determined. The force of the argument from freedom to morality is that the moral law is the principle that confers the capacity for free agency and through which it is exercised, and is thus its formal principle.

The idea of freedom in the second argument can be interpreted as a set of general presuppositions about rational deliberation—for example, that in engaging in deliberation, one presupposes that one can determine oneself by one's own judgement based solely on one's grasp of the reasons, and so on—that is to say, independently of influence external to reason. The necessity of *acting* under the idea of freedom applies these presuppositions to practical deliberation. In deliberating about how to act, one presupposes that one has the capacity to set desire or any potential consideration aside if one sees reason to, to judge what one has most reason to do, and to determine one's choice by this judgement. In short it is a necessary feature of the self-consciousness of a rational agent that one takes oneself to have the capacity to determine choice independently of any influences external to one's reasoning.

If a rational agent necessarily acts under the idea of freedom and the moral law is the basic principle of a free will, then any rational agent is bound by the moral law. But these are still analytical arguments and do not show that we have the relevant rational capacities. Kant needs to establish something about us, and the task that he takes on is to show that we have purely rational capacities that make us members of an 'intelligible world'.

⁷ More complete discussions of the argument of *Groundwork* III are found in Hill 1992: ch. 6 and 2002b: 93–108, 132–44; Henrich 1998; Korsgaard 1996: ch. 6; Allison 1990: chs 11 and 12; Ameriks 2003: chs 6 and 9; Schönecker 2006; Timmermann 2007: 120–51; and Guyer 2009.

An intelligible world is a world governed by rational norms of thought and action. To be a member of such a world is to be subject to such norms, and to have a capacity for spontaneity, or norm-guided self-determining activity that is independent of sensibility and cannot be understood in terms of naturalistic causal laws. If we have such intelligible capacities, we are warranted in ascribing free agency to ourselves. Our capacity to take an interest in and be moved by respect for moral principles appears to be such a capacity. But to appeal to ordinary moral consciousness at this point (to establish that we are free and thus subject to morality) would be to argue in a circle (*G* 4: 449–450). Instead Kant appeals to features of theoretical reasoning—that it constructs ideas that go beyond anything given in sensibility, that it marks out the limits of the understanding, and that it generates regulative ideals to guide the employment of the understanding. The capacity for theoretical reason displays a spontaneity and independence of sensibility that makes us ‘members of an intelligible world’ and thus provides non-moral grounds for ascribing freedom to ourselves (*G* 4: 450–453).

This line of thought is intended to show that we human beings are members of an intelligible world who necessarily act under the idea of freedom and are warranted in ascribing free agency to ourselves. But we are still sensible creatures with desires and an interest in happiness that can conflict with the demands of morality. The deduction proper (*G* 4: 453–5) appears to rest on the claim that because we identify with our intelligible self as our proper self (*G* 4: 457, 461)—perhaps because it is our rational capacities that constitute the self—its law has full and immediate authority in cases of conflict with our sensible interests. To summarize the overall trajectory of the argument: Kant argues for the necessary authority of morality by arguing that the moral law is the formal principle of free agency and that we necessarily act under the idea of freedom and identify with our free rational capacities as our proper self.⁸

A standing problem for Kant’s claim that the moral law is the basic principle of free agency is how choice can be free when it is not guided by moral principles—for example, when a choice is guided by a maxim that is desire-based or that is contrary to morality. Clearly Kant should and means to hold that we can freely act on non-moral grounds and are responsible for such choices, and in other works he makes it clear that desires or incentives never cause choice directly, but lead to choice only through an act of spontaneity on the part of an agent.⁹ One way to address this concern is to hold that the capacity to act from moral principles is sufficient for free choice, whether or not it is fully exercised. An agent with that capacity who acts on his own judgement acts freely, even when he acts on a desire-based or morally unworthy maxim.

⁸ Note how this argument would address the worries about the ‘common idea of duty’ that appear once it is on the table (cited at the beginning of this chapter). If the Categorical Imperative is the internal norm of a necessary self-conception, it is indeed reasonable to acknowledge its deliberative priority. The fact that this self-conception includes a robust idea of freedom speaks to the concern that ordinary moral thought presupposes an untenable conception of our motivational capacities.

⁹ This is the so-called ‘Incorporation Thesis’, stated at *Rel* 6: 24, that a free will is never determined by any incentive ‘except insofar as the human being has incorporated it into his maxim’. For discussion, see Allison 1990: 5ff., 39ff., and Reath 2006: 12–13, 17–21.

Another approach is to attribute to Kant the view that all rational choice is constitutively aimed at good in the sense that it is based on maxims that are taken by the agent to satisfy a standard of universal validity. According to this conception of rational agency, FUL is the formal principle of free agency in that it tacitly guides all rational choice. The freedom and spontaneity of rational volition is captured by the fact that it is guided normatively rather than causally, and moreover by its own internal norm. How would bad choice fit into this picture? Here one needs to hold that even bad choice understands itself to satisfy a condition of universal validity and is tacitly guided by FUL. Since it is guided by the formal norm of volition, it is free. At the same time the choice is bad because it is defectively or incompletely guided by this norm—for example it is guided by bad reasoning or a misrepresentation of this norm. This conception of choice as constitutively aimed at good is controversial, both philosophically and as an interpretation of Kant. But in holding that all choice is in some way guided by the formal norm of volition, it provides a way to make out the view that even choice on desire-based or morally unworthy maxims is free.¹⁰

Like the *Groundwork*, the *Critique of Practical Reason* aims to establish the authority of the fundamental principle of common-sense morality, but in the context of Kant's overall critical system.¹¹ The task of the second *Critique* is 'merely to show that there is pure practical reason' (*KpV* 5: 3)—that is, to show, contrary to empiricism, that reason by itself yields a basic principle of conduct, independently of empirically given aims, and that practical reason is not limited to instrumental and prudential reasoning. This principle is, of course, the moral law.

To show that pure reason is practical, Kant begins with a definition of a 'practical law' as a principle with authority for any rational agent, and he moves analytically from this concept to a statement of the 'fundamental law of pure practical reason', which is recognizable as the moral law (*KpV* 5: 19–32). This argument parallels the argument in the *Groundwork* II from the concept of a categorical imperative to the statement of the FUL, and it contains many points of overlap with the earlier work. For example, Kant argues that the fundamental moral principle must be a formal principle—one whose normative force depends upon its form rather than an interest in its purpose or 'matter'—that moral theories that base the fundamental principle on an external object presented to the will lead to heteronomy and cannot ground true practical laws, and that the fundamental moral principle must be a principle of autonomy that the rational will gives to itself (*KpV* 5: 27, 39–41, 33). Furthermore, the *Critique* argues for the same analytic connection between freedom and morality (*KpV* 5: 28–30).

However, there are at least two important differences between these works. In *Groundwork* III Kant attempts a 'deduction' of the moral law, and as we have seen, one component of this argument is that general features of rationality, including the spontaneity of

¹⁰ This approach is endorsed in different ways by Korsgaard 2009, especially ch. 5; Herman 2007: chs 10 and 11; and Engstrom 2009. Hill argues against such a reading in his 2002a: ch. 8.

¹¹ For overviews of the second *Critique*, see Reath 1997 and Engstrom 2002. Reath and Timmermann 2010 is a volume of critical essays on the *Critique*.

theoretical reason, give us non-moral grounds for ascribing free agency to ourselves. However in the second *Critique*, Kant says that a 'deduction' of the moral law is neither possible nor necessary, and that its authority is established directly as a 'fact of reason' (*KpV* 5: 31, 42, 46–7). Furthermore, he claims that grounds for ascribing free agency to ourselves are found only in moral consciousness. 'Morality first discloses to us the concept of freedom' and in place of the 'vainly sought deduction of the moral principle', this principle 'serves as the principle of the deduction of an inscrutable faculty which no experience could prove... namely the faculty of freedom' (*KpV* 5: 30, 47). Kant retains the idea that the moral law is the formal principle of free agency, but now holds that the capacity for free agency is grounded in our recognition of the authority of and ability to act from moral principle.

The basic idea that the moral law is given as a fact of reason is stated in the following passage:

Consciousness of this fundamental law may be called a fact of reason [*Faktum der Vernunft*] because one cannot reason it out from antecedent data of reason, for example, from consciousness of freedom (since this is not antecedently given to us) and because it instead forces itself upon us of itself as a synthetic a priori proposition that is not based on any intuition, either pure or empirical. (*KpV* 5: 31)

This passage holds that the authority of morality is ultimately self-standing and not based on anything outside of reflective moral consciousness, for example on general features of rationality or an independently given notion of freedom. Kant is not abandoning the project of justifying the authority of morality, but rather adopts a different approach. One influential interpretation takes the fact of reason to be our consciousness of the authority of morality in everyday thought, judgement, and feeling.¹² For example, in our common-sense judgements about right action (*KpV* 5: 30, 32), recognition of the distinction between moral reasons and reasons based in happiness (*KpV* 5: 91–3), and in the feeling of respect for morality (*KpV* 5: 72–89) we find that on reflection we do accept the authority of morality and cannot reject it without significant loss. A further 'credential' for the moral law is that moral consciousness reveals our free agency, a capacity with which we identify (*KpV* 5: 48, 86–87). Commentators have pointed out that '*Faktum*' is taken from the Latin '*factum*' and should be understood as a deed or action, and early in the *Critique* Kant writes that the reality of pure practical reason is given 'through the deed' [*durch die Tat*] (*KpV* 5: 3). Thus Kant's claim that the moral law is given as a fact of reason is that its authority is established through what reason does in us—that moral consciousness is our awareness of the activity of pure practical reason in us.¹³

There is scholarly consensus that the fact of reason represents an approach to foundational issues that is different from that of *Groundwork* III. But opinion divides over how

¹² See Rawls 2000: 255–60, 268–71.

¹³ Here see Engstrom 2002: xli–xliii. For further discussion see Kleingeld 2010: 57–65.

deep these differences are, over their relative merits, and how close either approach comes to succeeding. These are ongoing philosophical issues that cannot be settled in this chapter.¹⁴

21.3 KANT'S NORMATIVE THEORY

The main sources for Kant's normative theory are the *Groundwork* and the *Metaphysics of Morals*. The *Groundwork* is concerned primarily with foundational issues, but Kant illustrates the application of the Categorical Imperative in both its Universal Law and the Humanity versions through a set of examples. Clearly he takes the Categorical Imperative to be the basis of a deliberative procedure that leads to a set of substantive moral principles. The *Metaphysics of Morals* is intended to give a more complete presentation of his normative theory by applying the fundamental principle to the conditions of human life. Many questions about both the interpretation and the assessment of Kant's normative theory remain unresolved. For example, there is disagreement over how best to understand the Categorical Imperative and whether it provides an adequate guide to moral judgement, and the relation between the Categorical Imperative and the system of principles in the *Metaphysics of Morals* is not entirely transparent. I shall offer only a brief outline of his normative theory, beginning with the Categorical Imperative and focusing on the Universal Law version. Difficult issues of application (how to deal with moral conflict, for example) are beyond the scope of this chapter.

The Categorical Imperative applies a condition of universal validity to an agent's maxim of action. A maxim is an agent's subjective principle of action—the principle on which a subject acts or considers acting in some situation, which captures an agent's sincere understanding of what he intends to do and why he thinks it worth choosing.¹⁵ In Kant's examples, maxims contain an action kind as well as the end and features of the agent's circumstances that are taken to support it. We might think of maxims as representations of the practical reasoning that goes into a potential or actual choice. The focus on maxims indicates that Kant's theory assesses volition or choice in terms of the practical reasoning that underwrites and is taken to support it.

The Formula of Universal Law leads to a deliberative procedure for assessing the universal validity of a proposed maxim of action: one must be able both to adopt the maxim and will that anyone adopt the maxim (will it as universal law) without inconsistency or

¹⁴ For general discussions of these issues, see Henrich 1994; Rawls 2000: 253–72; Allison 1990: ch. 13; Ameriks 2003: ch. 10; Sussman 2008; Kleingeld 2010; and Timmermann 2010. Ameriks and Wood (2008: 135) are generally critical of the fact of reason approach, while Henrich, Allison, and Rawls defend it. Rawls, for example, does not think that it is a reversion to dogmatism or intuitionism and holds that the idea that morality is self-authenticating is a genuine advance (2000: 266–8).

¹⁵ Rawls stresses that a maxim must be a lucid and sincere statement of an agent's reasons in order to deal with the problem of 'rigged maxims'—artificial maxims designed to evade the criterion of universal validity. See Rawls 2000: 167–70 and Hill 2002b: 67ff.

irrationality. Here Kant draws a distinction between maxims that cannot be conceived as universal law and maxims that are not rationally willed as universal law given certain necessary interests (G 4: 424). The first kind of failure of universality has come to be called a 'contradiction in conception' and the second a 'contradiction in will' (O'Neill 1989: 96–101). The reasoning associated with the first imposes a condition of universal validity on discretionary choice, while that associated with the second grounds duties to adopt certain obligatory ends that Kant in the *Doctrine of Virtue* terms 'duties of virtue' (*MdS* 6: 383).

Kant's maxim of deceptive promising is a maxim that cannot even be conceived as universal law because it is self-defeating if universalized. If everyone were free to make deceptive promises for reasons of self-interest, the practice of binding oneself through a promise and the background of trust on which it depends would be undermined. In such a world, either the idea of a false promise would be incoherent since the practice of promising would no longer exist, or there would be no point in making a false promise since it would not be believed. Thus one cannot consistently will both the maxim and its universalization. Maxims that involve cheating and free-riding fail of universality in the same way. Deception attempts to manipulate another through his rational agency, and focusing on this feature of deception suggests a general strategy for showing that maxims that involve interference with the freedom of another, threat and coercion, exploitation, and so on cannot be willed as universal without inconsistency. All such maxims aim to interfere with or to control the agency of another for one's own purposes. But as rational agents who act under the idea of freedom, we necessarily conceive of ourselves to have the capacity to act from our own judgements about reasons and have an interest in exercising that capacity. To will as universal law a maxim that involves interference with or control of the agency of others for one's own purposes is to will that the general conditions of rational agency be undermined, and is inconsistent with this necessary interest. Along these lines, one might argue that such maxims cannot be conceived as universal law for agents who act under the idea of freedom. In the *Groundwork*, Kant claims that maxims that cannot be conceived as universal law are contrary to perfect or strict duty. He intends this form of reasoning to identify a set of action types undertaken for certain kinds of reasons that are impermissible, and thus to set limits on the ways in which we may pursue our ends.

Kant's maxim of indifference leads to a contradiction in will when universalized. The maxim is that while one is never to interfere with the freedom or property of another, one is never to assist others in need (unless doing so is to one's advantage) (G 4: 423). Kant notes that there is no inconceivability in this maxim holding universally, but that it would be irrational to will it as universal law. As rational agents with the capacity to set our own ends, we necessarily have an interest in being able to achieve them and in the conditions that support the effective exercise of our agency and end-setting capacities. Given our limited powers and vulnerabilities, a social climate in which people are willing to assist others in need is one of these conditions. Willing the maxim of indifference as universal law conflicts with this necessary interest and thus produces a 'contradiction in one's will'. A similar form of reasoning can be used to show that one has a

duty to develop one's natural talents. Rather than identifying ways of acting that are impermissible, this form of reasoning leads to imperfect duties—duties to adopt certain general ends or policies (in addition to one's personal ends and interest in happiness).

To provide a rationale for Kant's distinction between these two kinds of reasoning that explains why they lead to different kinds of duties, one might invoke the different ways in which they draw on the conditions of rational agency. Contradiction in conception identifies as wrong maxims whose universalization undermines the conditions of rational agency. The failure of universalizability here is due to the fact that such maxims directly interfere with or exploit rational agency, and they are appropriately classified as violations of perfect duty. By contrast, contradiction in will rejects maxims that fail to support the effective exercise of rational agency—maxims whose universalization is in conflict with our necessary interest in the effective exercise of our end-setting capacities.¹⁶ This form of reasoning leads to the ends of virtue, which are required for 'positive agreement' with humanity as an end in itself. This line of thought connects FUL and FH by bringing out the fact that maxims that do not satisfy the condition of universal validity are in different ways failures to respect rational nature as an end in itself.

Commentators disagree about the role that Kant intends for the FUL and about its adequacy as a general normative principle. For example, problems arise when it is applied to quite specific or qualified maxims, because most are consistently universalized, including many that are intuitively impermissible. Because of such problems, some philosophers argue for giving primacy to FH, despite Kant's claim that FUL provides the 'strict method' for moral appraisal (*G*4: 436).¹⁷ FUL certainly does not provide a mechanical procedure for assessing maxims of action. Philosophers who have defended its viability argue that it is best applied to maxims of a fairly high level of generality, to yield general moral principles that require further judgement for application to specific circumstances. Another worry is that since FUL most naturally assesses permissibility, it only leads to positive moral requirements indirectly by showing that certain maxims are impermissible. But one might want a more direct route to positive moral requirements (e.g. of honesty, fidelity, or beneficence). To address this concern one might apply the condition of universal validity directly to certain necessary human interests.¹⁸ For example, applying the condition of universal validity to the necessary interest in the exercise of rational agency leads directly to a set of principles that both protect and support the effective exercise of rational agency.

¹⁶ For discussion see O'Neill 1989: 96–7, 133–4; Herman 1993: 125–7, 225–30; and Reath 2006: 211–20.

¹⁷ For discussion of some of these problems, including the problems of 'false negatives' and 'false positives', see Wood 1999: 97–110. He argues that FH leads to a more viable approach to moral reasoning. See Wood 1999: 111–56, 182–91, and Wood 2002: 12–13. For other discussions of the application of FH, see Hill 1992: 38–57 and 2002b: 79–83, O'Neill 1989: 105–25, and Korsgaard 1996: 106–32. Philosophers who have defended FUL include O'Neill 1989: 81–104, Herman 1993, Korsgaard 1996: 77–105, and Engstrom 2009.

¹⁸ Such an approach is developed in Engstrom 2009: 184–240.

The *Metaphysics of Morals* distinguishes the domains of right (*Recht*) and ethics, and accordingly contains two parts—the *Doctrine of Right* and the *Doctrine of Virtue*. Each domain is governed by a fundamental principle that is nominally a version of the Categorical Imperative. All duties are principled constraints on free choice. Right and ethics regulate different domains of freedom—outer and inner freedom—and they do so in different ways.

Right regulates external action in so far as it affects the freedom of others in accordance with the necessary conditions of equal outer freedom. Kant's 'Universal Principle of Right' reads: 'An action is right if it can coexist with everyone's freedom in accordance with a universal law' (*MdS* 6: 230). This principle is a pared down version of the contradiction in conception component of FUL that focuses on actions outwardly described. It imposes a formal constraint on actions taken in pursuit of one's purposes and implies that an action is wrong if it interferes with an action that is consistent with the equal freedom of all. The contours of right are determined by the fact that this principle only regulates outer freedom. Because of this, it cannot require compliance from any particular incentive, nor does it require that agents adopt any particular ends. Further, right creates title to employ coercion to protect equal freedom, since coercion that opposes 'hindrances to freedom' is consistent with the equal freedom of all and therefore right (*MdS* 6: 231). From these points it follows that the 'law-giving' employed by right may be 'external'—that is, the incentives that constrain choice to conditions of right may appeal to self-interest and need not include any thought of duty (*MdS* 6: 218–20). Finally, right creates strict requirements, requiring specific actions and omissions. (*MdS* 6: 388–9). The Universal Principle of Right is the basis of Kant's legal and political theory, including his account of private right (property, contract, and status) and his theory of the state.¹⁹

Ethics goes beyond right by requiring the adoption of certain ends and introduces the important concept of a duty of virtue.²⁰ A duty of virtue is a requirement of reason to adopt certain general ends, values, and fundamental attitudes towards persons—in effect a duty to adopt some fundamental maxim (*MdS* 6: 389). The 'supreme principle of the doctrine of virtue' is: 'Act in accord with a maxim of ends that it can be a universal law for everyone to have' (*MdS* 6: 396). One should expect this principle to align with the contradiction in will component of FUL (or the positive dimension of FH). The most general duty of virtue is to adopt the fundamental principle of respecting humanity as an end in itself. Kant also stresses the more specific ends of one's own natural and moral perfection and promoting the happiness of others. Moral perfection involves cultivating the disposition to act from the motive of duty—a fundamental commitment to moral principle and the priority of moral considerations. It includes the disposition to comply

¹⁹ For discussion see Ripstein 2009a and 2009b. He argues that Kant's theory of right is based on a conception of external freedom as independence from the will of another and holds that political institutions are necessary conditions of equal freedom as independence.

²⁰ For general discussions of the *Doctrine of Virtue* see Wood 2002 and Hill 2010, as well as the essays in Timmons 2002 and Denis 2010.

with duties of justice from the motive of duty. In other words it is a duty of virtue to make 'the right of humanity' one's end—to take on the commitment to honour the juridical rights of others out of respect for morality, thus going beyond what right requires (*MdS* 6: 390, 395).

In various ways, ethics is concerned with the conditions of internal freedom, or self-constraint through principles of pure practical reason (*MdS* 6: 396). Because ends are always adopted through free choice and one can never be constrained by others to adopt an end, duties of virtue are not externally enforceable. Thus the ends of virtue require self-constraint through a conception of duty, and in that respect 'ethical law-giving'—the incentive attached to duties of virtue—is internal (*MdS* 6: 218–20, 380–4). Since duties of virtue are duties to adopt general maxims (maxims of ends) they are of 'wide obligation'. They do not require specific actions and leave agents latitude in how to act on these ends and general principles. But Kant also notes that 'a wide duty is not to be taken as permission to make exceptions to the maxims of action but only as permission to limit one maxim of duty by another... by which in fact the field for the practice of virtue is widened' (*MdS* 6: 390). Even though agents have discretion in implementing the ends of virtue, Kant appears to assign them an important role in shaping an agent's overall system of ends.²¹

The *Doctrine of Virtue* develops a classification in which duties to oneself figure prominently. In addition to the imperfect duties of natural and moral self-development, Kant includes perfect duties to oneself. Violations of perfect duty to oneself include suicide, and interestingly lying (since it involves misuse of one's capacity to communicate), avarice (which displays 'slavish subjection of oneself to the goods that contribute to happiness' [*MdS* 6: 434]), and servility (as a basic failure of self-respect). These are perfect duties since they are duties to preserve one's natural and moral capacities, but duties of virtue since they are satisfied by adopting certain basic attitudes towards oneself. Kant also includes duties to develop receptivity to moral norms, to cultivate conscience, and to develop the forms of self-knowledge needed for good willing and virtue (*MdS* 6: 399–403, 441–2. Cf. Hill 2010: 246–9). Perhaps the unifying theme in these duties is the preservation and development of the capacity for inner freedom and self-government.

The duties of virtue to others fall into the categories of duties of love and respect. Kant does not understand love and respect in this context as feelings, but rather as maxims or active practical attitudes towards others (*MdS* 6: 449). Duties of love require active concern for the well-being of others—beneficence and the cultivation of a capacity to share others' feelings that supports beneficence (*MdS* 6: 452–7). One respects others by acknowledging their equal standing—'by limiting our self-esteem by the dignity of humanity in another person' (*MdS* 6: 449). Among other things, respect requires restraining the tendency to affirm one's own worth in defective ways, for example through arrogance ('a demand that others think little of themselves in

²¹ Barbara Herman has argued that the ends of virtue may be understood as ends of all rational choice by shaping one's conception of happiness from the inside. See Herman 2007: 254–75.

comparison with us' [*MdS* 6: 465]) or mocking the faults of others. Kant's discussion of these particular duties should not be understood as an attempt at an exhaustive list, but rather as an illustration of what is called for by respect for humanity in oneself and others.

BIBLIOGRAPHY

Kant's Works

Citations will use the abbreviations given below, and will be to the Prussian Academy of Sciences edition of Kant's collected works (Berlin: de Gruyter, 1900–).

- G *Groundwork of the Metaphysics of Morals* in *Kant: Practical Philosophy*, tr. and ed. M. J. Gregor. Cambridge/New York: Cambridge University Press, 1996.
- Idea *Idea for a Universal History with a Cosmopolitan Aim* in *Kant: Anthropology, History and Education*, tr. and ed. R. B. Louden and G. Zöllner. Cambridge/New York: Cambridge University Press, 2008.
- KpV *Critique of Practical Reason* in *Kant: Practical Philosophy*.
- LE *Lectures on Ethics*, ed. P. Heath and J. B. Schneewind, tr. P. Heath. Cambridge/New York: Cambridge University Press, 1997.
- MdS *The Metaphysics of Morals* in *Kant: Practical Philosophy*.
- PP *Toward Perpetual Peace* in *Kant: Practical Philosophy*.
- Rel *Religion within the Boundaries of Mere Reason*, tr. and ed. A. Wood and G. di Giovanni. Cambridge/New York: Cambridge University Press, 1998.
- TP *On the Common Saying: That may be Correct in Theory but it is of no use in Practice* in *Kant: Practical Philosophy*.

Secondary Sources

- Allison, H. 1990. *Kant's Theory of Freedom*. Cambridge/New York: Cambridge University Press.
- 1996. *Idealism and Freedom: Essays on Kant's Theoretical and Practical Philosophy*. Cambridge/New York: Cambridge University Press.
- Ameriks, K. 2003. *Interpreting Kant's Critiques*. Oxford/New York: Oxford University Press.
- Denis, L. (ed.) 2010. *Kant's 'Metaphysics of Morals': A Critical Guide*. Cambridge/New York: Cambridge University Press.
- Engstrom, S. 2002. 'Introduction', in Kant, *Critique of Practical Reason*, tr. Werner S. Pluhar. Indianapolis/Cambridge, MA: Hackett Publishing, xi–liv.
- 2009. *The Form of Practical Knowledge*. Cambridge, MA: Harvard University Press.
- Guyer, P. 1995. 'The possibility of the categorical imperative', *Philosophical Review* 104: 353–85. Reprinted in Paul Guyer, *Kant on Freedom, Law, and Happiness*. Cambridge/New York: Cambridge University Press, 2000, 172–206.
- (ed.) 1998. *Groundwork of the Metaphysics of Morals: Critical Essays*. Lanham, MD: Rowman & Littlefield.

- 2009. 'Problems with freedom: Kant's argument in *Groundwork* III and its subsequent emendations,' in Timmermann (ed.), 2009, 172–202.
- Henrich, D. 1994. 'The concept of moral insight and Kant's doctrine of the fact of reason,' in D. Henrich, *The Unity of Reason*, ed. and tr. Richard Velkley. Cambridge, MA: Harvard University Press, 55–87.
- 1998. 'The deduction of the moral law: the reasons for the obscurity of the final section of Kant's *Groundwork of the Metaphysics of Morals*,' in Guyer (ed.), 1998, 303–41.
- Herman, B. 1993. *The Practice of Moral Judgment*. Cambridge, MA: Harvard University Press.
- 2007. *Moral Literacy*. Cambridge, MA: Harvard University Press.
- Hill, T. E. Jr. 1992. *Dignity and Practical Reason*. Ithaca: Cornell University Press.
- 2002a. *Human Welfare and Moral Worth*. Oxford/New York: Oxford University Press.
- 2002b. 'Editor's Introduction' and 'Analysis of arguments,' in Kant, *Groundwork for the Metaphysics of Morals*, tr. A. Zweig and ed. T. E. Hill, Jr. Oxford/New York: Oxford University Press, 19–177.
- 2010. 'Kant's *Tugendlehre* as normative ethics,' in Denis (ed.), 2010, 234–55.
- Horn, C. and Schönecker, D. (eds) 2006. *Groundwork for the Metaphysics of Morals*. Berlin: Walter de Gruyter.
- Kerstein, S. 2002. *Kant's Search for the Supreme Principle of Morality*. Cambridge/New York: Cambridge University Press.
- Kleingeld, P. 2010. 'Moral consciousness and the 'fact of reason',' in Reath and Timmermann (eds), 2010, 55–72.
- Korsgaard, C. M. 1996. *Creating the Kingdom of Ends*. Cambridge/New York: Cambridge University Press.
- 2008. *The Constitution of Agency*. Oxford/New York: Oxford University Press.
- 2009. *Self-Constitution*. Oxford/New York: Oxford University Press.
- O'Neill, O. 1989. *Constructions of Reason*. Cambridge/New York: Cambridge University Press.
- Prichard, H. A. 2002. 'Does moral philosophy rest on a mistake?', in his *Moral Writings*. Oxford/New York: Oxford University Press, 7–20.
- Rawls, J. 2000. *Lectures on the History of Moral Philosophy*. Cambridge, MA: Harvard University Press.
- Reath, A. 1997. 'Introduction,' in Kant, *Critique of Practical Reason*, tr. M. J. Gregor. Cambridge/New York: Cambridge University Press, vii–xxxii.
- 2006. *Agency and Autonomy in Kant's Moral Theory*. Oxford/New York: Oxford University Press.
- 2010. 'Formal principles and the form of a law,' in Reath and Timmermann (eds), 2010, 31–54.
- 2012. 'Formal approaches to Kant's formula of humanity,' in S. Baiasu and M. Timmons (eds), *Kant on Practical Justification*. Oxford/New York: Oxford University Press.
- 2013. 'The ground of practical laws,' in S. Bacin, A. Ferrarin, C. La Rocca, and M. Ruffing (eds), *Kant and Philosophy in a Cosmopolitan Sense: Proceedings of the XIth International Kant Congress*. Berlin: De Gruyter.
- and Timmermann, J. (eds) 2010. *Kant's 'Critique of Practical Reason': A Critical Guide*. Cambridge/New York: Cambridge University Press.
- Ripstein, A. 2009a. 'Kant on law and justice,' in T. E. Hill, Jr. (ed.), *The Blackwell Guide to Kant's Ethics*. Malden, MA: Wiley-Blackwell Publishing, 161–78.
- 2009b. *Force and Freedom: Kant's Legal and Political Philosophy*. Cambridge, MA: Harvard University Press.

- Schönecker, D. 2006. 'How is a categorical imperative possible? Kant's deduction of the categorical imperative', in Horn and Schönecker (eds), 2006, 301–24.
- Sensen, O. 2009. 'Kant's conception of human dignity', *Kant-Studien* 100: 309–31.
- 2010. 'Dignity and the formula of humanity', in Timmermann (ed.), 2010, 102–18.
- Sussman, D. 2008. 'From deduction to deed: Kant's grounding of the moral law', *Kantian Review* 13: 52–81.
- Timmermann, J. 2007. *Kant's Groundwork of the Metaphysics of Morals: A Commentary*. Cambridge/New York: Cambridge University Press.
- 2010. 'Reversal or retreat: Kant's deductions of freedom and morality', in Reath and Timmermann (eds), 2010, 73–89.
- (ed.) 2009. *Kant's 'Groundwork of the Metaphysics of Morals': A Critical Guide*. Cambridge/New York: Cambridge University Press.
- Timmons, M. (ed.) 2002. *Kant's Metaphysics of Morals: Interpretive Essays*. Oxford/New York: Oxford University Press.
- Wood, A. W. 1999. *Kant's Ethical Thought*. Cambridge/New York: Cambridge University Press.
- 2002. 'The final form of Kant's practical philosophy', in Timmons (ed.), 2002, 1–21.
- 2008. *Kantian Ethics*. Cambridge/New York: Cambridge University Press.