

Epistemology

Classic Problems and Contemporary Responses

SECOND EDITION

Laurence Bonjour

ROWMAN & LITTLEFIELD PUBLISHERS, INC.
Lanham • Boulder • New York • Toronto • Plymouth, UK



Contents

	Preface	vii
Chapter 1	Introduction	1
Part One	The Classical Problems of Epistemology	
Chapter 2	Descartes's Epistemology	9
Chapter 3	The Concept of Knowledge	23
Chapter 4	The Problem of Induction	47
Chapter 5	A Priori Justification and Knowledge	71
Chapter 6	Immediate Experience	97
Chapter 7	Knowledge of the External World	119
Chapter 8	Some Further Epistemological Issues: Other Minds, Testimony, and Memory	149
Part Two	Contemporary Responses to the Cartesian Program	
	Introduction to Part Two	175
Chapter 9	Foundationalism and Coherentism	177
Chapter 10	Internalism and Externalism	203

Chapter 11	Quine and Naturalized Epistemology	221
Chapter 12	Knowledge and Skepticism	237
	Conclusion	257
	Questions for Thought and Discussion	261
	Notes	273
	Glossary	307
	Annotated Bibliography	319
	Index	325
	About the Author	331

CHAPTER ONE



Introduction

The book you are reading is an introduction to the philosophical subject of *epistemology*. As a first stab, epistemology is the philosophical study of *knowledge*: its nature, its requirements, and its limitations. The best way to begin our inquiry into this area is to try to get some idea, in an initial and tentative way, of why and in what way knowledge seems to deserve or even require philosophical investigation and scrutiny—so much so, as it turns out, that epistemology has often been regarded as the most central area of philosophy in the period since the Renaissance. And to do that, it will be useful to say just a little about the general character of philosophy itself.

Philosophy has been described in many different ways, not all of them entirely consistent with each other. But perhaps the most helpful characterization at a general level is that philosophy is the search for *reflective understanding*: in the words of a prominent recent philosopher, the effort “to see how things in the broadest possible sense of the term hang together in the broadest possible sense of the term.”¹ As this might already suggest, the philosopher is particularly concerned with broad and general topics or areas: the nature and makeup of human beings, the basic ingredients and structure of reality, the nature and basis of value. Most of the general topics that the philosopher investigates can also be studied from other points of view, especially from the perspective of empirical science. But while the philosopher may make use of the results of these other investigations, his or her focus is different: more general, more abstract, and aimed in a distinctive way at intellectual *problems* that arise in the effort to understand, places where our

thinking seems to get tied into knots or tangles that are difficult to unravel, hard to make clear sense of. It is the presence of problems of this sort that makes a subject of particular concern to philosophers. And it turns out that that knowledge is a subject area in which the problems are especially difficult, pervasive, and troubling in their implications.

The most central and important of these problems will constitute the main subject matter of this book, and specific accounts of them will come later. But our goal for now is to get some initial idea of how and why such problems arise, of why knowledge, perhaps contrary to your first impressions, is puzzling or problematic in ways that make it difficult to achieve an intellectually satisfying understanding of it. One place to start is with a rough list of the various sorts of things that *seem* from a common-sense standpoint to be reasonably clear cases or instances of knowledge. To keep the project manageable, I will relativize the list to my own case, but such lists could obviously be similarly constructed for others or for whole groups of people. (Indeed, all of you who are reading this should try to construct a parallel list for yourselves.)

Here are some plausible general categories and specific examples of things that I know or at least confidently seem to myself to know:

1. Facts about my present subjective experiences or states of consciousness: that I feel an itch in my left thigh; that I am thinking about how to explain the problems pertaining to knowledge; that there is a large and variegated patch of green in the middle of my visual field.
2. Facts about my presently perceived physical environment, including my own body: that I have two hands; that there is a computer screen before me; that music is playing in the background; that there are large evergreen trees outside my window.
3. Facts about the larger perceptible and social world beyond my present experience: that my wife is presently teaching her class at the University of Washington; that there is a large lake a few blocks from my house; that there is a large range of mountains called “the Rocky Mountains” several hundred miles east of here; that there are several million people in New York City; that there are two main governments in the British Isles, one centered in London and the other in Dublin.
4. Facts about my personal past, the past that I actually experienced: that there was a black-capped chickadee on my bird-feeder this morning; that I took my dogs to an off-leash park last Sunday; that I used to live in Texas; that I have had various specific physical injuries at different times; that I taught various specific courses in the past (though here many details are fuzzy or altogether lacking).

5. Facts about the historical past that were not part of my personal experience, though they were experienced at least in part by others: that my wife grew up in Spokane; that George Bush was elected president in 2000 and again in 2004; that there was a worldwide depression in the 1930s; that the United States was first a British possession and then achieved independence under the leadership of George Washington; that the Roman Empire once controlled a very large area of the world.
6. Facts about the experiences and mental states of other people and at least some animals, in the past and sometimes in the present: that my wife was anxious this morning about her first class of the term; that my dogs are excited by the prospect of a walk; that many of the people at the concert last week enjoyed and were enthusiastic about the performance; that a certain student in one of my classes was very puzzled during a certain lecture last quarter; that one of my colleagues is often angry at the administration; that an injured protester (observed on television) was in severe pain.
7. Facts about the dispositional and character traits of myself and others (both people and some animals), again in both the past and present and extending into the future: that I am a rather cautious person; that some of my colleagues are not very responsible; that one of my dogs is easily frightened; that some people are afraid of water; that many students tend to be somewhat lazy about studying.
8. General and causal facts concerning observable objects and processes: that small amounts of sugar will always dissolve in large quantities of water; that green apples (of varieties that turn red or yellow when ripe) taste very sour; that indoor plants will eventually die if they don't receive water; that a thrown baseball will bounce off a cement wall; that gasoline ignites very easily.
9. Facts about future events: that the pane of glass I am holding will break (as it slips out of my hands above the paved driveway); that it will rain again in Seattle; that my research quarter will come to an end; that the 2012 presidential election will take place; that I will eventually die.
10. Facts that were or are outside the range of anyone's direct observation or that could not in principle be observed: that gases consist of tiny molecules; that the pinpoints of light in the sky are in reality large stars; that computers store information via magnetic coding; that evolution occurred; that the picture in my television set is produced by electrons striking the back of the screen; that it is very hot in the center of the sun.

11. Facts the knowledge of which does not seem to depend on sensory experience at all²: that $2 + 5 = 7$; that triangles have three sides; that anyone who is a bachelor must be unmarried; that when a certain container A is larger in volume than a certain other container B, and container B is in turn larger in volume than a third container C, then container A must be larger than container C; that if the surface of a ball is uniformly red at a certain time, then it is not also uniformly green at that time; that either today is Wednesday or today is not Wednesday.

And this list is obviously quite incomplete. Each of the lists of specific examples could be extended in various directions (try doing some of this for yourself). And there are also further general categories that many people would want to include, though almost all would agree that these are more questionable: especially those facts supposedly corresponding to moral and religious beliefs.

As we will see, there are problems and issues that can be raised about each of these apparent categories of knowledge. Perhaps the most obvious questions to ask right now are these: First, what does it *mean* to say that I *know* each of these various things? What conditions or criteria or standards must be satisfied for such a claim of knowledge to be true or correct? Second, supposing that I do in fact know these things, *how* do I know them? What is the *source* or *basis* of my knowledge? In some cases, the rough answer to this second question seems fairly obvious: I know about my immediate perceived environment via *sensory experience*, about my past history via *memory*, about the mental states of other people via observations of their bodily *behavior* (including especially their verbal behavior: what they say or seem to say). But further questions can be raised about how each of these alleged sources of knowledge works—and about whether it is genuinely reliable, whether it leads to true (or at least mostly true) results. And for many of the other general categories of apparent knowledge, even a rough answer to the question about its source or basis is much less obvious. How can we know facts about the future? How can we know facts about unobservable entities? How can we know facts like those in category 11, where sensory experience seems not to be involved at all? (Note also the important assumption being made throughout the list, one which is both natural and will turn out to be correct but must still eventually be discussed, that it is only things that are *true* that can be known, that are even candidates for knowledge.)

A further, though still closely related question arises from the reflection that there are also obviously even larger numbers of facts in each of the

indicated categories that I do not know. Some of these I could come to know with varying degrees of effort, but many of them would be difficult or impossible for me to know. So what then is the difference between the two sets of items, the known and the unknown? Again, rough answers suggest themselves for many of the categories, but elaborating these in detail often raises difficult questions.

One more important question that can be asked right now is whether I really do know all of the things that I think I do (or that common sense would say that I do)—or, much more radically, whether I really know any of them at all. What initially gives force to this question (along with uncertainties about how the various sorts of knowledge are obtained) is the familiar fact that sometimes I turn out not in fact to know something that I thought that I knew: that my dog is outside (the door was ajar and he slipped back in); that there are only three books on the table (there is another book hidden under one of the ones that I see); that there is a drugstore on a certain corner (it has burned down or closed); that there is a robin in the yard (it is really a varied thrush); that a certain student is following my lecture (she has merely learned when to nod or smile, but actually, as will be revealed when she tries to answer a question, has no real grasp of what I am saying); that a certain person is honest (he is really just a good liar); that vitamin C prevents colds (it really has no effect of this sort). As these examples reveal, it is easiest to find clear examples of apparent but nongenuine knowledge in categories 3, 7, 8, and 9, but there is no obvious reason to think that mistakes are confined to these categories, as opposed to just being harder to discern in the others. (Whether mistakes of this sort are possible in *all* of the categories, most particularly 1 and 11, is a more difficult issue, one that will be discussed later on.) Another point suggested by the examples is that the clearest instances of seeming knowledge that turns out not to be genuine are those in which the claim in question is discovered to be false (again reflecting the idea that only truths can be known). But it should not be assumed, and will in fact turn out not to be true, that this is the only way in which a claim of knowledge can be mistaken.

The concern raised by cases of apparent knowledge that turns out not to be genuine, of what we might call “failed knowledge,” may seem relatively minor, unthreatening, and easily dealt with. From a common-sense standpoint, such cases are relatively infrequent and seemingly easy, at least in principle, to identify. Thus it is unclear that they should be taken as symptoms of a serious problem. But there are two reasons why such a response seems too easy, not really intellectually satisfying. One is the point already noted that merely the fact that easily noticeable cases of failed knowledge are rare provides

no clear reason for thinking that less easily discernible ones are not much more common, perhaps even quite pervasive. If our efforts at knowledge can sometimes seem to be successful when they actually are not, why could this not occur much more commonly than we think without our being able to tell—to know—that it does? Real confidence on this point seems to demand at the very least a much clearer understanding of how knowledge works, of what determines whether apparent instances of knowledge are genuine. And the second point is that such an understanding would be intellectually valuable in any case, even if the common-sense reaction to the problem of failed knowledge is basically correct.

It is this concern that apparent knowledge might not be genuine which motivates the French philosopher René Descartes, often described as both the father of modern philosophy and the father of epistemology, at the beginning of his famous *Meditations on First Philosophy* (1641):

Several years have now passed since I first realized how numerous were the false opinions that in my youth I had taken to be true, and thus how doubtful were all those that I had subsequently built upon them. And thus I realized that once in my life I had to raze everything to the ground and *begin again from the original foundations*, if I wanted to establish anything firm and lasting in the sciences. [13]³

The problem in question was certainly much more obvious in Descartes's time, when modern science was in its infancy and the cross-currents of conflicting opinion and doctrine were much harder to sort out and evaluate. But the problem for us is essentially the same, and it is at least not obvious that there is any easy and unproblematic solution to it to be found. The central risk is that in trying to decide whether we really know one thing we will inadvertently appeal to other things that we think we know, but about which we are in fact mistaken. And this is probably the basic reason for the very radical character of Descartes's proposed solution, suggested in the second of the quoted sentences, one that we will examine in the next chapter.

One last question of a preliminary sort: How much does it *matter* whether we know what we think we know? Why do we *care* about knowledge—in particular, what is it about knowledge that really matters for our lives? My eventual suggestion will be that it is in fact not so much knowledge itself but rather certain of its key ingredients that are our main concern. But this is getting ahead of ourselves and must await later discussion.⁴

We turn then, in the next chapter, to a discussion of Descartes's historically seminal epistemological program and of the basic principles that underlie it.

CHAPTER TWO



Descartes's Epistemology

As already noticed briefly in the first chapter, the work that is arguably the starting point of modern epistemology is Descartes's *Meditations on First Philosophy* (first published in 1641). It is likely that many readers of the present book are already familiar with the *Meditations* and the engaging though perhaps also somewhat overly picturesque scenario that Descartes offers there. The main aim of the present chapter is not to offer yet another discussion and evaluation of that scenario and of the specific arguments and conclusions that Descartes offers in connection with it. Though we will have to pay some attention to the specific details of the *Meditations*, my main concern in this chapter is to discern and extract the underlying epistemological principles or assumptions that Descartes is relying on and, to some extent, defending there—which I will refer to as the principles of Cartesian Epistemology.¹ It will turn out that these Cartesian principles provide a surprisingly good guide to the central issues that have been the focus of epistemological discussion from Descartes's time all the way to our own.

The Method of Doubt

We have already taken note, in the previous chapter, of Descartes's starting point. He has come to realize that very many of the things he has previously believed are false, and the question is what he should do about this. This is a question worth thinking about with some care. What would *you* do if you realized that many of your beliefs were mistaken, but had no very firm idea

of which ones or how many? One obvious alternative would be to continue to examine and scrutinize your various beliefs and opinions individually, looking for mistakes and trying to correct them. But the problem with this, also briefly noticed earlier, is that such an examination of a particular belief would inevitably rely in large part on your other beliefs and convictions, particularly on the underlying principles that you accept, explicitly or implicitly, concerning how to identify beliefs that are false and how to arrive at beliefs that are true.² And if some or all of these other beliefs and principles should turn out themselves to be mistaken, then the whole project of identifying and eliminating mistaken beliefs would very likely be doomed to failure, since you would be as likely to retain old errors and even introduce new ones as to weed out the existing ones.

At least in part for this reason, Descartes proposes something much more radical: to tentatively reject any view or opinion or principle that is not “completely certain and indubitable,” any for which he can find “some reason for doubt,” some way in which the claim in question might be false in spite of whatever apparent reasons or basis have led him to accept it so far [13]. Here it is important to understand that the way in which a particular belief might be false does not have to be probable or even very plausible—it is enough that it is merely *possible*, something that cannot be *conclusively* ruled out. Anything for which such a basis for doubt can be found is something that *might* conceivably be false and so is something that cannot be accepted or relied on if the goal is to conclusively eliminate all error.³ (It might of course be questioned whether the complete elimination of error is a reasonable goal, one that we have any realistic chance of achieving.)

There are several stages to the resulting progression of doubt, as Descartes considers different kinds of beliefs and the ways in which they might be mistaken, but it will be enough for our purposes here to focus on the final and decisive one: the famous “evil genius” hypothesis.

I will suppose . . . an evil genius, supremely powerful and clever, who has directed his entire effort at deceiving me. I will regard the heavens, the air, the earth, colors, shapes, sounds, and all external things as nothing but the bedeviling hoaxes of my dreams, with which he lays snares for my credulity. I will regard myself as not having hands, or eyes, or flesh, or blood, or any senses, but as nevertheless falsely believing all these things. [16–17]

According to Descartes, such an evil genius (in effect a being with God’s alleged omnipotence, but differing from more standard versions of God in being bent on deception) would be capable not only of deceiving me about

the material world (including the contents of such sciences as physics and astronomy) and about my own physical nature, but also even about such areas as arithmetic and geometry: "May I not, in like fashion, be deceived every time I add two or three or count the sides of a square?" [15]

To repeat, Descartes is not saying that it is probable or even at all plausible that such a being exists; indeed, he would probably concede that the existence of such an evil genius is extremely unlikely (though this too could obviously be doubted!). But that its existence cannot be conclusively ruled out is enough to provide a *possible* basis for doubt. And thus by the end of Meditation 1, it begins to look as though Descartes has found a reason to doubt *every belief he has*, whether about the material world or about such abstract subjects as arithmetic and geometry. The reason is simply the mere possibility that such an evil genius exists.

The *Cogito*

Does anything at all survive this process of systematic doubt? Descartes initially takes seriously the possibility that it may be "not within [his] power to know anything true," or perhaps rather that he can only "know for certain that nothing is certain" [17]. But this turns out in the end not to be so. For as he famously argues, there is at least one thing that cannot be doubted on this basis, something about which even the evil genius cannot deceive him, namely, his own existence:

But doubtless I did exist, if I persuaded myself of something. But there is some deceiver or other who is supremely powerful and supremely sly and who is always deliberately deceiving me. Then too there is no doubt that I exist, if he is deceiving me. And let him do his best at deception, he will never bring it about that I am nothing so long as I shall think that I am something. Thus . . . "I am, I exist" is necessarily true every time I utter it or conceive it in my mind. [18]

Though he does not use exactly this wording here, the gist of this argument is captured in the famous Latin formula "*Cogito ergo sum*," "I think, therefore I am," and it has come to be referred to simply as the *Cogito*.

While there are many questions that can and have been raised about the *Cogito*,⁴ Descartes's basic claim that his belief in his own existence cannot be doubted, that this is something that he cannot be mistaken in believing or accepting, seems plainly correct. (Doesn't it? Think about this for yourself before proceeding.) What is not clear, however, at least initially, is that this

result can contribute very much to Descartes's overall project of eliminating error from his beliefs and thus perhaps arriving at a substantial body of knowledge that is certified to be error-free. (The mere elimination of error could of course be achieved, in principle at least, by simply believing nothing at all.) Descartes is careful to make clear that that result yielded by the *Cogito* is not that the flesh-and-blood, biologically constituted, historically located person René Descartes exists, for the evil genius could still obviously deceive him about the physical and biological and historical aspects of his nature. The secure and indubitable conclusion, he says, is only that he exists as "a thinking thing; that is, a mind, or intellect, or understanding, or reason" [19]. And this, even if correct, seems to amount to very little. If the evil genius could still deceive him about everything else, then the Method of Doubt seems to have left Descartes in a situation of extreme, albeit not quite complete *skepticism*: a situation in which his knowledge is confined to this single, crucially important but still extremely limited fact.

Descartes does not, however, view the result of the *Cogito* as being limited to this extent:

But what then am I? A thing that thinks. What is that? A thing that doubts, understands, affirms, denies, wills, refuses, and that also imagines and senses.

. . . Is it not the very same "I" who now doubts almost everything, who nevertheless understands something, who affirms that this one thing is true, who denies other things, who desires to know more, who wishes not to be deceived, who imagines many things even against my will, who also notices many things that appear to come from the senses? What is there in all of this that is not every bit as true as the fact that I exist—even if I am always asleep or even if my creator makes every effort to mislead me? . . . For example, I now see a light, I hear a noise, I feel heat. These things are false since I am asleep. Yet I certainly do seem to see, hear, and feel warmth. This cannot be false. [20]

Descartes's claim in this passage, a claim that is absolutely crucial for his subsequent argument, is that the immunity from even possible doubt, the *indubitability* that is a feature of the claim about his own existence, is also in the same way a feature of his awareness of his specific conscious states of mind, his specific thoughts and desires and sensory experiences—that the evil genius could no more deceive him about the contents of those states of mind than about his own existence. And this in turn gives him an essential further starting point, over and above the bare fact of his own existence, for the project of reconstructing his knowledge.

But is Descartes right that the evil genius could not deceive him about the contents of his own mental states? The issue is difficult, and Descartes's

claim here is certainly far less obvious than the analogous claim about his own existence. Consider as an example the awareness of a particular sensory content, such as my visual experience of a large green coniferous tree directly in front of me. Now the evil genius could surely deceive me about whether there is really a tree there, that is, could cause me to believe that there is a tree when there is not. It⁵ could also seemingly, though somewhat less obviously, deceive me about the significance of the sensory experience I am having, for example, could cause me to believe mistakenly that my experience is of the sort that depicts or is usually caused by or associated with a large green coniferous tree.⁶ But could it deceive me even about the existence or character of the specific sensory experience itself? (Think about this for yourself before proceeding, but don't leap too quickly to a conclusion.)

Well, why couldn't the evil genius deceive me about this? Couldn't it, being omnipotent, produce in me the *belief* that I was having such-and-such a specific sort of experience when actually I was not? Suppose that it can: it makes me believe that I am having a visual experience of a green square (to take a somewhat simpler case) when it is in fact false that I am having such an experience.⁷ And wouldn't this amount to deceiving me about even the existence of the experience? But think carefully here: according to this supposition, I believe that I am having an experience that I am not in fact having. Do you think that this is really possible? Could I really believe that I am having an experience of a green square (or of pain or of the taste of fudge), when I am not really having such an experience? Wouldn't I at once notice the discrepancy between the belief and my actual experience and so cease to accept the belief?⁸

Though he never explicitly considers this issue, I think that Descartes would have responded to it in the way just suggested: As long as the evil genius produced only such a belief in me without also producing the actual conscious sensory experience itself, I would be deceived by such a belief, if at all, only for the briefest instant. The falsity of the momentary belief would be immediately apparent to me by comparing it with whatever conscious experience I actually was having.⁹ (And, of course, if the evil genius also produced in me the relevant sort of conscious sensory experience, then I would no longer be being deceived about its existence.) There will be more to be said later about the issues in the vicinity of this question,¹⁰ but for now I propose to grant Descartes this further claim, at least provisionally, and proceed to examine the use he makes of it.

So by using his Method of Doubt, Descartes has tentatively rejected the vast majority of his beliefs, but not quite all of them. Two important kinds of beliefs have, he claims, survived the application of the Method

of Doubt: (1) the belief that he exists as a thinking thing; and (2) the many specific beliefs that he has about the contents of his various specific experiences or states of mind.

The Existence of God

Descartes now has what he describes as his “first instance of knowledge” [24]: he knows that he exists and that he has states of mind of various specific sorts. But how is he to go beyond this still pretty meager beginning? The only very obvious way to get from such a purely subjective starting point to further conclusions of any sort about the world *outside* his mind is to find some sort of rationally cogent *inference* from the former to the latter, from the premise that he has such-and-such specific states of mind to the conclusion that something of such-and-such a specific sort exists in the mind-external world. If there are no rationally compelling inferences of this sort to be found, then it seems that Descartes’s knowledge will be confined forever to his own mind and its contents. This would still be a severe sort of skepticism, even though slightly less severe than the one that would limit his knowledge to the mere fact of his existence.

Is there any cogent inference of this general sort to be found, any rationally legitimate way of inferring from the contents of our subjective mental states to facts about the external, at least primarily material world? Perhaps the most prevalent view from Descartes’s time to our own, and especially in recent discussions, has been that there is not, that Descartes and the others who have followed his lead have backed themselves into a corner from which there is no escape.¹¹ Whether this is indeed so is quite possibly the most difficult of all epistemological issues. We will return later to the question of whether this pessimistic assessment is correct¹² and still later to the issue of whether there are viable epistemological approaches that can somehow avoid the issue entirely.¹³ For the moment, our task is to examine the general structure of Descartes’s own approach.

Considered at a very abstract level, Descartes’s strategy is to argue that (1) the fact that the content of his mental states has a certain specific feature (or features) can only be *explained* by supposing that (2) that feature is caused by and correctly represents something existing outside of his mind. Since he already knows that his mental states have the specific content that they do, he can then infer that the external cause in question must exist. What is needed then are defensible specific instances of this general pattern of argument.

Descartes considers briefly [26–27] the possibility that he might be able to infer from (1a) the fact that he has perceptual ideas of various kinds of

material objects to (2a) the existence of actual external objects that those ideas resemble¹⁴ and that produce them. But to be justified in claiming that facts of the sort indicated by (1a) can only be explained by facts of the sort indicated by (2a), he would need some background premise or principle to this effect, one that he already somehow *knows* to be true if he is to thereby know the resulting conclusion. According to Descartes, however, his reason for thinking (prior to the Doubt) that facts of the sort indicated by (1a) must be explained by facts of the sort indicated by (2a) is only that he “has been so taught by nature,” that is, that he is “driven by a spontaneous impulse to believe this,” an impulse that he eventually characterizes dismissively as “blind.” And this, he argues, is plainly not good enough. Such spontaneous impulses have often led him astray; and (a deeper point that is only suggested but not really stated explicitly) they involve no insight into how or why the claim in question must be true.

The specific argument that Descartes eventually endorses [28–34] is instead that (1b) he has an idea of God, understood as “a certain substance that is infinite, independent, supremely intelligent and supremely powerful, and that created [him] along with everything else that exists,” and that the existence of this particular idea can only be explained by supposing that (2b) it is ultimately caused by a being actually having those characteristics, that is, by God himself (or herself), who therefore must exist. (And, as we will eventually see, it is by appeal to the supposed fact of God’s existence that Descartes attempts to reconstruct his other knowledge of external reality in a way that is allegedly free from error.)

This argument also obviously requires some sort of background premise or principle that establishes that (1b) can only be explained by supposing that (2b) is true. Unfortunately, the principle that Descartes actually suggests is extremely implausible, indeed difficult to really make very clear sense of. It is the principle that “there must be at least as much reality in the efficient and total cause as there is in the effect of that same cause” [28]. Here “reality” means something like *perfection*; and the sorts of reality to which the principle applies are supposed to include both “formal reality,” that is, the reality (or perfection) that a thing has by virtue of its actual existence and qualities, and “objective reality,” that is, the reality (or perfection) that an idea supposedly has by virtue of the formal reality that what it represents would have, if it existed.¹⁵ The suggestion is then that these two seemingly very different sorts of “reality” (or perfection) are nonetheless on a par from the standpoint of causation, that is, that what causes an idea of something must have as much reality (formal or objective) as the object represented by that idea would have; and hence that the objective reality (or perfection)

of Descartes's idea of God can ultimately be explained (since an infinite sequence of ideas is impossible) only by the existence of something having the same degree of *formal* reality (or perfection) as the idea has *objective* reality (or perfection), that is, by the actual existence of God (since only God has that degree of perfection).

The "Light of Nature"

There is no way, in my judgment, to make either the foregoing argument or the causal principle that underlies it at all plausible, the most obvious problem being that merely objective reality seems obviously easier, "cheaper" to produce than formal reality, thus allowing the idea of God to be produced by something much less exalted in its characteristics than God himself (or herself), for example, by human imagination. Our immediate concern, however, is to understand the epistemological status that this causal principle is supposed by Descartes to have—even though he is surely wrong that it actually has it. Descartes's claim is that the causal principle has a status that is different from and epistemologically superior to that of the principle discussed earlier concerning ideas of material objects and the objects that they supposedly resemble. Whereas his belief in the latter principle results merely from a spontaneous but "blind" impulse, the causal principle is revealed to him by what he refers to as the "light of nature," whose results "cannot in any way be doubtful" [26].

But what exactly is this "light of nature," and why are the beliefs or convictions that it produces supposed to have this status? Descartes refers to it as a cognitive "faculty" and says that "there can be no other faculty that [he] can trust as much as this light and which could teach that [the things revealed by the light of reason] are not true" [26–27]. Somewhat more helpfully, he describes the results produced by this faculty as "evident" and as "manifestly true" [28]. The underlying idea seems to be that the causal principle and other beliefs and convictions (if there are any) that result from the "light of nature" are *self-evidently* true, that is, are things that can be seen to be true simply by thinking about their content. It is this self-evidence that Descartes somewhat picturesquely describes as being revealed by the "light of nature." And in virtue of being self-evident, beliefs or convictions having the status that the causal principle is alleged to have can seemingly be known independently of any reliance on sensory and introspective experience: known a priori, as later philosophers would put it (though Descartes does not use this phrase).

To repeat, it is more than doubtful that Descartes's causal principle actually has this status. But even if this particular candidate for the status of self-

evidence is unsuccessful, it seems pretty obvious on reflection that Descartes needs *something* having this general sort of status if he is going to infer successfully from the contingent fact that he has such-and-such specific mental states (especially states of sensory experience) to the existence of specific kinds of external, especially material reality. Such an inference will, as we have already seen, require a *known* connecting principle of some sort, a principle saying that if someone has mental states with those specific contents, then it follows (somehow) that a certain sort of external reality must exist as well. But how is any such principle itself to be known, since what it says is not a fact merely about mental states? To say that it too is inferred from the fact that Descartes has mental states with various specific contents would mean that the knowledge of this principle would have to depend on *another* known connecting principle, one saying this time that if certain specific mental contents occur then the first connecting principle must be true. And then how is this second principle to be known? To say that it is also known in this same way would then require a *third* known connecting principle, and so on, leading to an infinite and apparently vicious *regress* of such principles, each dependent on the next, none of which would ultimately be known, since the series could never be completed. And the only apparent way to avoid this regress is to say that *some* principle (and it may as well be the first one in the sequence) can be known without reliance on this sort of inference, that is, can be known independently of the fact that Descartes has certain specific mental, especially sensory states.¹⁶ And the only way that this can apparently be so is if the principle is self-evident in the way just described.¹⁷

If Descartes is right that there are beliefs or principles having this status, then he has seemingly identified a *second* sort of possible knowledge that he can use as a starting point for further reconstruction, even if the specific instance he appeals to is highly dubious: if there are claims or principles that are genuinely self-evident, then they can be used to supplement his knowledge of the contents of his own mental states (and of his own existence), thus possibly providing a basis for inference to further knowledge, including knowledge of the material world. This idea of self-evidence also raises a number of problems and issues that we will consider later on in this book.¹⁸ But there is one specific difficulty, growing out of Descartes's own position, that needs to be discussed now, in concluding the present section.

In developing the idea of self-evidence that seems to underlie Descartes's appeal to the "light of nature," we have temporarily lost sight of the specific problem that motivates his whole discussion, namely the concern that his various beliefs and convictions might result from the actions of the envisaged evil genius who uses all of its power to deceive. This is indeed pretty much the way that Descartes himself proceeds, but we must now ask the

obvious question: Couldn't the evil genius deceive Descartes about the causal principle itself, making it *seem* to be self-evident, *seem* to be revealed by the "light of nature," even though it is actually false (and analogously for any other allegedly self-evident principle to which Descartes might appeal)? Indeed, at a point prior to the specific discussion of the causal principle and of the resulting argument for God's existence, Descartes himself lends force to this question by mentioning the truth of the proposition that two plus three equals five (or rather, equivalently, the falsehood of the proposition that two plus three does not equal five), surely an obvious example of a claim that can be plausibly regarded as self-evident, as something of which he cannot be certain until the worry about the evil genius has somehow been laid to rest [25].¹⁹

Moreover, reflection on this point suggests an even deeper problem for Descartes's position: not only is it far from clear that self-evident claims can escape the doubt that results from the evil genius hypothesis; but, even worse, Descartes's attempt to meet this doubt turns out to be a circular, question-begging argument (involving the so-called Cartesian circle). Descartes proposes to alleviate the doubt by proving the existence of a perfectly good God, who is therefore not a deceiver, and whose existence thus rules out the existence of an all-powerful evil genius. The proof relies, as we have seen, on the causal principle, which in turn depends on the underlying principle that self-evident claims revealed by the "light of nature" are true. But this last principle is not secure from doubt, according to Descartes himself, as long as the existence of the evil genius has not been ruled out. The resulting circular argument thus moves from the general principle that self-evident claims are true to the specific causal principle to the existence of a nondeceiving God to the nonexistence of the evil genius to the conclusion that self-evident claims are true and can be trusted. It thus establishes the nonexistence of the evil genius only by relying on a general principle that cannot be known to be trustworthy until that nonexistence has already been established, thus rendering the argument circular and so futile.

Though this objection to Descartes's actual argument is quite clear and pretty obviously fatal, it is not obvious what further conclusions we should draw from it. While it might seem at first to suggest that the practice of accepting claims and principles on the basis of their supposed self-evidence does not yield knowledge after all and accordingly should be rejected, it is unclear, as Descartes himself suggests in the passage quoted earlier, what the alternative to self-evidence might be, at least with regard to beliefs or principles warranting inferences that go beyond the contents of one's own mental states. Such beliefs or principles cannot be justified by appeal to the

mental states that they attempt to go beyond, and self-evidence seems to be the only other possibility. Again the threat of skepticism looms.

An alternative suggestion is that what the objection shows is that Descartes's implicit standard for knowledge is too demanding: that knowledge does not after all require overcoming all *possible* doubt. We will return to this issue later.²⁰ For now, I turn to an examination of how Descartes, having (as he supposes) established the existence of a nondeceiving God (and thus having eliminated the possibility of the evil genius), tries on that basis to reconstruct the rest of his knowledge, in particular knowledge of the material world.

Knowledge of the Material World

The account of knowledge of the material world that Descartes offers in the last of the six Meditations is in fact disappointingly thin. We have already taken brief note of the central theme: that the God whose existence has allegedly been established, being perfectly good, cannot be a deceiver. And since God "has given [him] a great inclination to believe" that his sensory ideas "issue from corporeal things," Descartes says that he cannot see "how God could be understood not to be a deceiver, if these ideas were to issue from a source other than corporeal things," which accordingly must exist [52].

Even this most minimal conclusion about the material world could hardly be more shaky, relying as it does on proofs of the existence of God that few other than Descartes would accept and on further claims about what such a being would or would not do that are questionable at best. But even if these problems are provisionally set aside, a further, more immediate question is *how much* he can know in this way about such "corporeal things" (beyond the bare alleged fact of their existence). Descartes can hardly claim that God's not being a deceiver means that all of the specific beliefs about the material world that he or we arrive at via sensory experience are guaranteed to be correct, since it is too obvious, to him as to us, that these beliefs are sometimes internally contradictory. But then which, if any, of his or our more specific beliefs about the material world, beliefs about the existence of specific sorts of objects in particular places at particular times, can be salvaged from the doubt on this basis?

Descartes makes a number of remarks that bear on this question, but none that yield a very clear and definite answer. The main ones are the following: First, since our "sensory grasp" of material objects "is in many cases very obscure and confused," we have no reason to think that "all bodies exist exactly as I perceive them by sense" [52]. Second, we can know

that material bodies have all of the features that are clearly and distinctly understood, “that is, everything, considered in a general sense, that is encompassed in the object of pure mathematics” [52]. This seems to mean only that we can know that material objects have the general kinds of qualities subsequently labeled “primary qualities” (by John Locke²¹ and others): such qualities as size, shape, and motion; but not necessarily that we can know that specific instances of these qualities are present in a particular case (“for example, that the sun is of such and such a size or shape”). Third, for other kinds of perceived qualities, the ones that do not lend themselves to mathematical measurement (“colors, sounds, odors, tastes, levels of heat, . . . grades of roughness, and the like”), we can conclude only “that in the bodies from which these different perceptions proceed there are differences [of some sort] corresponding to the different perceptions—though perhaps the latter do not resemble [that is, are not accurately represented by] the former” [53]. Fourth, our perceptions are still adequate to the primary purpose of “signifying to the mind what things are useful or harmful” to the person, even though they tell us nothing about “the essence of bodies located outside us” (that is, nothing very specific about the true natures of material bodies) “except quite obscurely and confusedly” [55]. Descartes sums all of this up by saying that we should not doubt that there is “some truth” in our perceptions, adding that the fact that God is not a deceiver means that where there is falsity in our opinions, he must have also given us a faculty that allows us, at least in principle, to correct our mistakes [53].

Thus, despite the rather upbeat tone on which the *Meditations* ends, Descartes’s attempted reconstruction of our knowledge actually salvages little that is very specific from the doubt induced by the evil genius hypothesis, and thus leaves us in a state of severe, albeit not total skepticism with regard to knowledge of the material world. We will have to consider in later chapters²² whether it is possible to do any better in this regard.

It is worth noting in passing, however, that there are at least traces in Descartes’s discussion of an argument that might prove more successful and that at least avoids Descartes’s extremely dubious reliance on theology. At certain points in his discussion, Descartes notices two important facts about our sensory experiences of the material world: First, our sensory experience is involuntary, independent of our will [26, 49]. Second, our various sensory experiences are, in general, related to each other in such a way as to fit together into a cohesive whole, thus differing significantly from the fragmentary experiences characteristic of dreams [58–59]. Taken together, these two facts seem to demand *some* sort of explanation (think about why this is so), with the claim that the experiences in question are systematically caused by

and are reflections of an independently existing world representing at least one obvious explanatory possibility.²³ We will consider later whether or not this general sort of argument has any real hope of success.²⁴

The Principles of Cartesian Epistemology

On the basis of the foregoing discussion, we are now in a position to formulate the central principles or basic assumptions of Cartesian epistemology, principles that have largely shaped the subsequent 300-plus years of epistemological discussion (though often, especially in recent years, only by providing a target for criticism):

1. *The concept of knowledge.* The view that has been standardly ascribed to Descartes is that only beliefs that are *infallible*, beliefs that are *guaranteed* to be true, can really count as knowledge. Descartes never actually states such a view, but the way that he employs the idea of knowledge in relation to the doubt suggests it pretty clearly. (Notice that the propositions that are the objects of such beliefs need not be *necessary* truths, that is, need not be true in every possible world: my own existence is, alas, merely contingent, but I still cannot, as the *Cogito* argument shows, be mistaken about it.) Presumably the person who has knowledge must actually have the infallible belief in question; and Descartes seems to suggest that this belief must also be very strongly held: the person in question must have *no doubt at all* that the proposition in question is true. And, finally, it seems clear that for Descartes the person must *realize* that the belief is infallible, must see or grasp the *reason* why its truth is guaranteed (since a belief that is in fact infallible but not recognized as such could still be doubted). Thus we have the following three-part Cartesian account of knowledge: knowledge is a *belief* held with no doubt for which the person has a *reason*²⁵ that guarantees *truth*.
2. *The rational or a priori basis of knowledge.* One initial basis for knowledge is provided by claims that are revealed by the “natural light,” that is, that are self-evident. Claims that have this status are knowable a priori, without reliance on sensory or introspective experience. Things known in this way thus provide one starting point or *foundation* for knowledge, on the basis of which other kinds of knowledge, including most importantly knowledge of the material world, can perhaps be inferred.
3. *The empirical basis of knowledge.* According to Descartes, the specific contents of one’s own conscious states of mind, including beliefs,

desires, sensory states, and many others, are known with the same certainty as one's own existence. This knowledge resulting from *immediate experience* thus provides a second starting point or *foundation* for further knowledge.

4. *The inference to the external, material world.* Everything else that we know, especially knowledge of the material world, is known via inference from these two foundational elements. The general form of such an inference that Descartes's discussion suggests is an *explanatory* inference, in which the reason for accepting various claims about the material world is that they provide the best explanation for facts about the contents of our mental states, especially our sensory states, with this inference being governed by self-evident principles. Descartes's own version of this inference uses the existence of God as an intermediate step and is extremely dubious. But there is at least a hint of a different version, appealing to the involuntary and cohesive character of our sensory experience, that *might* prove more successful.

Whether it is possible to build a tenable epistemology around these principles, despite Descartes's own rather obvious failure to do so, is an issue that will occupy us in various ways for most of the present book. The conclusion most widely accepted by recent philosophers is that the answer to this question is "no," that an acceptable epistemology, if possible at all, will have to depart very substantially from these Cartesian principles. My own belief is that this conclusion has been too hastily drawn and in fact that the principles of Cartesian epistemology are, when appropriately generalized and supplemented and with only minor corrections, still quite defensible as the basis of a satisfactory epistemological account—though it will take most of the rest of the book to make even a preliminary case for this conclusion. Subsequent chapters in Part I consider the issues raised by each of these principles in the order listed, though with a digression after the first to deal with an important issue that Descartes does not consider and a digression at the end to deal with some further questions. In Part II, we then consider the most important of the contemporary criticisms of, and alternatives to, the Cartesian epistemological program.