

MORAL EPISTEMOLOGY

Aaron Zimmerman

CONTENTS

Acknowledgments	ix
Chapter 1 Moral epistemology: content and method	1
1.1 What is moral epistemology?	1
1.2 Socrates, Gettier, and the definition of “knowledge”	3
1.3 The standard method: levels of inquiry	9
1.4 Theories of moral knowledge: an overview	14
1.5 Chapter summary	22
1.6 Further reading	23
Chapter 2 Moral disagreement	25
2.1 Disagreement and skepticism	25
2.2 Moral contextualism	33
2.3 Chapter summary	40
2.4 Further reading	41
Chapter 3 Moral nihilism	42
3.1 Moral skepticism characterized	42
3.2 The death of god	43
3.3 Mackie’s queerness	47
3.4 Motives internalism	54
3.5 Reasons internalism	61
3.6 Chapter summary	69
3.7 Further reading	71

Chapter 4 The skeptic and the intuitionist	73
4.1 The Pyrrhonian problematic	73
4.2 Non-inferential moral knowledge	76
4.3 Chapter summary	103
4.4 Further reading	105
Chapter 5 Deductive moral knowledge	107
5.1 On deducing “ought” from “is”	107
5.2 In search of an epistemologically valuable moral deduction	113
5.3 Assessing the epistemological value of our deduction	124
5.4 Chapter summary	138
5.5 Further reading	139
Chapter 6 Abductive moral knowledge	141
6.1 Moral inference to the best explanation	141
6.2 Chapter summary	149
6.3 Further reading	150
Chapter 7 The reliability of our moral judgments	151
7.1 Acquiring moral concepts and exercising objectivity	151
7.2 Chapter summary	168
7.3 Further reading	169
Chapter 8 Epilogue: challenges to moral epistemology	171
8.1 Frege, Moore, and the definition of “immorality”	171
8.2 Common-sense objections to non-cognitivism	180
8.3 The Frege–Geach problems: semantics v. pragmatics	182
8.4 Non-cognitivist forms of validity	186
8.5 Chapter summary	193
8.6 Further reading	193
Glossary of philosophical terms	195
Notes	204
Works cited	219
Index	241

1

MORAL EPISTEMOLOGY: CONTENT AND METHOD

1.1 What is moral epistemology?

Roughly speaking, moral epistemology is the study of whether and how we know right from wrong. This colloquial characterization is only “roughly” correct because as epistemologists we are concerned with more than just knowledge, and as moral theorists our interests extend beyond mere right and wrong. So, for instance, once we know that a proposition is false, we know that those who believe it do not know it. But this need not end our critical evaluation of the believer or believers in question. We may still ask whether they were misled by what was otherwise excellent evidence, or whether, instead, they lacked on balance good reasons for believing what they did. We can ask whether they were led astray by “internal” problems like poor vision, bad memory, or defective methods of reasoning, or whether the cause of error was instead “external.” Were the lighting conditions poor? Were they, perhaps, deliberately tricked by a crafty adversary? And we can ask whether those who do not know those propositions they believe are generally reliable on the matters at hand or whether they quite often go astray. Should we trust them in the future, or have they earned our suspicion?

Similarly, once we know that an action is immoral we know that those who perform it thereby do something wrong. But we may still want to know what makes the action bad and how the nature of these particular instances of immorality ought to shape our critical reactions. Is the act uncaring or cruel? Is it unjust or unfair? Do the perpetrators deserve blame for what they've done, or did they have a good excuse? Are they just rotten people, or is this act of immorality an exception to an otherwise acceptable pattern of behavior?

In sum, epistemology is concerned with knowledge and the truth required for it, but it is also concerned with belief, justification, reasons, evidence, cognitive malfunction, proper functioning, reliability, and a host of cognate notions. Moral philosophy is concerned with morally right and wrong actions and the moral goodness and badness endemic to them, but it is also concerned with virtues and vices such as kindness and cruelty, fairness and greed; it explores the nature of moral obligations and rights, and the more or less general rules that we must observe to fulfill the former and avoid violating the latter; and it has a great deal to say about moral excellence and culpability and the attitudes, rewards, and punishments that we ought to level at those who act in morally laudable or blameworthy ways.

Moral epistemology thus explores the application of an enormous and somewhat varied set of concepts to a range of behaviors and institutions that are, if anything, even more numerous and varied. In consequence, the field is an exceedingly difficult one to circumscribe. So, for example, as moral epistemologists we are concerned with knowledge and ignorance regarding the morally right thing to do; the way to arrive at justified or well-grounded beliefs as to which actions and institutions are just; an enumeration of the sort of psychological maladies and sociological conditions that result in an improper appreciation of the viciousness of cruelty; and so on for each such combination of the many things separately investigated by mainstream epistemologists and moral philosophers. Knowing right from wrong is no more than a chunk of the iceberg's visible portion.

The looming multiplication of topics means that work in moral epistemology must of necessity be either wholly superficial or rather drastically limited in scope, and I aim to partially avoid the first of these vices by embracing the second to a greater degree than I would otherwise like. For this reason, among others, I will focus the discussion to follow on different views of basic moral knowledge and justification: our knowledge of the premises of those moral arguments we offer to one another in contrast with their conclusions; the justification with which we hold our most common

moral beliefs; the assumptions almost all of us make when we consider these matters. As a result, I will only touch on the difficulties endemic when we try to “weigh” conflicting considerations so as to arrive at an all-things-considered verdict about a particular scenario of moral interest. That is, I will have relatively little to say about which if any of the numerous mutually exclusive courses of action available to a person at any given time are the morally right or permissible options for her to pursue. And I will address only in passing our judgments about whether and how a person who is forced to weigh competing moral considerations can come to know her all-things-considered moral duty or what is all-things-considered the morally best course of action for her to undertake.¹ I won’t ignore these topics entirely, but because there is little current consensus on them, a survey of the difficulties involved is the only way to avoid an overly dogmatic presentation.

And there are two other advantages to this approach. First, it allows us to begin at the beginning with those moral beliefs and judgments that are conceptually and developmentally most fundamental. And second, it establishes a forum for the discussion of moral skepticism: the view that we cannot know right from wrong, either because evidence sufficient to support knowledge is not forthcoming, or because there are no moral facts to be known. Of course, by focusing on our most basic moral beliefs we are prevented from providing much if anything in the way of a guide to those already competent moral judges who are trying to figure out how to resolve the moral dilemmas (real or imagined) that they have encountered in trying to lead good lives. At best, we can hope to provide moral people with a better understanding of their knowledge, while supplying the ignorant and incompetent – who nevertheless possess the intelligence needed to follow our discussion – with an account of their deficiencies (cf. Aristotle, *Nicomachean Ethics*, 1095b [Aristotle, 1984, 1729–1867]).

1.2 Socrates, Gettier, and the definition of “knowledge”

One of the most intriguing questions in our field is whether virtue can be taught. Religious leaders and ethics professors, the writers of self-help books, and the principals of reform schools all claim to possess the kind of knowledge of virtue they must have if they are to teach it to their students or disciples. Can we learn how to be virtuous from a book? Does the acquisition of moral knowledge require training? Or is it, perhaps, largely innate? A number of different hypotheses come readily to mind. Perhaps some

exceptional people can teach themselves virtue from the Torah, the Bible, the Koran, the Hindu Vedas, the writings of the Buddha, or the works of the great moral philosophers. After all, the physicist and climate apostate Freeman Dyson is supposed to have taught himself calculus from an encyclopedia entry. Why can't the privileged few learn virtue in the same way? (When a sweet and loving child emerges from a horribly debauched environment, the attribution of "moral genius" is almost irresistible.) Perhaps, though, other people, indeed most children, really do need the flesh-and-blood instruction, training, and encouragement that good parents and teachers try to provide. Those of us who learned calculus in high school or college only managed to do so by asking a number of questions and solving a whole range of exercises and problems. Why should learning virtue be any easier? Indeed, it might turn out that some people lack the innate equipment to ever acquire virtue, no matter how much help they are given, and no matter how forcefully they are coerced into the pursuit. Surely, there are some kids – if only those with severe learning disabilities – who couldn't learn calculus if their lives depended on it. Mightn't virtue also be unattainable for some? Might some children – if only those with psychological problems of a rather drastic sort – be innately incapable of acquiring moral knowledge from even the most caring, perceptive, gifted communicator? When, as Shakespeare says, "good wombs have borne bad sons," must anguished parents find recourse in either the hospitalization or imprisonment of their children?

These issues have a long and storied history. Indeed, they were hotly debated in Athens over 400 years before the birth of Christ, during the time of Socrates, Plato, and Aristotle, the greatest philosophers of antiquity. The sophists were intellectuals who claimed to be able to teach virtue. But you cannot teach what you do not know. So do the sophists then know how we ought to behave? If they do, they would seem eminently qualified to lead. In the final analysis we are the state. So someone who knows what we should do must know what the state should do. Shouldn't the leader of the nation be someone who can articulate its proper mission and instruct us on the best means to its attainment? Shouldn't the true teachers of virtue then lead us all?

In Plato's dialogue *Meno*, Socrates discusses these issues with Meno, a Thessalian aristocrat. The best men – Pericles and Thucydides among them – sometimes produce bad sons. No doubt, they would have produced good children were it in their power to do so, and this suggests that virtue cannot be taught. But the sophist Protagoras was paid to teach virtue for more than

forty years. Surely, his ability to maintain a paying clientele over so long a period speaks in favor of his expertise. The question is therefore extraordinarily difficult to resolve. After pursuing several lines of attack, Socrates introduces a novel hypothesis. When a virtuous person does the right thing, this is not an accidental matter. In fact, we will only judge that someone is virtuous if we are confident that he will act justly in the absence of some unforeseeable accident or unlucky circumstance. But, for all that, we must admit that a good person will not be able to share his virtue with his children unless they are blessed by nature (or the gods) in some way or other. Perhaps then the righteous man has the *right opinion* as to how we should act, but he lacks genuine *moral knowledge*. Perhaps the true opinion explains his reliably virtuous actions and the lack of knowledge explains his inability to communicate virtue to his offspring. “If it is not through knowledge, the only alternative is that it is through right opinion that statesmen follow the right course for their cities” (99b–99c).

But what is the difference between knowledge and true belief? Though Socrates claims to know very little, he tells us that he is absolutely certain that there is a difference between these two states of mind (98b). The preceding discussion gropes toward an account.

True opinions, as long as they remain, are a fine thing and all they do is good, but they are not willing to remain long, and they escape from a man’s mind, so that they are not worth much until one ties them down by giving an account of the reasons why. After they are tied down, in the first place they become knowledge, and then they remain in place. This is why knowledge is prized higher than correct opinion, and knowledge differs from correct opinion in being tied down. (97a–98)

Socrates’ tentative claim here is that one’s correct opinion on some matter can “become” knowledge once one has acquired some grasp of the reasons why it is true. Having an explanation of some fact solidifies or deepens one’s conviction in its truth, helps one remember it, and enables one to communicate it to others. Knowledge, according to Socrates, differs from true belief in all these respects; and if we keep these differences in mind, we will credit virtuous people with certain true opinions as to what is just and good, but we will persist in denying them any moral knowledge.

There are, I think, few contemporary thinkers who would endorse the hypothesis left on the table at the *Meno*’s end.² First, for a person to actually

be virtuous, she would seem to need more than a set of correct opinions on moral matters. A virtuous person must be compassionate, loving, brave, and kind; and it is unlikely that these largely emotional capacities can be correctly identified with the possession of moral views that just happen to be true. Perhaps, as we will discuss, there is a kind of wisdom that really is sufficient for virtue, but wise people have more going for them than true opinions. Second, it is far from obvious that those who know something must be able to teach it to all those they wish to instruct; so why deny the virtuous moral knowledge simply because they do not invariably teach their children to be good?

Third, it is not at all clear that the children of the virtuous are ignorant of virtue. Mightn't they know how they ought to behave, and yet fail to act as they know they ought? Perhaps the virtuous do succeed in providing their children with moral knowledge of a kind, and yet moral knowledge is insufficient for moral action. This last question would hound Socrates throughout his days and trouble Plato and Aristotle a great deal. Indeed, as we will see, it remains a central topic for moral epistemologists working today.

Still, even if we reject Socrates' tentative explanation as to why virtue is not so easily inherited from the virtuous, the distinction behind his hypothesis holds considerable interest. Can someone have the right opinion on moral matters and yet fail to possess moral knowledge? To evaluate Socrates' positive answer to this question, we need to assess his description of knowledge. And it turns out that this is no idle enterprise, as for thousands of years philosophers looked to his remarks for guidance. Mightn't Socrates have provided us with the materials we need to *define* "knowledge" in terms of its difference from opinions that just happen to be true? Can we use his comments to formulate a relatively short, informative account of the kind of thing "knowledge" denotes, an account that might supply someone ignorant of this expression with an adequate grasp of its meaning?

Well, though Socrates doesn't offer us anything like a definition of "knowledge" in the *Meno*, theorists inspired by him did.³ For instance, Roderick Chisholm would go on to define "knowledge" as a true belief held with adequate evidence (Chisholm, 1957), and A. J. Ayer would define it as conviction in some truth of which one has the right to be certain (Ayer, 1956/1990). Though interesting in their own right, the details of these accounts needn't detain us here. For, in a landmark work, Edmund Gettier (1963) was widely credited with refuting them – and all similar analyses – by supplying compelling examples of people who seemingly

fail to know facts that they nevertheless justifiably believe. In the wake of Gettier's essay, the quest to provide Socratic definitions of "knowledge" has gradually ground to a halt.

Suppose, to gloss one of Gettier's examples, that the boss tells me that I am getting a promotion. And suppose, that as I know I only have \$20 left to my name, I quite reasonably infer that the man getting the promotion only has \$20 left to his name. Indeed, though it is exceedingly coy, we might suppose that in response to an inquiring colleague I go on to assert what I have here inferred. "Who is getting the job?" he asks. "Well, I'm not at liberty to disclose his name," I answer, "Though I will say that he's someone who's been reduced to his last \$20." Now imagine that for some reason or other the boss has lied to me, and it is really Jones who is getting the bump-up. But yet imagine too that, as chance may have it, Jones is in the exact same financial position as I am. He also has just \$20 left to his name. Then I will be justified in believing that the man being promoted has only \$20 left to his name, and this belief will be true, but the accidental nature of its truth will dissuade most of us from thinking of it as knowledge. I'm right in believing that the guy getting the promotion only has \$20, but I do not know that this is so.

Now if testimony can provide us with good evidence, I have good evidence for what I believe in the case at hand; and, again, what I believe is in fact true. But my belief still fails to constitute knowledge; so Chisholm's account must be rejected. And I surely have the right to trust those – like my boss – whose testimony I have no reason to doubt. So Ayer's account cannot be quite right either. If our ordinary thinking about the matter is to be respected, knowledge cannot be equated with justified, true belief.

Philosophers responded to Gettier's examples by requiring, in one way or another, that it be *no accident* that one's belief is true if it is to be properly characterized as knowledge, with Alvin Goldman (1967, 1976, 1986) and Robert Nozick (1981, ch. 3) providing what were perhaps the most widely discussed analyses of this kind. But these and all subsequent attempts to reflect on our ordinary thinking about knowledge so as to arrive at a relatively simple, interesting, explanatory account of the phenomenon failed to secure widespread acceptance (Shope, 1983). For this reason, among others, many contemporary theorists now find themselves agreeing with Timothy Williamson's (2000) claim that "knowledge" expresses a relatively simple concept that resists reductive definition or analysis.⁴ This isn't to say that epistemology is now a dead discipline. We can still investigate knowledge in general, and moral knowledge in particular. But it now seems as

though we are going to have to accomplish this task without the aid of a widely accepted definition of “knowledge.”

Indeed, many epistemologists now draw similar conclusions when they turn to cases in which someone forms a false opinion through no fault of her own – cases in which even the most careful investigator would be led into error. Suppose, again, that the boss has lied to me, and I am wrong in thinking that I will be promoted. Still, so long as I have no reason to suspect that the man is lying, my confidence in the promotion is entirely *reasonable*. I am, we would say, entirely *justified* in drawing a false conclusion in this context. But in what does having a justified belief consist? Can we adequately define “justification” as it is used in cases such as these? Can we supply a brief, insightful account of the phenomenon that “justification” is used to denote, an account that might supply someone ignorant of this term with an adequate grasp of its meaning?

Again, recent history is littered with proposals that have yet to secure agreement. We have accounts drawn from the definitions of Chisholm and Ayer cited above: someone is justified in believing something just in case she has adequate evidence of its truth; someone is justified in believing something if she has a right to be convinced that it is the case (cf. Feldman and Conee, 1985; Pollock, 1986). But philosophers often classified as “naturalists” or “externalists” have argued for the inadequacy of these equations. Instead, they suggest, one is justified in believing something when one’s belief is generated by a reliable mechanism or procedure (Goldman, 1986). Or perhaps one is justified in believing something just in case one’s belief results from the exercise of an epistemic skill or ability, the proper functioning of a psychological module or set of modules, or the expression of some epistemic virtue (Sosa, 1980, 2007; Greco, 1993; Zagzebski, 1996).

The details of these debates needn’t concern us here. We need only register the suspicion, most expressly voiced by William Alston (2005), that there is no single concept associated with “justification” even when it is limited in its application to opinion, credence, or conviction. Perhaps, that is, our differing evaluations of the ways in which we sometimes go right and sometimes go wrong when forming, maintaining, and revising of our beliefs track distinct properties that are nevertheless all important in one way or another. Again, it seems we must set about examining the various ways in which our moral beliefs might be said to be justified or unjustified without first having anything like a definition of “justification” in place.

1.3 The standard method: levels of inquiry

The search for definitions has proved inconclusive at best. So let us turn away from the analytic project and note the sense in which theorizing about knowledge in general – and moral knowledge in particular – must begin with observation of human behavior and human psychology.

We can start with an examination of moral theories. Surely, if we are to develop a view as to how people and institutions ought morally to act we must have on hand some description of the ways in which they actually do act. How do we behave? What causes or explains our acting in these ways? Which behaviors are constant across space and time, and which behaviors vary?

Even “theorists of the ideal” bent on describing how a moral utopia would function must concern themselves with the best that common observation, psychology, and sociology have to offer. After all, the imagined utopia is supposed to be a community of people, not angels. If an imagined ideal state is to represent a genuinely human possibility, its conjurer must take into account our distinctively human abilities and frailties (Flanagan, 1993).

Zero-level moral inquiry: a description of the motives and behaviors of people and institutions.

The next step on the way to a moral theory consists in a description of our *critical practices* with regard to the actions and motives we’ve identified at the zero level. Which actions do we think of as morally right and which do we think of as morally wrong? Which institutions and practices fill us with moral condemnation, and which agents inspire our awe and admiration? Which of these evaluations, criticisms, and emotional reactions vary across time and differ between geographically isolated communities – which even vary between different people in a given community – and which exhibit greater constancy?

When a philosopher writes of her “intuition” that the behavior of some agent in a hypothetical scenario is wrong, unjust, or blameworthy – if she says that someone would be wrong to push a fat man in front of a trolley to save the lives of those it would otherwise trample (Thomson, 1976, 2008), or that a person is morally blameworthy if, because he doesn’t want to ruin his clothes, he blithely walks past a child drowning in shallow water (Singer, 1972) – she is perhaps best understood as engaging in this kind

of first-level moral theory. If everyone shares her intuitions, she will have described critical practices and judgments that are universally engaged in and assented to. But even if her intuitions do not extend beyond herself and her readership, her effort will be of some utility, as she will have helped articulate the evaluations of that particular community.

First-level moral inquiry: a description of the distinctively moral evaluations (e.g. criticism and praise) that people level at the motives and behavior of people and institutions.⁵

Still, moral theory *proper* begins where first-level moral inquiry ends. Once we have identified the critical practices of a community we can then try to critically evaluate those critical practices themselves. Evaluative practices change, so we don't have to continue resenting, condemning, and calling "wrong" all and only those things that we have called "wrong" up to this point in time. Thus, we can ask, is there any sense to be made of our often pre-reflective (Haidt, Koller, and Dias, 1993; Haidt, 2001) use of "right" and "wrong"? Do we apportion moral praise and blame in conformity with a set of tacit rules, or will our intuitive moral classifications elude even the most careful attempts at codification? If there are rules that can be extracted from the pattern of reactions we've identified at the first level, what if anything can be said in favor of retaining them?

Second-level moral inquiry: an evaluation (critical review) of the distinctively moral evaluations uncovered by first-level moral inquiry.

I've said that proper academic moral theorizing takes place at the second level. But there are philosophers who mingle the first and second levels in the style of an Emily Post by recommending or endorsing every moral evaluation attributable to common thought. Still, where this approach does not result in immediate incoherence – an incoherence mirroring the practices being described – it ensures for itself a conservative outcome. It leaves the theorist thinking that a radical moral theory (such as Singer's strain of utilitarianism) is obviously untenable simply because it fails to describe "our" moral verdicts.

The need for second-level inquiry presses itself upon us with particular force when we have identified a set of evaluative practices that differ significantly from those we find ourselves embracing. Thus, when she is exposed to a liberal society, a member of an orthodox Jewish or Muslim community

who is capable of the relevant form of reflection may wonder whether chastity of the sort she is practicing really is a virtue – as her parents, teachers, and friends maintain – or whether, instead, her community is wrong to condemn, look down upon, and call “wrong” all physical contact between unmarried, unrelated people of differing sexes.

Of course, to evaluate a set of evaluative practices a person must use some means of evaluation, and our subject’s critique will have its greatest impact if she assesses her community’s moral view using that community’s own concepts and methods of criticism. Consistency or coherence is therefore one of the most powerful tools employed in moral inquiry conducted at the second level (Aristotle, *Eudemian Ethics*, 1214b1–1215b1.15 [Aristotle (384–322 BC/1984)]). So, for instance, almost all of us agree with Singer (1972) and Peter Unger (1996) that it is morally wrong for someone to walk past a child drowning in shallow water if the passerby’s only motive for not wading in is a desire to preserve her expensive footwear. But most of us also think that it is morally acceptable for those so inclined to buy expensive shoes for the pleasure that it brings. So suppose someone surfing online happens to have the Oxfam and Christian Louboutin websites open before her: clicking on link A will help save human lives, whereas clicking on B will garner nothing more significant than a fancy pair of high heels. Those I have surveyed tend to think that though clicking on B is somewhat selfish, it is not immoral or morally wrong. But is there a relevant difference between clicking on B and walking past a drowning child when the harm that could be prevented is equivalent, and the goal of action in both cases is the possession or preservation of the same pair of shoes? Is there a difference significant enough to warrant our thinking that the passerby acts immorally, but the somewhat self-centered shopper does not? Perhaps there is. But if we cannot articulate such a difference, the critical scheme we’ve identified at the first level will seem impugned by a second-level review, and rational inquirers will feel forced to change their ways of thinking (Berlin, 1955–56; Nozick, 1974, 277–79). We might conclude, with Singer and Unger, that the requirements of benevolence are more stringent than we typically assume, or we might instead decide that indifference to human suffering is in fact morally permissible (Thomson, 1971).

But though it is a powerful tool, coherence needn’t be thought of as the only consideration employed in second-level moral inquiry. Perhaps a minority community of mavericks can reflect on the moral system operative in their society – and the moral “intuitions” common among its populace – and recognize that though these practices are in some sense coherent,

they are leading to more misery and suffering than is necessary, and are morally objectionable on these grounds alone. Perhaps people are being uniformly criticized for harmless pleasures, or condemned as “unnatural” for exhibiting behaviors to which no one should object. I have in mind here the intellectual currents that gave rise to Socrates’ criticisms of uncritical religious piety and the British Sentimentalists’ rejection of the monkish “virtues” of self-abnegation. But left-leaning moral theorists might also include the birth of the young Hegelians and Marx’s critique of the now prevalent bourgeois attitude toward the distribution of property.

John Rawls’ (1971) less ambitious criticisms of this same socioeconomic structure might provide another example. For though Rawls does use the coherence of his principles with our considered judgments about particular cases as a tool in arguing against the kinds of severe inequality now prevalent in nations the world over, he also invokes seemingly independently grounded psychological and sociological hypotheses. These include propositions about the nature of envy, the conditions necessary for social stability, and metaphysical claims about the separateness of people.⁶

Of course, attempts to undermine or radically revise common morality are often horribly mistaken. (As, I would argue, were Friedrich Nietzsche’s [1886/1966] reactionary attacks on democratic ideals.) But if a thinker not wholly of his time can mount a successful second-level critique, moral theorists can use better or more accurate schemes of evaluation to effectively evaluate worse or less accurate schemes of evaluation, where “better” and “worse,” “accurate” and “inaccurate” are not measured in terms of coherence alone.

Epistemological inquiry exhibits the same three-level structure present in the moral case. It properly begins at the zero level with a description of the beliefs we actually hold: What do we believe? What causes or explains our holding these beliefs? On which issues do we agree and on which do we part ways? How have our beliefs changed across time, and how do they differ among geographically and culturally distant peoples?

It then proceeds to a first-level description of our evaluation of these beliefs and believers. Under what conditions will we say that agents “know” what they believe? When do we say that though they fail to know they are nevertheless “justified” in believing what they do or that they fail to know through no fault of their own? And when do we say that people or their opinions are “irrational,” “unjustified,” “gullible,” “hasty,” “dogmatic,” “unreliable,” or “overly skeptical”? Which of these evaluations are constant across the speakers in our community and on which do they differ? Which

vary across communities or within communities across times (Weinberg, Nichols, and Stich, 2001)?

Finally, second-level epistemology consists in a critical evaluation of our first-level critical evaluations. Is there, as skeptics often allege, some deep-seated incoherence in our attributions of knowledge? Do people fail to know much of what we credit them with knowing? Or are we perhaps overly stingy with application of “knows” and similar terms? In the end, we might conclude that “common sense” should be left alone. Or we might decide that the skeptics are right. But we might discover that the critics have gone entirely the wrong way in rejecting our ordinary attributions of knowledge, when in fact harsh Cartesian strictures have so infiltrated the population at large that genuine possessors of knowledge are commonly being denied the status that is their due (James, 1897/1956, 18).

Properly conceived, distinctively *moral* epistemology results when links are established between epistemological and moral inquiries conducted at one or more of the levels that we have identified. Suppose, for instance, that some of the beliefs that epistemologists enumerate in their zero-level inquiries have distinctively moral content. Suppose, that is, that alongside our scientific belief that $E = mc^2$ and our belief that the Earth is over four billion years old, we must account for our moral belief that greed is a vice that ought to be discouraged, and our conviction that infidelity is immoral. And suppose that the very evaluations we level at those of our beliefs that have non-moral content are also appropriately leveled at beliefs with moral content. Suppose, that is, that as we actually think of properly instructed children learning and therein coming to know that the Earth is over four billion years old, we can similarly speak of children learning and therein coming to know that it is immoral to be selfish, mean and unjust. If these suppositions are made, then first-level epistemology will contain first-level moral epistemology as a part, and a second-level epistemological project can aim its sights at those moral beliefs that we’ve uncovered at the first level. Perhaps though we claim that properly raised children come to know that it is immoral to be cruel and selfish, they can’t really come to know any such thing, as our assignments of knowledge in these cases are incoherent or just plain false. Or perhaps (to go the other way) though many people in our society think that we are not yet justified in drawing definitive conclusions regarding the exact circumstances in which the abortion of a pregnancy is morally impermissible, we are already in a position to figure this out. In sum, if we have *moral beliefs* that are relevantly like our non-moral beliefs, moral epistemology is a legitimate line of inquiry.

Working from the moral to the epistemological, we can establish the same claim by showing that moral judgments and beliefs relevantly similar to our non-moral judgments and beliefs figure among the evaluations we uncover when pursuing our first-level moral inquiries. There is, of course, no doubt that some of what P. F. Strawson (1962) calls our “reactive attitudes” either stop short of belief or go beyond it. We often respond to perceived immorality with anger, indignation, loathing, resentment, guilt, and disapprobation, and react to what we take to be moral kindness and sacrifice with admiration, approval, love, and pride. But we also say that certain actions are morally wrong and others morally right, some morally obligatory and some impermissible, some vicious and other virtuous. If when we say these things we are at least sometimes expressing moral beliefs rather than (or in addition to) morally fraught emotions and sentiments, and we can meaningfully apply epistemological concepts to the states of mind in question, we can then inaugurate the kind of inquiry distinctive to moral epistemology by asking whether we know that certain acts are wrong, whether we’re justified in believing that others are virtuous, and so on.

Moral epistemology is, thus, a fairly natural line of inquiry to pursue. We should note, though, that there are theorists who deny the very existence of beliefs with distinctively moral content. For instance, some philosophers equate moral knowledge with certain kinds of value-neutral knowledge; and some philosophers argue that there are no moral beliefs, or insist that if there are moral beliefs, these states of mind are so unlike our non-evaluative beliefs that they cannot be coherently assessed in epistemic terms. Those interested in these challenges to the very coherence of moral epistemology will find them discussed, at some length, in the final chapter.

1.4 Theories of moral knowledge: an overview

Epistemology is often divided into two questions: the “what” and the “how” (e.g. Sosa, 1980):

- the “what” question: what do we know?
- the “how” question: how do we know it?

Distinctively moral epistemology might then be divided into:

- the “m-what” question: what moral facts do we know?
- the “m-how” question: how do we know them?

The moral skeptic issues rather stark responses to these traditional questions. The “m-what” question asks what we know of morality. “Nothing,” the skeptic answers. The “m-how” question asks how we know those moral facts we know. “We don’t,” is the skeptic’s reply. According to the skeptic, we have no moral knowledge, and since we have no moral knowledge, there is no such knowledge to explain.

Let’s call attempts to respond to skepticism *defensive epistemology*. In partial contrast, we have *constructive epistemology*, which first assumes the existence of some kind of knowledge and then tries to explain the knowledge that has been assumed. Again, if we apply this distinction to distinctively moral epistemology, we arrive at:

- *defensive moral epistemology*: an attempt to show that we have moral knowledge;
- *constructive moral epistemology*: attempted explanations of the moral knowledge we have.

Which project should we pursue here? Should we try to answer the moral skeptic first, and assure ourselves that we have some moral knowledge? If we do this, we can address our “m-how” question directly. We can say, “We have shown that we know that this kind of thing is immoral and this other kind of thing is not,” and then ask, “Now how do we know these facts?”

But we needn’t follow this approach, as we might skip over defensive epistemology either for once or for good. After all, most of us are not moral skeptics. And those of us who believe in moral knowledge are genuinely curious as to how we – and those around us – came to know that, say, theft is wrong and dishonesty a vice. If we put this curiosity on the back burner, and begin by trying to prove that, at some point or other, we really did come to know these moral facts, we may never get around to positive explanations of the moral knowledge most of us think we have. In this event, moral skepticism will have proved itself a nuisance, a source of philosophically debilitating distractions.

Nevertheless, despite this very real danger, I adopt a “skepticism first” approach in the pages to come. I do so for three – somewhat interrelated – reasons. First, moral skepticism differs from other common forms of skepticism in its depth, popularity, and consequences for ordinary thought and practice. Admittedly, philosophers still discuss skepticism about other minds and skepticism about non-observational knowledge more generally. Even full-blown skepticism about the external world has yet to disappear from

view. But most epistemologists working in these areas are fairly certain that we do in fact know that there are other people who think and feel roughly as we do; that we do in fact know that objects will fall when dropped near the Earth's surface; and that those of us with hands do in fact know that we have these appendages. The bulk of the literature in these other areas of epistemology is therefore justly dedicated to the constructive task of explaining just how we know these things. In contrast, there still exist many philosophers today who doubt that there are moral facts or doubt that we have any knowledge of morality. Moral skepticism therefore remains a live option – a “relevant alternative” to non-skeptical accounts of moral judgment.

And there is a second reason for beginning with a discussion of moral skepticism, a reason having to do with the interrelation of the defensive and constructive projects we have described. It is my sincere hope that when you finish this book you will have killed two birds with one stone, as responding to skeptical arguments against the very existence of moral knowledge will have provided you with an excellent understanding of the kind of moral knowledge we do in fact have. By seeing what must be done to answer our “m-what” question in a non-skeptical manner, we will have arrived at a better sense of the various non-skeptical answers to our “m-how” question that might be defended.

I will admit, though, that my attitude toward these matters would differ if I thought the moral skeptic's challenges could not be answered. But, in fact, I think exposure to the best arguments for moral skepticism, and the best responses that can be given to those arguments, will leave you convinced that we do in fact have a substantial body of moral knowledge. Indeed, this is the one prejudice I have indulged in the pages to come. While I have done my best to present debates over whether and how we know right from wrong in a neutral, even-handed way, I have done nothing to hide my sense that even the strongest arguments for moral skepticism still come up short. If I am right about this, we needn't settle for the conditional question, “Supposing we do have moral knowledge, what is it like?” We can instead investigate a body of knowledge we are fairly confident exists. We can ask, without qualification, “How is our very real moral knowledge best described and explained?”

Nevertheless, before plunging into an examination of moral skepticism, I want to give you some sense of the field's more constructive aspects. This will help dispel the impression that moral epistemologists are unduly fixated on skepticism, and give us a glance at the many interesting debates that continue to litter the field.

To begin constructing even the most rudimentary theory of moral knowledge, we must have in mind some item of moral knowledge that needs explaining. We must assume, that is, if only for the sake of discussion, that someone, somewhere, knows some moral fact. But which example should we choose? Should we examine someone's knowledge of a *general* claim, like the claim that adultery is wrong? Or should we instead look at knowledge of a *particular* claim, say the average man's knowledge that Senator Ensign acted immorally when cheating on his wife with his aide's spouse? Suppose we split the difference. We begin with knowledge of a particular case and examine the claim that knowledge of particulars always stems from knowledge of some more general moral principle.

We can take as our point of departure Charles Dickens' widely beloved *David Copperfield*, a work that contains a host of details drawn from the author's own troubled childhood. Early on in the tale, Copperfield is sent to Salem House, a boarding school run by the despicable Mr Creakle. Copperfield's assessment of the man's cruelty bears repeating in full.

I should think there never can have been a man who enjoyed his profession more than Mr. Creakle did. He had a delight in cutting at the boys, which was like the satisfaction of a craving appetite. I am confident that he couldn't resist a chubby boy, especially; that there was a fascination in such a subject, which made him restless in his mind, until he had scored and marked him for the day. I was chubby myself and ought to know. I am sure when I think of the fellow now, my blood rises against him with the disinterested indignation I should feel if I could have known all about him without having ever been in his power; but it rises hotly, because I know him to have been an incapable brute, who had no more right to be possessed of the great trust he held, than to be Lord High Admiral, or Commander-in-chief: in either of which capacities, it is probable that he would have done infinitely less mischief. (Dickens, 1849–50/1997, 85–86)⁷

We will have ample opportunity to return to this case in what follows. For now, I want to consider, in brief, the frame of mind Copperfield evinces as he comes to appreciate the sadistic motives that led Creakle to inflict so much suffering on his wards, and moves from this realization to a belief in the headmaster's cruelty and immorality. We are assuming, for the sake of inquiry, that Copperfield here *knows* that Creakle acted in a cruel and immoral fashion. But just *how* does Copperfield manage to acquire this knowledge?

According to one school of thought, Copperfield can just *see* that Creakle is acting immorally. That is, Copperfield knows that Creakle's actions are wrong without drawing this conclusion from distinct premises of any kind. Consider that many (if not most) epistemologists think we can know what we are feeling and thinking by simply introspecting. I don't need to infer that I am in pain from distinct premises or supporting evidence. Instead, the knowledge is wholly *non-inferential*. A smaller but still substantial number of epistemologists think of our perceptual knowledge in the same way. I just know by looking that there is something red in front of me. I needn't argue my way to this conclusion from other things that I know, or supply additional reasons to support my perceptual belief. I can just see that there is something red there. A growing number of philosophers would add certain instances of moral knowledge to this class of judgments. According to this group, just as I know by simply looking that there is something red in front of me, Copperfield knows by simply observing Creakle's actions that the headmaster is acting immorally. We will discuss this perceptual model of moral knowledge in section 4.2.

An instinct more prevalent among moral epistemologists is to depict Copperfield as inferring Creakle's immorality from the value-neutral facts of the case. That is, we might think of the boy as: (a) first using his observations of Creakle's behavior, and his observations of its effects on the students, to verify that the headmaster enjoyed causing the boys to suffer; and as then (b) using his knowledge of a general fact detailing the "defining essence" of cruelty to infer that the man acted cruelly. We can model this two-step inference as follows:

- 1 Creakle enjoyed causing others to suffer.
- 2 Someone acts cruelly just in case he takes enjoyment in causing others to suffer.

Therefore,

- 3 Creakle acted cruelly.

Now let us suppose, if only for the sake of discussion, that Copperfield's observations really do provide him with knowledge of the ways in which he and the other chubby boys have suffered at Creakle's hand; and that these same observations provide Copperfield with knowledge of the frame of mind in which his teacher delivered the lashings. These assumptions allow us to conclude that premise (1) is not only true, but that it is also *known* to be true by Copperfield. And yet it would seem that knowledge of premise (1) is not yet moral knowledge. Instead, it is the kind of knowledge philosophers and psychologists label "knowledge of other minds" (see 8.1 for discussion).

But let us now also assume that “cruelty” is accurately defined by the *Oxford English Dictionary* as something like “enjoying causing another to suffer.” This allows us to conclude that premise (2) is in fact true: indeed, that it is true by definition. But we must still ask, as epistemologists, how Copperfield manages to know its truth. Surely, Copperfield’s knowledge of (2) is not “based on” observations in the same way, or to the same degree, as is his knowledge of (1). It isn’t as though Copperfield first learns what cruelty is by fixing on Creakle’s motives. Instead, he seems to bring this item of knowledge to the case at hand. So how does he know the fact in question? According to some philosophers, though Copperfield will not have non-inferential knowledge of particular evaluative facts like (3), he can have non-inferential knowledge of general evaluative facts like (2). Just as we know, by “reflecting on our concepts,” that bachelors are unmarried and vixens are foxes, Copperfield knows, by reflection alone, that it is cruel to relish the infliction of suffering. Theories of this kind are also discussed in section 4.2.

But there are other options here. According to a distinct set of epistemologists, Copperfield’s knowledge of Creakle’s immorality is inferential, but the premises of the inference needn’t include moral facts that Copperfield knows in an intuitive or non-inferential way. Instead, the boy knows that Creakle has acted immorally by inferring as much from a set of wholly non-moral or value-neutral premises that he knows. On this view, Copperfield infers an “ought” from an “is.” Indeed, some would argue that the requisite inference will prove deductive in form, that the value-neutral premises of Copperfield’s inference will entail its value-laden conclusion. These deductive models of moral knowledge are the topic of chapter 5.

Other philosophers would reject altogether the deductive inference that we presented above as (1)–(3). To their way of thinking, Copperfield needn’t deduce (3) from his knowledge of (1) and (2), as the boy’s knowledge of Creakle’s immorality is in fact abductive in nature. According to these accounts, Copperfield infers Creakle’s immorality as the best explanation of the man’s behavior, even though he knows that other explanations are possible. Abductive moral knowledge is discussed in chapter 6.

Epistemologists have also asked questions about (a) the causal origin of our moral beliefs, and (b) the differing roles that reflection and emotion play in generating moral knowledge like Copperfield’s. We have observed that Copperfield’s knowledge of premise (1) – that Creakle enjoyed striking the boys – is based on observation in a fairly direct way. And we remarked that Copperfield’s knowledge of premise (2) – the general fact that someone acts

cruelly when he relishes the infliction of suffering – differs in this regard, as Copperfield brings this item of knowledge to the scenes of corporal punishment Dickens describes. But is Copperfield’s knowledge of the nature of cruelty wholly independent from his experiences?

When writing the definition of “cruelty” the compilers of the OED were trying their best to report common usage of the term, usage that is guided by the ordinary English speaker’s beliefs about which people and actions are cruel and which are not. How, we might ask, does a speaker arrive at a fully general belief in the cruelty of taking enjoyment in the infliction of pain and harm? Is this knowledge similar to our knowledge of mathematics in its reflective or a priori nature? Or does it more closely resemble our knowledge of those scientific laws we establish through experimentation and observation?⁸

In the *Meno*, Socrates repeats a theory he had heard spoken among “priests and priestesses,” a story repeated by Pindar, “and many others divine among the poets,” a tale apparently embraced by Plato in the *Republic*.

As the soul is immortal, has been born often and has seen all things here and in the underworld, there is nothing which it has not learned; so it is not surprising that it can recollect the things it knew before, both about virtue and other things. (81c)

And, though the *doctrine of recollection* here described is no longer taken seriously, there is fairly substantial evidence that certain aspects of a normal human being’s capacity for empathy and sympathy are innately specified or genetically encoded. Indeed, there is at least some evidence that non-human primates possess a sense of fairness, however crude in nature (de Waal, 2006). We will discuss these matters in chapter 7.

Moral nativists go beyond the data we have at present to posit not only innate emotional capacities, and innate modes of inference, but a body of innate moral beliefs (Dwyer, 1999; Harman, 2000a; Hauser, 2006; Mikhail, 2008; cf. Kamm, 1993). The purportedly innate aspects of our ability to use and comprehend natural language provide the moral nativists with their model – a comparison that can be traced to Rawls (1971). Noam Chomsky (1957, 1986, 1988, 1995) has posited an innate universal grammar that enables children to learn a natural language – like English, Japanese, or Spanish – from the limited data presented to them during infancy. Similarly, nativists suggest, there is an innately known set of moral principles that enables children to acquire a moral competence from the limited instructions of their parents, teachers, and peers. Consider then Copperfield’s understanding of the cruelty

inherent in relishing the infliction of harm, or his grasp of the immorality people display when indulging in cruelty. Mightn't this understanding constitute one part of his innate knowledge of our shared "moral grammar?"

Critics of the nativist approach point to the range of moral disagreement and the differences among moral codes that we find at different times and places. Chomsky's hypotheses are supported by the common principles to which all natural languages conform. In contrast, anti-nativists argue, there are no universally accepted moral principles, or if there are some common moral rules, they are either empty or nearly empty in content (Sripada 2008a,b; Prinz 2008a,b).

There is at present very little empirical work that would move a neutral observer to favor either account of moral competence over its opponent. Indeed, it would seem that the controversy's resolution must await further articulation of the rules or principles that would constitute a universal "moral grammar" were one to exist (Harman, 2000a). But some of the issues relevant to the debate are broached in chapters 2 and 7 below.

Notice, though, that even if we go beyond the data and posit not only innate emotional capacities and innate modes of inference, but a body of innate moral beliefs, we won't have yet demonstrated or explained the existence of moral *knowledge*. Cognitive psychologists have uncovered a number of areas in which natural, nearly universal ways of thinking consistently lead us into error (Kahneman, Slovic, and Tversky, 1982; Kahneman and Tversky, 1996). So the innate specification of a cognitive structure is no guarantee of its truth or reliability. Nor is the nearly universal acceptance of an opinion sufficient to establish the justification with which we hold it, or the rationality of its retention. What, then, do we need to learn about Copperfield to figure out whether he truly *knows* what it is for an action to be cruel? And if, as we have been assuming, he does know premise (2), how did he arrive at this knowledge? Similar questions might be asked of a normal person's belief that it is unfair to break covenants on a whim or unjust to lie for monetary gain. Is it rational to persist in these beliefs? In virtue of what do they constitute knowledge?

Mathematics and science are commonly regarded as paradigm routes to knowledge. So attempts to compare various aspects of Copperfield's knowledge of immorality to our most fervently held mathematical and scientific beliefs are natural enough. But which (if either) is the more appropriate comparison? Is Copperfield's moral knowledge relevantly like our knowledge of math? Is it wholly based on abstract reflection? Or is it more like our knowledge of science? Did past encounters with sadism and hypocrisy play an essential role in its genesis?

Moral rationalists like Kant argue that we can know general facts like premise (2) in a wholly reflective manner: a way of thinking that resembles mathematical thought in its *a priori* or independence from experimentation, observation, and experience. Moral empiricists like Hume argue, in contrast, that moral knowledge depends for its existence on our emotional experiences, and is therefore *a posteriori* in nature. This historically important debate informs much of chapters 4–7.

We have asked whether Copperfield’s knowledge of the immorality or cruelty of Creakle’s actions is arrived at or sustained with an inference, or whether it is instead non-inferential in genesis; we have asked as to the nature of the inference that might support Copperfield’s belief; whether it might be deductive or non-deductive in nature; how Copperfield might know its premises; whether his knowledge of its premises might be innate or learned; and whether this knowledge might be grounded in pure reflection, or must instead depend for its existence on his emotional or affective experiences. But once we have some sense of how Copperfield has arrived at his moral views we will also want to inquire into the *reliability* of the process. We have supposed that Creakle did in fact act immorally when whipping the boys placed in his care, and this supposition entails that Copperfield’s opinion of his headmaster’s actions is accurate in relevant respects. But is the truth of Copperfield’s belief a matter of luck? Does the method by which he comes to believe in Creakle’s immorality always issue in true or accurate judgments? And if, as seems plausible, the method Copperfield employs is not infallible, when does it break down and why does it do so? How large and varied is the range of cases in which Copperfield arrives at the truth when reaching moral conclusions in the manner Dickens has here described? Do the successes greatly outnumber the failures? And are the successes sufficiently common to warrant our self-consciously retaining the form of moral reasoning Copperfield here employs? We will return to these questions in chapter 7.

1.5 Chapter summary

Moral epistemology consists in the study of whether and how we know right from wrong along with a host of related topics. The field is therefore extraordinarily difficult to circumscribe. Philosophers have yet to agree on definitions of key moral and epistemological terms such as “knowledge,” “justification,” “virtue,” and “immorality.” The Socratic definitions of “knowledge” that once held sway were refuted by Gettier’s counter-

examples; and alternative analyses have proven controversial. Nevertheless, there is nothing barring us from launching into an investigation of moral knowledge and related phenomena without first having definitions of “knowledge” and “morality” in place. We can rely on our common understanding of these terms.

Moral epistemology is best pursued using the method of levels. Zero-level inquiry consists in a description of our moral beliefs; first-level inquiry consists in a description of our differing epistemological evaluations of our moral beliefs; and second-level inquiry consists in an evaluation of the epistemic evaluations uncovered at the first level. Coherence is the most powerful tool used by philosophers pursuing second-level moral epistemology, but there may be other ways to rationally revise common thinking about our knowledge of morality and related matters.

Defensive epistemology attempts to respond to skepticism; constructive epistemology assumes the existence of knowledge and tries to explain its nature and genesis. Though we begin our study in earnest with defensive moral epistemology, it is helpful to first give some sense of the constructive issues that have engaged researchers in the field. Theorists disagree over whether our knowledge of particular moral facts is inferential or non-inferential in nature. Those who think our knowledge of particular moral facts is inferential disagree over the nature of these inferences and our knowledge of their premises. Some theorists argue that moral inference is deductive; some that it is abductive. Some argue that our knowledge of general moral principles is innate; some that it is acquired. Some argue that this knowledge is based on reflection or understanding alone and is therefore a priori in nature, whereas others argue that moral knowledge is a posteriori because grounded in emotional experience. Theorists also differ in their opinions of the reliability of those processes that give rise to our most basic moral beliefs. Some think that our common methods of moral reasoning are too unreliable and must be supplanted with novel, more trustworthy ways of thinking. Some disagree.

1.6 Further reading

There are a number of excellent anthologies of work in analytic epistemology; these include Ernest Sosa and Jaegwon Kim, *Epistemology: An Anthology* (2000) and Matthias Steup and Sosa’s *Contemporary Debates in Epistemology* (2005). Gettier’s short paper “Is Justified True Belief Knowledge?” (1963) is indispensable, and Robert Shope’s *The Analysis of Knowing* (1983) does a fine

job of describing the differing sources of dissatisfaction with post-Gettier analysis of “knowledge.” Timothy Williamson’s influential *Knowledge and Its Limits* (2000) argues for a conception of epistemology that would preclude reductive analyses of “knowledge,” and William Alston’s *Beyond “Justification”* (2005) urges a non-reductive treatment of other important epistemic concepts.

W. V. O. Quine’s “Epistemology Naturalized” (1969) pictures a reduction of epistemology to psychology, an approach criticized by Jaegwon Kim in his “What Is ‘Naturalized Epistemology’?” (1988). The more nuanced method of levels described above has its origins in the drive to reflective equilibrium introduced by Nelson Goodman in *Fact, Fiction, and Forecast* (1955) and applied to ethical inquiry by John Rawls in his monumental work *A Theory of Justice* (1971). Norman Daniels discusses variations on the method of reflective equilibrium in his *Justice and Justification* (1996), and Stephen Stich warns against its uncritical use in both “Reflective Equilibrium, Analytic Epistemology, and the Problem of Diversity” (1988) and *The Fragmentation of Reason* (1990).

Robert Audi’s “Moral Knowledge and Ethical Pluralism” (1999a) provides a nice overview of the issues that have divided constructive theorists of moral knowledge, and the collection of essays *Moral Knowledge?* (1996), edited by Walter Sinnott-Armstrong and Mark Timmons, contains a representative sampling of the field. Debates over the existence of innate moral principles take center stage in Sinnott-Armstrong (ed.), *Moral Psychology*, vol. I, *The Evolution of Morality: Adaptation and Innateness* (2008a) and Peter Carruthers, Stephen Laurence, and Stephen Stich (eds.), *The Innate Mind*, vol. III, *Foundations and the Future* (2007).