# MORAL KNOWLEDGE AND ETHICAL CHARACTER

ROBERT AUDI

# Moral Knowledge
# and Ethical Character

*This page intentionally left blank*

# Moral Knowledge
# and Ethical Character

Robert Audi

## Oxford University Press

Oxford   New York

Athens   Auckland   Bangkok   Bogota   Bombay   Buenos Aires
Calcutta   Cape Town   Dar es Salaam   Delhi   Florence   Hong Kong
Istanbul   Karachi   Kuala Lumpur   Madras   Madrid   Melbourne
Mexico City   Nairobi   Paris   Singapore   Taipei   Tokyo   Toronto   Warsaw

and associated companies in
Berlin   Ibadan

# Preface and Acknowledgments

This book presents a large-scale position in ethical theory through twelve interconnected essays on four of its major dimensions. The first is moral epistemology, which concerns the possibility of knowledge and justification in ethics; the second is the ontology of ethics, which addresses the nature of moral properties and concepts and their relation to the natural world; the third is moral psychology, which concerns the nature and scope of moral motivation and judgment, their internalization in character, their influence on the will, and the extent of our responsibility for them; and the fourth is the foundations of ethics, which (in this book) concerns particularly the question of how motivational and valuational elements may be basic for moral reasons for action. The overall ethical theory I present combines four elements, each addressing one of the dimensions of ethical theory just sketched: a moral epistemology that is internalist, qualifiedly rationalistic, and moderately intuitionist; an ontology of ethics that is realist, pluralist, and non-reductively naturalistic; a moral psychology that countenances unconscious motivation and cognition, posits a good measure of control and moral responsibility for aspects of character, and describes how morality is internalized in us and reflected in action; and an account of value that posits a plurality of intrinsic goods and evils and yields a theory of how they provide objective reasons for action.

Five of the essays, comprising nearly half of the volume — chapters 3, 6, and 10 through 12— have not been previously published. In each part of the book, the essays are ordered chronologically. This order is appropriate to their content, and the same holds for the ordering of the four parts. Many chapters build on accounts, ideas, or distinctions introduced earlier. My aim has been to order the essays to provide as much cumulative development of my position as possible while also making each essay and each part of the book essentially self-contained. This should be a significant convenience to readers interested in some one part or a selection among the essays. The previously published essays have not been altered apart from minor changes chiefly intended to put the notes into a uniform style, and there is some overlap. This is mainly among chapters 2, 4, and 5, but where overlap occurs it is

generally developmental, and it yields some gain in the depth or comprehensiveness with which important questions are treated.

Those considering the book for teaching ethics may find it useful to conceive it in terms of *clusters* of essays, including a number of groupings not precisely corresponding to one of the volume's four parts. Among the essentially self-contained clusters are (1) *intuitionism* (or ethical rationalism), particularly chapters 2–5 and 12; (2) *moral psychology*, chapters 6–11; (3) *moral epistemology*, especially chapters 1–4; (4) the *foundations of ethics*, chapters 2 and 9–12; (5) *virtue ethics*, chapters 6–8 and perhaps 11; (6) the *metaphysics of ethics*, mainly chapters 4, 5, and 11; (7) *general metaethics* (objectivity, realism-anti/realism, moral explanation, justification in ethics), chapters 2, 3, 5, 10, 11, and perhaps 12. The concluding, synoptic chapter is appropriate to any of these clusters.
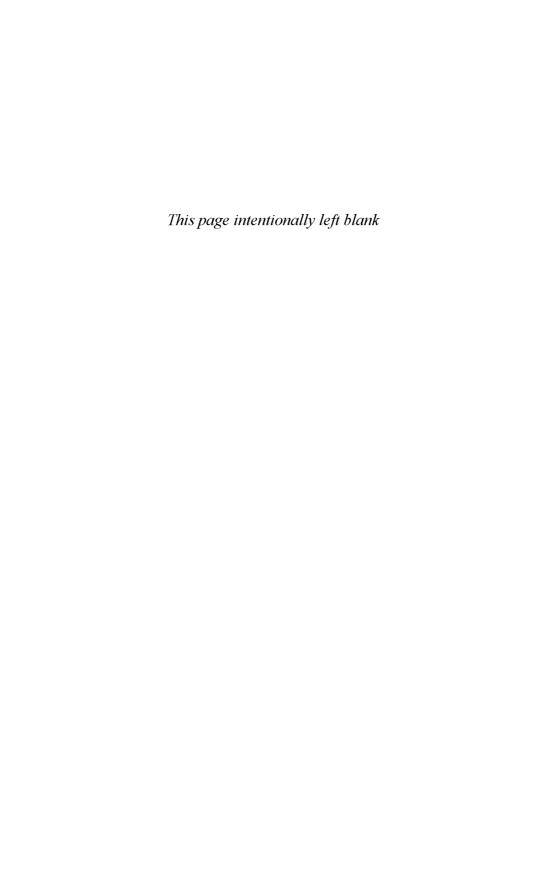
Kane, Christine Korsgaard, Craig Lawson, Joseph Mendola, Michael Meyer, James Montmarquet, Michael Moore, Paul Moser, Jeffrey Poland, Nelson Potter, William Prior, Elizabeth Radcliffe, Peter Railton, Amélie Rorty, Bruce Russell, Geoffrey Sayre-McCord, Robert Schopp, Stefan Sencerz, Michael Slote, James Sterba, Nicholas Sturgeon, Laurence Thomas, Judith Thomson, William Tolhurst, Nick Zangwill, Michael J. Zimmerman, and, especially in relation to action theory and moral psychology, Michael Bratman, Hugh McCann, Alfred Mele, and Raimo Tuomela. Through their writings, and in our occasional discussions, I have learned from the work of Philippa Foot, Richard Hare, Alasdair MacIntyre, Thomas Nagel, John Rawls, Thomas Scanlon, and Bernard Williams. Comments by two anonymous readers for the Press were helpful in the later stages of revision. Significant proportions of this book have also benefited from my interactions with the faculties and visiting philosophers at the University of Notre Dame, where I held a fellowship in 1987–88, Dartmouth College, where in 1994 I served as the Senior Fellow in the College's Institute on Moral Knowledge; and at Santa Clara University, where in the same year I served as Fagothey Distinguished Visiting Professor.

*Lincoln, Nebraska*                                                                                    R. A.
January 1997

*This page intentionally left blank*

# Contents

*This page intentionally left blank*

# Moral Knowledge
## and Ethical Character

*This page intentionally left blank*

# Introduction

## Four Dimensions
## of Ethical Theory

In the field of ethics the integration between theory and its everyday applications has commonly been close, and all the great moral philosophers have proposed both normative standards for day-to-day conduct and theories of the foundations of ethics. This book is written in the belief that ethical theory should not be isolated, either from concrete questions about what is right or wrong or from other parts of philosophy. From the rise of positivism in the 1930s, however, and at least into the 1960s, moral philosophy was preoccupied not only with ethical theory but, for the most part, with a fairly narrow part of it: the logic of moral concepts and, correspondingly, the uses of moral language. By the 1970s the tide had turned, and at least in the Anglo-American world applied ethics burgeoned in and outside academia and began to occupy a large part of the ethics curriculum in most departments of philosophy. But applied ethics requires theories and principles to apply, and at least since the early 1980s ethical theory in the broadest sense has regained strength and wider interest in the general field of ethics and particularly in philosophy.

    A distinctive feature of ethical theory now is its self-conscious realization that other areas of philosophy — especially epistemology and metaphysics, philosophy of mind and action, and philosophy of natural and social science — have much to contribute to it. This book is intended to advance ethical theory in a way that does full justice to the integrity of moral philosophy as traditionally understood but also brings to bear on its problems a perspective drawn from extensive work in epistemology and the philosophy of mind and action. The four areas of ethical theory to which the parts of the book are mainly addressed are moral epistemology, the metaphysics of ethics, moral psychology, and the foundations of ethics. What follows is a sketch of the territory covered.

## I. Ethical Knowledge, Intuition, and Moral Skepticism

Part I develops a position in moral epistemology. It explores the moral epistemology of some major ethical views, most notably those of Kant, Mill, and W. D. Ross; it

3

constructs a theory of the nature and grounds of the justification and knowledge that moral judgments, at their best, can express; it develops an account of ethical intuitions and their place in moral theory; and it shows how moral skepticism can be met by some of the arguments effective in dealing with skepticism about non-moral matters.

Chapter 1, "Internalism and Externalism in Moral Epistemology," shows how a major division in general epistemology applies in ethics. Epistemically *internalist* epistemologies ground justification and knowledge in elements accessible to reflection, such as sensory states and memory impressions; *externalist* epistemologies ground justification and knowledge in elements not thus accessible, such as the objective reliability of the processes causing belief: roughly, the relative frequency with which those elements lead to truth. Kantian ethics is argued to be internalist, in part because it grounds knowledge of basic duties in a priori reason; utilitarian ethics is externalist, grounding such knowledge in evidence concerning the consequences of acts. On one count, internalism is shown to be preferable: it better explains the importance of internal factors, such as motives and ideals, in appraising character. Good character is not a matter of what one's behavior causes; it is more nearly a matter of what causes one's behavior.

The task of chapter 2, "Intuitionism, Pluralism, and the Foundations of Ethics," is to develop an internalist, fallibilistic ethical intuitionism that can merit a serious place among contemporary ethical theories. The point of departure is Ross. In part through developing a more fine-grained account of self-evidence than Ross or other intuitionists provided, a reconstructed Rossian intuitionism is shown to be viable without some of the objectionable features Ross apparently attributed to it. As compared with his view, the more comprehensive intuitionism developed allows wider scope for moral reflection and inference as both correctives and systematizers of our moral standards. It is also free of some often alleged defects of intuitionism: arbitrariness, dogmatism, and an implausible philosophy of mind.

Chapter 3, "Skepticism in Theory and Practice," consolidates and defends the results of the first two chapters by articulating and, so far as possible in a single essay, meeting the challenge of skepticism. Just as skeptics about theoretical reason find a logical gap between our evidence for factual propositions and their truth, skeptics about practical reason find a logical gap between our moral reasons for action (or other reasons for action) and the overall moral goodness (or overall justification) of action. If we can vindicate common sense against theoretical skepticism (as I have tried to do in a series of epistemological works), we are at least in a good position to vindicate it against skepticism about the justificatory force of reasons for action. Skepticism about our behavioral realization of goodness has much in common with skepticism about our cognitive realization of truth, and important points responding to skepticism about the latter can be adapted to respond to skepticism about the former.

## II. Moral Concepts and the Natural Order

Part II is devoted to a theoretical (mainly ontological) profile of moral concepts, such as those of justice, virtue, and obligation, and to appraising the view that moral

concepts and properties merit a realist interpretation. Its results support the moral epistemology of part I. Ethical concepts are shown to be understandable through a certain kind of reflection on them and to figure in principles for which intuitions are the most important single source — though not the only source — of evidence.

The first chapter of part II, "Moral Epistemology and the Supervenience of Ethical Concepts," concentrates on the implications of the special dependence (the *supervenience*) of moral concepts on non-moral ones. A person is morally good, for instance, on the basis of non-moral, psychological features, above all basic desires, governing beliefs, settled intentions, and other elements of character. Does this imply that we can know such moral truths as that someone is a good person in the same way we can know psychological truths about the person? I argue that it does not, although it does imply (on plausible assumptions) that there is an objective, "factual" route to justification of certain moral propositions. Being governed by altruistic intentions is excellent evidence of one's moral goodness. This defense of objectivity is a positive result for ethical theory, but it is neutral between empiricism and rationalism and indeed can be interpreted so as to be compatible with non-cognitivism.

Chapter 5, "Ethical Naturalism and the Explanatory Power of Moral Concepts," picks up where the previous chapter ends: given the dependence of moral concepts on non-moral ones, are the explanatory uses of moral concepts essentially like those of (non-moral) explanatory concepts employed in the sciences, or are they perhaps sui generis? In answering, I explore the most widely held and most important naturalistic version of moral realism. On this realist naturalism, the reality of moral properties is shown by a proper account of their causal role in explaining human conduct. My own position accommodates what is most important in this one but yields a different kind of realism: I accept a naturalistic, broadly causal account of moral explanation but argue that this account can be cogent even if it does not construe moral properties themselves as natural. My strategy is to naturalize moral explanations without naturalizing moral properties.

## III. Moral Psychology and Ethical Character

Part III concentrates on problems of moral psychology. It is devoted to extending, to a wider range of human conduct, the internalist, objectivist perspective on moral judgment and moral epistemology laid out in parts I and II. At one end in human behavior there is the generally regrettable conduct associated with self-deception, weakness of will, rationalization, and compulsion; at the other end there is virtuous action in its many forms.

Chapter 6, "Self-Deception, Rationalization, and the Ethics of Belief," shows how moral appraisal can reach below the surface of behavior and even of consciousness, where reasons for action can operate unnoticed by the agent. Explanation — naturalistically understood as in chapter 5 — is contrasted with mere rationalization. A broad picture of rational action is presented, in which moral conduct has a place even when it is not a response to the agent's moral judgment or grounded in the agent's moral reasoning. The chapter serves in part as an antidote to the intellectual-

ist idea that moral appraisal and responsibility belong only to conduct that is grounded in explicit practical reasoning, consciously performed, or at least motivated by conscious factors. The ethics of belief is shown to be one domain in which moral standards reach beyond reasoned conduct and even below the surface of consciousness.

In Chapter 7, "Responsible Action and Virtuous Character," the conception of moral responsibility set out in chapter 6 is extended to character. To be sure, we do not have direct control over our character, and even our indirect control over it is limited, particularly in childhood. But there are several respects in which we can control it. I describe these and distinguish three major kinds of responsibility for elements of character. Here internalism again appears, this time in clarifying the basis not of justification but of responsibility. The idea, in outline, is that one is responsible for conduct only if it either traces to something internal, such as an underlying intention or decision, or there is at least some appropriate internal element accessible to the agent, such as a decision to check to be sure the fire is out, which would have altered that conduct. One might be responsible for a fire either because one decided not to check or because one *could* have decided to and did not. If moral responsibility is ultimately internal in this way, that is some confirmation of the general view, developed in many parts of the book, that moral concepts are in the main internally grounded.

The concluding chapter in part III, "Acting from Virtue," brings virtue ethics into the perspective of the internalist objectivism and multilayered moral psychology of the book. In the ideal cases, moral conduct is not merely acceptable on the surface but grounded in some virtue of the agent, as where a promise is kept not from fear of reprisal but from fidelity. But what is it to act from virtue, as opposed to merely doing what a virtuous person does? And can such action be explained in terms of reasons, in the way suggested by the explanatory account outlined in parts I and II? Moreover, can such action be understood in relation to rules or general standards of conduct, as Kant and Mill and the intuitionists believed? In response to these questions I develop an account of acting from virtue that represents it as a kind of action for reasons — reasons appropriately connected with the agent's virtues as elements in character. These reasons can be viewed both as virtue-guided and as rule-governed: there is no incompatibility here. This is not in the least to say that virtue ethics is dispensable in favor of rule ethics. Far from it. The concluding section shows how, if we distinguish (as is often not done) between the theory of moral obligation and the theory of moral worth, we can see that virtue concepts can be absolutely fundamental in the latter even if they are not fundamental in the former.

## IV. Reason, Judgment, and Value

Part IV, drawing on the first eight chapters, presents an account of reasons for action, with the central focus on normative reasons, the kind in virtue of which (as where a child will be burned to death if one does not act) one in some sense *ought* to do the deed those reasons support. Normative reasons may also, in a broadly causal sense, explain conduct, but this explanatory power is not their defining feature. Chapter 9

connects normative reasons above all with desire; chapter 10 connects them with moral judgment; and chapter 11 connects them with values. On the theory of the foundations of ethics presented in these chapters, values are experientially realizable ends that, in autonomous agents, can yield guiding practical reasons: reasons for both desire and practical judgment.

Chapter 9, "Autonomy, Reason, and Desire," explores autonomy as a test case for both instrumentalism and objectivism. For instrumentalism, the proper function of reason is to serve desire: its role is instrumental. But autonomy implies governing oneself in accordance with reason broadly conceived. This suggests that reason should guide our desires and thereby motivate us to act rationally. But how can reason, as a capacity for discerning truth — and in that sense an intellectual faculty — motivate us? A broadly Kantian answer is that reason provides its own motivational power independently of desire. I find neither this position nor instrumentalism fully satisfactory. I argue that desire can be educable by reason even if reason does not automatically govern it by independent motivational power. Autonomy is neither mere efficiency in satisfying one's unbridled basic desires nor an automatic result of properly using a priori reason; it requires an integration between reason and desire.

Chapter 10, "Moral Judgment and Reasons for Action," like chapter 9, brings together issues in moral psychology with questions in the foundations of ethics. It reinforces the idea that rational agents exhibit an integration between reason and desire, but it also shows how deficiency in such integration can impede moral motivation even where the use of reason issues in accepting a moral judgment that would ordinarily produce motivation. This deficiency is a challenging case, since moral judgment is in some sense action-guiding and has been widely (and plausibly) taken to entail motivation. A main task, then, is to assess *motivational internalism*: roughly the view that some degree of motivation is internal to moral judgment. In doing this I distinguish a range of motivational internalist positions, and I show how motivation can support moral judgment even if it is not internal to such judgment. If, however, moral judgments need not motivate, how can they be essentially practical, action-guiding cognitions? The answer is in part that they can express normative reasons for action even if they do not express motivational reasons; they can point our way and in that sense direct us, even if they do not by themselves push us along the path. This answer is developed in relation to an account of moral judgment and practical reasons that places both in the context of a theory of the relation between intellect and will.

Chapter 11, "Intrinsic Value and the Dignity of Persons," reinforces a major result of chapter 10. If the latter and several earlier chapters are correct in the view that there are objective reasons for action, then we might expect some of these to reside in things of intrinsic value, for instance in elements constitutive of human flourishing. Naturally enough, I begin with Aristotle, constructing an analogue of his argument to the effect that if desire is not to be endless and futile, then there is something good in itself. I proceed to explore various concepts of the good, such as Mill's hedonism and a more pluralistic view on which various non-hedonic experiences can be (intrinsically) good or bad. With some concrete examples in place, I argue for the possibility of knowledge or justified belief about value. Given this anti-skeptical result, together with the plausible assumption that if there is anything of

intrinsic value, then there can be objective reasons for action, I argue that there can also be knowledge or justification about reasons for action. Some of this knowledge or justification is specifically moral. If, however, there are potentially competing values, such as one's own pleasure and service to others, then knowledge of one's overall moral obligation may be at best difficult to achieve. I argue that practical wisdom is indispensable in approaching this problem.

The concluding chapter, "The Moral Justification of Actions and the Goodness of Persons," summarizes and further defends the overall theory of the book, makes additional connections among the chapters, and outlines a normative ethical view that extends the moderate intuitionism suggested in chapter 2. This Kantian intuitionism is shown to have significant advantages over Ross's position (though that remains among the best intuitionist views in the field) and to provide an interpretation of how a Kantian normative ethics can be made more concrete and plausible than it is if its first-order ethical principles are taken to be deducible more or less directly from the categorical imperative. The overall ethical theory that emerges from the book combines a version of moral realism with a moderate intuitionism: it is epistemologically internalist, normatively objective, valuationally pluralist, and qualifiedly naturalistic.

# I

## Moral Epistemology

*This page intentionally left blank*

# 1

## Internalism and Externalism in Moral Epistemology

A major division has recently come to the fore in epistemology. On one side are internalist theories, on the other, externalist theories, most notably reliabilism. The distinction has proved fruitful in understanding justification and knowledge, but its bearing on moral philosophy has yet to be systematically explored. If internalist and externalist views can provide general accounts of justification or knowledge, they may be expected to apply to the justification of moral beliefs and to yield criteria for moral knowledge. There is reason to think, however, that internalism and externalism in epistemology have even wider implications, for they reflect broad assumptions about normative status and may thus apply not only to beliefs but to other propositional attitudes and to actions. If so, we should be able to formulate not only internalist and externalist accounts of justified moral belief but also parallel accounts of moral rightness, moral obligation, and other moral notions applicable to actions. This paper will develop such accounts, contrast them, and compare them with other views, including *motivational* internalism and externalism, the positions commonly given those names in the literature of ethics. The final section will briefly assess the two moral epistemologies as candidates to serve among the basic premises of ethical inquiry.

## I. Internalism and Externalism in Epistemology

We must start with an explicit understanding of internalism and externalism in epistemology. To simplify matters, I shall first focus only on justification. Justification has, in any case, been central in the controversy between the two positions. The reason is not that knowledge is felt to be less important but that internalists and externalists differ in their accounts of knowledge primarily *because* of differences over justification conceived as a central element in knowledge. Moreover, if we can understand the controversy in relation to justification, we can readily consider its application to some aspects of the notion of knowledge, and I shall indeed do so later.

## Accessibility

No simple characterization of either internalism or externalism has emerged from the recent literature as acceptable to all the major writers on the subject; but there is a tendency for the former to bear the conceptual weight and the latter — which is historically a much more recent theory — to be described in contrast with it.[1] There is also a tendency — sufficiently widespread to give us a good starting point — for internalism to be characterized by appeal to some notion of introspective accessibility. Roughly, the central internalist idea is that what justifies a belief is internal to the agent in the sense that the agent is aware of it or can become aware of it by introspection or introspective reflection, where the relevant kind of awareness grounds knowledge of, or justified belief about, the justifier, or at least provides a ground for the capacity to arrive at such knowledge or belief on the basis of the awareness. Consider a paradigm case of justification conceived along internalist lines. My sensory experience of a white expanse might justify my belief that there is something white before me; and clearly I am aware of this experience, or at least can become aware of it by introspection. The implications of 'introspection' should not be exaggerated: to achieve the appropriate kind of awareness I need only attend to my visual consciousness; neither elaborate reflection nor inward searching is required.

Internalists need not require a capacity to become aware of the justifier under any particular description or concept, such as the sophisticated concept of sensory experience. All the agent needs is some notion of the justifier such that, understood in terms of that notion, it can be plausibly thought to justify. By contrast, an externalist view takes what justifies a belief to be something *not* introspectively accessible, and in *that* sense external to the subject. The most common candidate is (causal) sustenance or production of the belief by a reliable process, such as perception, which — with certain qualifications — may be held to generate more true than false beliefs. Such generation is not only something inaccessible to introspection; to know or justifiedly believe that a belief possesses it would take high-level psychological knowledge based on considerable experience.

There are, of course, stronger and weaker conceptions of the appropriate kind of accessibility. Hence, there are stronger and weaker forms of internalism, and correspondingly purer forms of externalism, depending on *how much* of what yields justification (if any of it) is or is not accessible in the relevant sense. Regardless of the strength of an internalist or externalist view, a plausible account of either sort should observe two distinctions. One concerns the closeness of accessible elements to consciousness; the other concerns the difference between justifiers that are *possessed* and thereby accessible and, on the other hand, justifiers which, though *obtainable* through reflection, are not actually possessed. Let me develop this contrast.

### Internal Justification

The first distinction holds between justifiers present in consciousness and those which, while not present, are internally accessible because they can appropriately reveal themselves in consciousness. In the first category are occurrent memory impressions that might justify a belief — say that I have met a certain person now

mentioned to me as a former acquaintance; in the second are beliefs one holds, such as might supply me with premises for some view I maintain but do not now have in mind. The memory impressions, like visual ones, are actually *in* my consciousness now, while the beliefs are simply such that I can *become* aware of them should I need to indicate what justifies the view I hold for which they express premises.

Phenomenologically, the difference here is between the way occurrent conscious elements justify and the way dispositional elements do so. The case is also instructive because it illustrates that introspective reflection need not be explicitly concerned with facts about oneself but may take the form of a search for grounds, where the *target* of the search is propositions that might be evidence, rather than any beliefs or internal attitudes. This evidential targeting is as it should be. When a belief one holds is what constitutes one's internal justification for another belief, the best way to become aware of this justifying belief can be to search one's memory for a premise for the belief whose justification is in question. Often, we can best call up our beliefs by considering appropriate content categories, such as that of reasons for holding a certain view, rather than by scanning the cognitive field directly.

The second distinction is between *justified belief* and merely *justifiable belief* (for the agent). Suppose I am confronted by a mathematical problem which I do not want to trouble to solve. Imagine that I believe it has a certain solution, *x*, but hold this unjustifiedly, on the basis of Tom's testimony, which, given my past experience with him, I should not accept. In disparate strands, I may nonetheless *have* evidence that *would* justify my belief if I thought about the matter and thereby saw how my several unintegrated evidential beliefs support the belief. Here introspective reflection might *produce* a justification for believing the solution is *x*; but it would not *reveal* a justification already possessed. Call the first kind of reflection *generative*, the second *revelatory*.

Both notions belong in an overall internalist account of justification; but our main concern is with actual, not potential, justification: with justification one *has* (or at least has quite readily accessible), not with mere justifiability, that is, justification one can *get*. For instance, we want to act *with* moral justification, so that we *have* a justification for the action at the time we act; we are not content to act with only the capacity to find a justification if we reflect on how to piece one together from our relevant beliefs. The deeds characteristic of a moral agent proceed *from* the agent's moral virtue or sense of moral duty; they are not merely capable of being shown to be consistent with morality by retrospective reflection. Though the distinction is not sharp, I want to concentrate on actual moral justification rather than on justifiability. Externalists as well as internalists can make this distinction. An externalism would express it in terms of what *does* reliably generate a belief, as opposed to what *would* do so under appropriate conditions, such as carefully reflecting on possible grounds for it.

## II. Internalism and Externalism in Ethics

In extending epistemic internalism and externalism to ethics, we might begin with the case of moral beliefs, even though action is ultimately of greater moral interest.

If we assume a cognitivist view, how should we characterize the two positions with respect to, for instance, a justified belief of what we might call the promising rule: that people should generally keep their promises?

## The Epistemic Internalism of Kantian Ethics

To focus the problem, consider Kant as apparently an internalist in his epistemology.[2] He might hold that by reflection — conceived as an internal even if not an introspective process — I can become aware of an a priori principle I believe, namely the categorical imperative, and my introspective awareness of its apparent self-evidence, or at least of my beliefs of certain premises for it, can justify believing the imperative. Now the categorical imperative *is* a justification of the rule that one should keep one's promises, or at least is so together with premises which, on internal grounds, I also justifiedly believe. Thus, I have justified beliefs of premises that justify the promising rule, and on the basis of those beliefs I justifiedly believe the rule.

Some internalists might go further: even a child not yet in a position to be justified in believing the categorical imperative might have an accessible and adequate ground for believing the promising rule, namely, a belief that all its elders have taught this (or perhaps simply a memory *of their having done so*). Nothing precludes an internalist from allowing the authority of testimony to justify in these ways, even when the belief is of an a priori principle. Indeed, if testimonial generation of a belief may be considered reliable, an externalist may also allow testimony to justify believing the promising rule. In the full-blown adult case, however, a natural line for an externalist to take would be broadly utilitarian: beliefs of the promising rule might, on a utilitarian externalism, be justified by virtue of being sustained by beliefs to the effect that generally promise-keeping conduces to happiness, where this is a reliable process because by and large such beliefs are *true*, in which case following the promising rule does indeed tend to have good consequences.

Supposing it is clear in outline how epistemic internalism and externalism apply to moral beliefs, there is no obvious extension to the notions of obligation or rightness. These, after all, apply primarily to actions, which are not truth-valued and hence not appropriate objects of knowledge or belief. Nonetheless, the extension is quite possible. Suppose we think of morally right actions as morally justified and of (morally) obligatory ones as such that no alternative to them is morally justified. We can then try to fill out the idea that justification *in general* is an internal or external concept, so that behavioral justification is parallel, in the sense that what justifies an action is, or is not, appropriately accessible. Once again, it is instructive to work with Kant and Mill.

For Kant, what is obligatory upon me is what I have a duty to do. But as Kant saw, having a duty to do something is one thing, doing it in fulfillment of that duty quite another. For him, at least the paradigms of acting morally — that is, of performing acts having moral worth — are deeds done purely *from duty*. If I keep a promise to Jack in order to fulfill my duty to him, my action in so doing is morally justified. It is justified by its being my duty. Now, what kind of access do I have to *what* justifies it? Can I, by introspective reflection, be aware of it as my duty or perhaps know, or at least justifiedly believe, that it is my duty? A Kantian might well say

that I have all these capacities. I can, for example, call to mind some memory of my promising to do it. I am also aware of my belief that one must keep one's promises; and Kant might well argue that normal adults can become aware of beliefs they hold which express premises adequate to justify that principle in turn.[3] Clearly, then, on a Kantian view one has appropriately internal access to what it is that justifies one's morally justified actions.

It might be objected that these points only show that an action's *conformity* with duty is accessible, not its being justifiedly performed, that is, its being done *from* duty and thereby having moral worth. To assess this objection, consider what is relevant to my justifiedly keeping the promise when I do the promised deed. This is determined by three factors: first, the *content* of my will, as reflected in what I *intend*, say to keep my promise as opposed to avoiding trouble with the promisee (something I may also take the deed to accomplish); second, my sense for the moral appropriateness of that content, for example of keeping my promise being a *duty*; and third, my actually keeping the promise *as* an expression of that content and from that sense, say *in order to fulfill my duty*, as opposed to acting on some self-interested purpose.[4] It will not do if, for instance, I unwittingly — and so merely in accordance with, and not from, the content of my intention — give Jack the money owed him, unaware that he is masquerading as the carrier collecting for my newspaper subscription, for which I happen to owe the borrowed amount. Now, a Kantian may plausibly claim that when I act with moral worth, I have access to the grounding elements, for instance the sense of my duty to do the thing. Granted, I may not have access to the fact *that* my sense of duty grounds the action, say through this sense's suitably sustaining it in a broadly causal sense. But to require that is to require access not only to *what* justifies my action but also to *how* it does so. A first-order internalism need not require that we have this more extensive access.

### The Epistemic Externalism of Utilitarian Ethics

For Mill, on the other hand, an externalist account is a better fit. Indeed, it is not clear how any truly consequentialist view *can* be internalist (unless, implausibly, the relevant consequences are themselves suitably internal). What I have a duty to do, in the sense of what I morally ought, on balance, to do, is to contribute optimally to happiness (for the sake of argument let us make it simply the happiness of people now in the world and include reduction of suffering under the general notion of contributing to happiness). Thus, what justifies keeping the promise is the total consequences of the act, something that is certainly not introspectively accessible to me. To be aware of these consequences, or to know or form justified beliefs about this totality, I need external evidence about the world.

One implication of this consequentialist position is that I can fail to do what is right, and indeed do something seriously wrong, even when from my point of view there is no reasonable alternative. It must be said immediately, then, that no sensible utilitarian would consider me blameworthy if, on excellent evidence, I kept a promise expecting good consequences and brought about disaster, for instance by returning a car to someone who then went berserk with it. But this does not entail that what I did was right or obligatory after all; it simply makes my act morally *excus-*

*able*, as opposed to culpable. One might also protest that there must be two senses of 'right' or 'obligatory' here. But a utilitarian need not multiply senses, so long as a doctrine of excuses is at hand to do the relevant work.[5]

Note, too, that a truly utilitarian doctrine of excuses would itself be externalist; for instance, a lapse from duty would be excusable only if it stemmed from traits, in a very broad sense including sets of desires and beliefs as special cases of traits, which in general conduce (sufficiently) to happiness. Negligence, for example, would not be excusing, whereas trusting a credible but mistaken witness often would be. It might be thought that a utilitarian should hold that for agents to be praiseworthy, blameworthy, or excusable is, roughly, for praising, blaming, or excusing them (respectively) to be optimific. But even from a utilitarian point of view, this assimilates being praise*worthy* to being *worth praising*. Specifically, it conflates criteria for *having* an optimific character with criteria for justifiably *ascribing* it. It might be argued that utilitarianism makes the conflation natural; but even if that is so, the theory need not carry such baggage.

On Mill's view, then — or at least on the moral epistemology to which a Millian utilitarian seems committed — the consequences of an action are parallel to the reliable production of a belief. Both are external in a similar way, and justification is grounded in the respective causal processes: truth-conducive generation in the cognitive case, happiness-conducive tendency in the behavioral case. Beliefs are justified by what produces *them*, actions by what *they* produce. The direction of causation is different, but justification is external in both cases.

It is true that Mill allows us to take commonsense moral principles for granted; but that is only because, through collective human experience, they are justified by their consequences. For Kant, the consequences of my action do not matter to its moral justification — though this is easily misunderstood. The point is that if keeping the promise is my duty, then the moral justification of my doing this does not depend on its consequences, including even my friend's doing evil with the returned car; but internalism does *not* imply that *calculation* of consequences is not relevant to what I should do. If I have *reason* to believe the friend will go berserk, I may have an internal ground, related to consequences, for breaking my promise. At least on one reading of Kant, his principles need not be construed in the rigoristic way he apparently interpreted them, and an *assessment* of consequences, as opposed to consequences themselves, can be a properly internal ground of moral justification for action.

If a Kantian internalism can be stretched to allow consideration of consequences, it can also be contracted. Suppose that through no fault of my own I am made to hallucinate keeping my promise to Jack, and that the incident is so vivid and so ostensibly normal that I justifiedly believe that I have kept the promise. Any plausible moral theory would say that I am excusable; but a strongly internalist theory might go further: in the circumstances, it might be claimed, I willed, or tried to do, something which I conceived as keeping my promise. Morally, then, I might be thought to do all I could. Certainly my volitional act was both justified and obligatory. The only question is whether I have done my duty. We should not say I have if such duties attach to the overt behaviors we undertake to perform or if, when I regain my senses and discover the car still in my possession, I will have the *same* duty I originally had. But it would be consistent with both internalism and some of

Kant to say that we should conceive our basic actions, and hence our basic duties, *internally*: on this view, what we directly undertake to do in promising is to move our wills toward the promised overt act; whether we perform that overt act, and thus fulfill our secondary duty, is up to nature. Thus, I promise to do what I shall justifiedly take to be, say, a return of a car; if I am unlucky and only hallucinate returning the car, that is unfortunate, but it counts nothing against my moral character. If I then discover the car, I acquire a new obligation, namely, to do what I shall justifiedly take to be returning *it*. This view is not quite Kant's; but, in a way that seems quite Kantian, it makes morality wholly internal, a thing of the mind and will.

## Virtue Ethics

It should also be instructive to comment briefly on the application of epistemic internalism and externalism to virtue ethics, conceived generically as the view that what makes an action morally right is its appropriateness to some moral virtue of the agent. Thus, a just action, in a given set of circumstances calling for justice, is to be characterized not as one which (say) treats similar cases similarly (even if it may turn out to be equivalent to such an act) but as one that is the sort of thing a just person would do in those circumstances. Now clearly, the notion of a virtuous person, say one who is just and kind, can be characterized either in terms of what is introspectively accessible, say the agent's beliefs, desires, and settled intentions, or by appeal to such external factors as traits whose manifestation in action tends to conduce to happiness.

It might be argued that we cannot be introspectively aware of, or know introspectively that we have, a *trait*, since this is a long-term feature a person can be known to have only through inductive behavioral evidence. A virtue ethics, therefore, must be externalist. But even if, on internal grounds, one cannot have knowledge of one's traits, it would not follow that one cannot, on such grounds, justifiedly believe one has a trait, as opposed to knowing this. Moreover, a cautious internalism in virtue ethics might only require that, for an act to be right, the agent is capable of being aware of a relevant underlying trait in a suitable way, for instance through an awareness of beliefs and intentions that manifest it. This might be conceived as an awareness of one's character *at the time*. Thus, if, as part of one's just disposition, one acts in fulfillment of an intention to treat two people equally, the moral rightness of the action may be grounded in the intention *as* manifesting the disposition, even if one cannot (as in the case of a very young child) *see* it as such a manifestation. One may need some rudimentary moral concepts, but perhaps not the concept of justice itself. Being just seems possible, in a minimal sense (and in ways described in chapter 8), for agents not yet capable of conceiving themselves as just, rather as — to recall Aristotle's point — habitually acting justly as a young adolescent may, is in some sense prior to becoming a just person.[6]

## III. The Range of Internalist and Externalist Moral Theories

We can better understand both internalist and externalist theories in ethics if we explore their connections with certain other important general categories. I have in

mind especially the bearing of the distinction on six issues: the contrast between objective and subjective interpretations of moral concepts; the debate between internalism and externalism in the theory of moral motivation; the rationalism-empiricism controversy; the problem of reductive naturalism; the question of noncognitivism; and the priority of moral reasons over other kinds. It would take a great deal of space even to describe these issues in detail. My aim here is simply to work with the issues in outline, providing only the amount of detail needed to clarify internalism and externalism in moral epistemology.

## Subjectivism

It is important to see that internalism is not necessarily a kind of subjectivism, in any of three important senses. First, it does not imply *psychologism*, the view that a moral judgment, such as that one must help one's aging parents, is about oneself, say about one's settled distaste for neglecting them. Second, it does not imply, and in plausible versions disallows, the *rejection of moral supervenience*, construed roughly as the view that (1) if two people are alike in their non-moral properties — especially in beliefs, wants, and experiences — neither can have a moral reason for a certain action which is not equally a moral reason for the other to do the same thing, and (2) they have such moral reasons in virtue of having certain non-moral properties. And third, internalism does not imply *subjectivistic relativism*: the view that there are no external moral reasons, in the sense of reasons for people to do certain things whether or not they are motivated, directly or indirectly, to do them — and thus subjectively so inclined.[7] An internalist might hold that there *are* reasons for us to keep our promises, even if nothing we want makes that seem to us to pay off; for we might have internally rooted justification for believing that we should keep them (whether we do believe it or not). A plausible internalism might hold that the existence of such a reason implies that we *can* be motivated to act accordingly; but this is a weak motivational requirement which seems plainly fulfilled by normal persons with respect to anything plausibly considered an external reason. An externalist, of course, can also hold these things. But externalism does not naturally take the form of a subjectivist view in any of the three senses. An externalist could, however, treat rightness as a matter of external facts about oneself, such as one's attitudinal and behavioral tendencies under complex hypothetical conditions, and could also deny that there are external reasons for acting. But the natural positions for externalists are not subjectivist and do not deny that the moral justification they are at pains to ground provides at least some degree of external reason for action in accordance with it.

## Motivational Internalism and Externalism

One might think that there should be more than a terminological coincidence between epistemic and motivational internalism, but there is little reason to believe so. The term 'motivational internalism' has been used in numerous ways, and here I take it broadly, as designating the view that believing (and so, normally, assentingly judging) that an action is one's overall (moral) obligation non-trivially entails having

some degree of motivation to perform it. Minimally, the degree of motivation is such that if one does the thing in question, this action is explainable as performed at least in part on the basis of the agent's belief that it was obligatory.[8] It is *natural* for an internalist about moral justification to hold this; for the thesis is especially plausible if one notes that the kinds of factors which morally justify an action — such as an awareness of one's having promised to do the deed — are well suited to motivate it and often *cited* in explaining it. But it is open to an internalist about justification to hold that the agent, particularly if not assumed to be rational, *need* not be so motivated. For an externalist about moral justification, on the other hand, motivational externalism may seem the likelier option. For what renders an action morally justified need not be in one's ken at all; hence, one could fail to be motivated to perform it. But this is poor reasoning if taken to show that *believing* one is obligated on balance to do something need not motivate one to do it; it argues only that *being* so obligated need not be motivating. Consider also an externalist, quasi-internal notion of excusability, a notion applicable where, though the action would do great harm, one believes, for excellent reason, that it is an overall duty and thus that one *must* do the thing. Doubtless a motivational internalism finds support in this sort of case, but the view is not required by anything in the (epistemic) externalist's premises. This is not to deny that motivational internalism is plausible on other grounds (not relevant here); my point is simply that neither internalism nor externalism in moral epistemology commits one to it.

## Rationalism and Empiricism

If Kantian moral theories are leading candidates for epistemically internalist accounts of moral justification, one might wonder if the distinction between these and externalist views of moral justification is neutral with respect to rationalism and empiricism. Doubtless it is not unnatural for an internalist about moral justification to be a rationalist, and (for reasons that will become apparent below) it is even more natural for a rationalist to be an internalist; but one could be an internalist or externalist virtue theorist without being a rationalist. Even a rule theorist might maintain that knowledge of the relevant rules is not a priori — either because no knowledge of non-formal principles is, or because moral knowledge (and justified moral beliefs) depend on experience, possibly including the deliverances of a moral sense conceived as an empirical faculty. Here we would also do well to consider Hume. In some passages, he seems to hold a psychologistic naturalism, on which moral appraisals of actions are really a kind of self-ascription.[9] Now one could be epistemically either internalist or externalist about what justifies the relevant self-referential beliefs, and so correspondingly internalistic or externalistic about what justifies moral action itself. Thus, my being morally justified in doing something might be a matter of my having appropriate grounds, internal or external, to believe that it would produce the appropriate moral feelings in me. Suppose, on the other hand, that we combined a Humean instrumentalism about rational action with an account of moral justification. On one version of this combined view, an action is morally justified when it is in one's long-run self-interest judged according to one's stable intrinsic desires viewed through one's own beliefs. Since one's beliefs and