

- Hume E David Hume, *Enquiry Concerning Human Understanding* and *Enquiry Concerning the Principles of Morals*, ed. Selby-Bigge and Nidditch. Oxford: Oxford University Press, 1995.
- Hume T David Hume, *A Treatise on Human Nature*, ed. Selby-Bigge. Oxford: Clarendon Press, 1967.
- Korsgaard CKE Christine Korsgaard, *Creating the Kingdom of Ends*. Cambridge: Cambridge University Press, 1996.
- Korsgaard FC Christine Korsgaard, "Fellow Creatures: Kantian Ethics and Our Duties to Animals," in Grethe B. Peterson (ed.), *The Tanner Lectures on Human Values*, Vol. 25. Salt Lake City: University of Utah Press, 2005.
- Korsgaard SN Christine Korsgaard, *The Sources of Normativity*. Cambridge: Cambridge University Press, 1996.
- Mill John Stuart Mill, *Utilitarianism*, ed. George Sher. Second edition. Indianapolis: Hackett, 2001.
- O'Neill Onora O'Neill, *Constructions of Reason*. Cambridge: Cambridge University Press, 1989.
- Rawls TCR John Rawls, "Two Concepts of Rules," *Philosophical Review* 64 (1955), pp. 3-32.
- Rawls TJ John Rawls, *A Theory of Justice*. Cambridge, Mass.: Harvard University Press, 1971.
- Rawls Lectures John Rawls, *Lectures on the History of Moral Philosophy*, ed. Barbara Herman. Cambridge, Mass.: Harvard University Press, 2000.
- Reath Andrews Reath, *Agency and Autonomy in Kant's Moral Theory: Selected Essays*. Oxford: Oxford University Press, 2006.
- Rousseau D Jean-Jacques Rousseau, *Discourse on the Origin of Inequality*, tr. Donald Cress. Indianapolis: Hackett, 1992.
- Rousseau E Jean-Jacques Rousseau, *Émile, or On Education*, tr. B. Foxley. New York: Dutton, 1969.
- Rousseau SC Jean-Jacques Rousseau, *On the Social Contract*, tr. Victor Gourevitch. Cambridge: Cambridge University Press, 1997. Cited by book and chapter.
- Scanlon T. M. Scanlon, *What we owe to each other*. Cambridge, Mass.: Harvard University Press, 1998.
- Sidgwick Henry Sidgwick, *The Methods of Ethics*. Indianapolis: Hackett, 1981.
- Williams Bernard Williams, *Moral Luck*. Cambridge: Cambridge University Press, 1981.

Reason

1. What Is Kantian Ethics?

Some recent moral philosophers draw a distinction between *Kant's ethics* and *Kantian ethics*.¹ *Kant's ethics* is contained in Kant's own writings: the *Groundwork*, the *Critique of Practical Reason*, the *Metaphysics of Morals*, and the others. It is the theory Kant himself put forward, the fundamental principle of morality as he formulated it, the system of duties as he presented it, even the moral conclusions he thought followed from them. To write about *Kant's ethics* is to interpret that theory, to show how its parts are supposed to fit together, to relate it to Kant's philosophy as a whole. *Kantian ethics*, on the other hand, is an ethical theory formulated in the basic spirit of Kant, drawing on and acknowledging a debt to what the author of the theory takes to be his insights in moral philosophy. Kantian ethics is not merely, or even mainly, an interpretation of what Kant said. It is put forward instead as a theoretical option in thinking about ethical questions and philosophical questions about ethics. It is answerable not to textual accuracy or exegetical standards of Kant interpretation but to the right standards for thinking philosophically about ethical theory and ethical issues.

It should be clearly understood, however, what these standards are – and what they are not. Some philosophers seem to think that each proposition in a theory must be argued for entirely on its own, using arguments that are supposed to persuade anyone at all, even someone with no sympathy whatever for the project in which the theory is engaged. That is a standard that no significant philosophical theory could ever meet. In fact, the best defense of any philosophical conception is always a more or less systematic exposition of it. It is reasonable to ask for arguments on behalf of individual claims, especially fundamental ones, but these too are to be understood in the context of the theory as a whole. A philosophical theory is best defended by letting us see clearly how it conceives its task, how it performs it, and how the resulting conception of the subject matter addresses the questions

reasonable people have about that subject matter. No philosophical theory is going to persuade everyone. What we should look for in a philosophical theory is one that, when presented in this comprehensive way, not only looks appealing, but its rejection also can be seen to incur significant intellectual costs that we should be reluctant to pay.

This means that "Kantian ethics" as I mean the term may sometimes look something like a sympathetic interpretation of Kant's writings, even if its aim is quite different. Kantian ethics, however, certainly may depart freely from what Kant wrote and thought. It may criticize and modify the theory Kant put forward as well as sympathetically interpret or defend it. The present book is intended as an exercise in Kantian ethics in this sense. But it will also have a lot to say about Kant's ethics. This is because I do not think the most defensible version of Kantian ethics needs to depart as far from what Kant thought and wrote as most recent practitioners of Kantian ethics do. What is needed instead, in many cases, is only a better understanding of Kant's own thoughts.

One way of understanding the term 'Kantian ethics,' however, involves the at least tacit assumption that we already know what ethics is (from currently fashionable ideas about the aims and methods of ethical theory). "Kantian ethics" is simply a matter of seeing what Kant has to contribute to this project. In my view, however, the main benefit of studying an important figure in the history of philosophy, such as Kant, is that doing so helps us learn that the current philosophical fashions are not the only way to think about things. Philosophers (like other people) have a deplorable tendency to think in terms of entrenched prejudices. On many points, I will criticize standard interpretations of Kant for having interpreted Kant in terms of fashionable assumptions about ethical theory that have frequently been imposed on his writings – sometimes with charitable intent, but often with profoundly distorting effect. In Chapter 3, I argue that Kant's conception of ethical theory – its aims, methods, and conception of ethical reasoning – differs significantly from prevailing conceptions.

A much better reason for developing Kantian ethics in ways that diverge from Kant himself is indicated by the wry title of Marcia Baron's book *Kantian Ethics (Almost) Without Apology*. Those who find Kantian ideas in ethics appealing also sometimes feel that there is something about this for which they need to apologize. No doubt some of Kant's opinions on particular ethical topics are – or at least seem at first glance to be – so out of touch with enlightened opinion today as to seem either ridiculous or repugnant. But I suspect that those who think we need to apologize for Kantianism in ethics are using these opinions only to confirm a certain traditional image of Kantian moral philosophy. Kant is seen exclusively as a representative of moralistic strictness and sternness, downright hostile to human happiness, mercilessly unsympathetic to human weakness, allowing no place in the moral life for natural human feelings and desires. People may sometimes

see an element of truth in this aspect of morality, but they view the Kantian version of this truth as wildly exaggerated, one-sided to the point of inhumanity. This image of Kant is colorfully presented by Simon Blackburn:

For Kant, so the contrast goes, there is indeed the Humean crew. But standing above them, in the quarter-deck, there is another voice – a voice with ultimate authority and ultimate power. This is the Captain, the will, yourself as an embodiment of pure practical reason, detached from all desires. The Captain himself is free. But he always stands ready to stop things going wrong with the crew's handling of the boat. Sometimes, it seems, the happiest ship will have no crew at all, but only a Captain . . . Thus the Kantian Captain. He is a peculiar figure, a dream – or nightmare – of pure, authentic self-control. He certainly appeals to our wish to be, ourselves, entirely the masters of our own lives, immune in all important respects from the gifts or burdens of our internal animal natures, or of our temperaments as they are formed by contingent nature, socialization, and external surrounds. Context-free, non-natural, and a complete stickler for duty, perhaps the Kantian self is nothing but the sublimation of a patriarchal, authoritarian fantasy.²

Even more flamboyant is the following remark by Richard Taylor:

I have known many admirers of Kant, and include myself with them; but if I were ever to find, as I luckily never have, a man who assured me that he really believed Kant's metaphysical morals, and that he modeled his own conduct and his relations with others after those principles, then my incredulity and distrust of him as a human being could not be greater than if he told me he regularly drowned children just to see them squirm.³

The starting point for a less fantastic image of Kantian ethics was well stated by John Rawls. He regarded Kantian ethics "not as a morality of austere command but an ethic of mutual respect and self-esteem" (Rawls TJ, p. 256). Kant was a philosopher of the Enlightenment – perhaps the greatest of all Enlightenment philosophers. For Kant, the principle of Enlightenment is: "Think for yourself!" (WA 8:35, O 8:146). This means: Take the responsibility for your own actions and convictions. Do not put yourself under the *tutelage* or *authority* of others or let them do your thinking for you, however much, in thinking for yourself, you may need to listen to their arguments or treat their expertise as good evidence in the formation of your own judgments. This principle is based on respect for yourself as a rational being, arising from the recognition of rational nature in your own person as an end in itself (G 4:429). The same principle, however, requires you to respect rational nature in the person of every other human being. Each human being, as rationally self-governing according to universally valid standards, has *dignity* or *absolute worth* (G 4:431–6). Because the worth of every human being is absolute, the worth of all persons is fundamentally equal.

Kant's moral outlook, in its fundamentals, is a characteristically Enlightenment outlook. In its time the Enlightenment was an important part of an

emerging intellectual movement, a way of thinking that still exists today. In the eighteenth century this outlook was strongly opposed by antirationalistic and traditionalist ways of thinking, and it is still under attack in our time both from antirationalists and conservatives. To see Kantian ethics only through the lens of malicious or condescending caricatures is therefore not only to misread an influential historical philosopher but also to blind yourself to a lot of the ongoing cultural life of modernity. If you are on the Enlightenment side of the ongoing struggle, then to confuse Kantian ethics with your own nightmares about moral authoritarianism is to mistake one of your closest friends for one of your worst enemies.

2. Human Nature

To Rawls's felicitous formulation I want to add something else almost as important. Kant's moral outlook is also fundamentally determined by a subtle, shrewd, historically self-conscious (and characteristically Enlightenment) conception of human nature and human psychology that most treatments of Kantian ethics (even sympathetic ones) have largely overlooked. This side of Kant owes a great deal to Jean-Jacques Rousseau, and it belongs to a radical tradition in the social criticism of modernity whose later representatives include Johann Gottlieb Fichte and Karl Marx. The Kantian mistrust of our empirical desires reflects a Rousseauian picture of the way our natural desires have been influenced by the loss of innocence – the restless competitiveness – characteristic of human beings in the social condition, especially as found in the social inequalities of what Rousseau and Kant called the “civilized” stage of human society but was later renamed “modern bourgeois society” or “capitalism.” Again, to miss this continuity is not only to misread Kant; it is badly to misread the history, and even the living reality, of the social order that is all around us.

Kant's famous mistrust of our empirical “inclinations” is mistrust of “nature” only insofar as our nature has been shaped by society. Kant asserts (as explicitly as it would be possible for him to do) that there is nothing at all in our “animality” – our animal instincts for survival, reproduction, and sociability – that could be called “evil” or held responsible for it. Our inclinations, considered in themselves, as expressions of our bodily or animal nature are entirely good and “display themselves openly” for what they are. Kant holds that they become evil only insofar as vices have been “grafted onto them” by “an invisible enemy, one who hides behind reason and is hence all the more dangerous” (R 6:26–7, 57). This enemy is competitiveness, social inequality, the passion for domination over others.

Rousseau called it *amour propre* (Rousseau D, pp. 36, 53–4, 90; Rousseau E, pp. 172–6). Kant has various names for it. Alluding to Montaigne, he calls it “unsociable sociability” (I 8:20),⁴ at other times “self-conceit” (KpV 5:73), or, finally, the “radical propensity to evil” (R 6:28–32). For Kant, as

for Rousseau, this propensity develops along with our reason, hence only in the social condition (R 6:27).

It is not the instigation of nature that arouses what should properly be called the passions, which wreak such great devastation in his originally good [animal] predisposition. His needs are but limited and his state of mind in providing for them is moderate and tranquil. He is poor (or considers himself so) only to the extent that he is anxious that other human beings will consider him poor and will despise him for it. Envy, addiction to power, avarice, and the malignant inclinations associated with these, assail his nature, which on its own is undemanding, *as soon as he is among human beings.* Nor is it necessary to assume that these are sunk into evil and are examples that lead him astray: it suffices that they are there, that they surround him, and that they are human beings, and they will mutually corrupt each other's moral disposition and make one another evil. (R 6:93–4)

Against those theories that want to ground ethics on natural feelings, inclinations, or passions (such as sympathy), Kant has two main objections. One is that feelings and inclinations do not suffice to ground clear and determinate principles for action. But the deeper objection is that in human beings, no feelings, empirical desires, or passions are merely “natural” – that is, good or innocent. All are at the same time *social* (and socially corrupted), so that the most we can expect from them is a correspondence to what is morally good that is contingent and at best precarious. Ethical theories grounded on them therefore might give the right results for a different species of rational creatures, a species that was asocial or whose sociability was not, like ours, infected with self-conceited ambition and a passionate need to dominate our fellows. When applied to us, such theories are either too naïve or too complacent, especially in the context of our more developed or “civilized” societies.

In other words, Kantian ethics is fundamentally committed to a radical critique of human social life, especially of social life in its “civilized” form. This critical tendency is not a mere ancillary feature or contingent concomitant of Kantian ethics. It conditions the fundamental conception of Kantian ethical theory. For it is Kant's view that our only resource in combating the radical evil of our social condition is the faculty of reason, whose development accompanies that of our propensity to evil, and which alone enables us to recognize evil for what it is. This is why moral principles for Kant must be a priori rather than empirical in origin, and why we cannot trust our natural feelings, inclinations, or passions to provide us with moral distinctions, judgments, and motives.

Our use of reason itself, of course, is subject to the very same subversion as natural feelings and desires. Ordinary moral thinking, Kant says, is therefore vulnerable to a “dialectic” in which we tend to quibble with the demands of morality or adjust them to our wishes (G 4:405). Wouldn't it be nice if we had some other faculty, or some infallible (divine) source of moral

wisdom that is not subject to such corruption? But even supposing we did have such a source, our use of it would still be conditioned by our own interpretation, which would necessarily be our own thinking, hence subject to the same fallibility and corruption. Some circumvent this inconvenience either by saying that they are taking the word of this source "literally," or else by attributing their interpretation to the same infallible sources. It is almost charming how naïvely they thereby assert what is now obviously only their own infallibility. Such blasphemous arrogance would be only comical if its real-world consequences were less monstrous.

Thus in the end there is no escaping the fact that human reason – feeble, fallible, imperfect, corrupted reason – is always our last resort, even our only ultimate resource, for criticizing everything, including our own misunderstandings and abuses of reason itself. Kant's "critique of reason" thus takes "reason" in both the objective and subjective genitive – it is a critique *carried out by reason upon reason*. We rely on reason to criticize feelings, desires, inspirations, revelations, and even reason itself, not because it is infallible but rather because it is only through reason that we have the capacity to criticize or correct anything at all.

Kant's ethical theory holds that every human being has equal dignity as an end in itself, but his theory of human nature and history is based on the idea that civilized human beings tend to assert their self-worth antagonistically in relation to others, seeking superiority over them. Kant is sensitive to this tendency at work in all our desires, and also to the way it leads us to deceive ourselves about our own motives, our merits, and about what morality demands of us. He therefore thinks we need to guard against our corrupt tendency to quibble with the strictness of the moral law and make exceptions to moral rules in our own favor. This is even the reason why Kant thinks we need moral *philosophy* in addition to moral common sense or "common rational moral cognition" (G 4:405).

Kant thinks that the chief benefit of our social condition, in combating the evils that come along with it, is the development of *reason* – which he understands as the capacity to regulate our conduct by universal principles of respect and concern that we are capable of sharing with other rational beings. Reason is a capacity for self-government (which Kant emphasizes that human beings exercise with only very limited success) based on mutual respect and free communication, yielding a system of principles people can all share, and aiming at what he calls a "realm of ends," a system of human ends that can be rationally shared between all people because the dignity and welfare of all rational beings are equally included in it.

3. Gender and Race

Through the intellectual and cultural movement the eighteenth century called "Enlightenment," modernity is still struggling to free itself from the

chains and the pollution of traditional ideas and traditional ways of life and find a path toward a more rational and decent human future. In the writings of eighteenth-century representatives of this movement, we sometimes find a torch we may still use to light our way. At other times, however, we see them fettered by the very traditions – cultural or religious – from which, in their best thoughts, they were still trying to free themselves. Kant may be the greatest philosopher of the Enlightenment, but in this way he is also typical of it.

Kant's view of women. There are some special worries in this respect about Kant's views on race and gender. Although Kant's ethics is based on radically egalitarian principles, Kant accepted quite complacently the social and political subordination of women that prevailed in his time, and in some of his writings on anthropology he expressed views that can be described only as racist. The enterprise of interpretation, moreover, is sufficiently holistic in character that we cannot automatically dismiss the thought that these views might possibly require us to qualify in disturbing ways the seemingly egalitarian principles on which Kantian ethics appears to rest. It has been maintained, for example, that when Kant speaks of the dignity or absolute worth of humanity or rational nature, the referent of these terms must be understood as restricted only to white males.⁵

Such an extreme conclusion as that, however, is rendered indefensible by Kant's explicit statements including women and human beings of any and all races as rational beings and hence as falling within the scope of principles of right. For example, Kant's entire theory of marriage right, however repugnant parts of it may be, is motivated mainly by the need to protect the rights and human dignity of women. It is nevertheless true that he regarded women as weaker than men not only physically but also intellectually and thought it appropriate that they should be in a permanent condition of civil guardianship (*Vormundschaft*), represented in the public sphere by their fathers or husbands (VA 7:209).

Kant is a subject of lively controversy among feminist philosophers, some of whom see his entire moral philosophy as nothing but an ideology of patriarchy and male supremacy, while others regard Kantian ethics as the original articulation of principles of morality and right that are indispensable to women's liberation and equality of the sexes.⁶ Some of these issues will be addressed later, in Chapter 13. It is also relevant to point out that the criticisms of the former group of feminists often tend to follow a pattern of Kant interpretation and criticism that is by no means characteristically feminist but familiar from Romantic, Hegelian, virtue ethics, and other older traditions.⁷ This is precisely the misreading of Kantian ethics I have criticized above and will continue to criticize, especially in Chapters 2 and 8.

Kant on the inferiority of nonwhite races. During the 1770s and 1780s, Kant became increasingly interested in the empirical study of human nature, and one side of this was the development of a theory of race. He held that