Introduction

This volume collects a dozen essays of mine written over the past thirty years, two of which are new to this volume. The essays cluster around a common theme, which I explain here.

Most of our beliefs are unreflectively arrived at. We go about our daily affairs, rarely stopping to reflect upon the epistemic status of our beliefs. When I am awakened by my alarm clock, I get dressed and ready for work with no thought at all about the epistemic status of my beliefs. Getting dressed and eating breakfast inevitably involves taking in an extraordinary amount of information about my environment: the location of the hot and cold water taps, where my clothing is, the height of each step on the stairway, what is in the refrigerator, and so on. All of this information is registered in me passively without a thought given to whether the beliefs produced within me are justified. My focus is not on epistemological matters; it is on getting dressed, eating breakfast, and getting to work.

And so it is with others as well, of course. Reflecting on our beliefs and their epistemic status is the exception rather than the rule. Non-human animals and very young children never engage in such reflection. The supremely overconfident do it far less than they should. And those crippled by self-doubt may engage in it to a pathological degree. But normal human adults do, at times, stop to reflect about whether they are believing as they should. Even if most of our beliefs, at most times, are not given a second thought, it does seem that the capacity for such reflective thinking, and the actual practice of engaging in it, is an important part of human intellectual life. The essays in this volume address just such second thoughts.

We give matters a second thought, in many cases, quite involuntarily. My wife was recently enjoying a pleasant stroll in Central Park, barely noticing the passersby, when she did a double take. Was that Leonardo DiCaprio who just walked by? The second thought was triggered in her, not something she chose to do as part of a carefully managed inquiry. But freely chosen or not, this second thought involved a reflective examination

1

of the epistemic credentials of her tentative identification. At other times, our second thoughts are very much a matter of choice. We purposefully stop to check upon beliefs already formed or hypotheses under consideration. We engage in epistemological investigation with the explicit goal of making sure that our beliefs are appropriately responsive to the facts.

Second thoughts about the epistemic standing of one's beliefs are the starting point of much epistemology. My focus here, however, is somewhat non-standard. For many epistemologists, much as they may agonize over the status of our beliefs about the physical world, about the future, about the past, and about other minds, they regard certain classes of our beliefs as especially resistant to the sort of doubt and second thoughts mentioned here. In the Cartesian tradition, beliefs about one's current mental states enjoy a special resistance to doubt. And for many epistemologists, beliefs which are a priori justified, including beliefs about the nature of good reason itself, enjoy a similarly special status. Such beliefs may, of course, be reflectively examined; one may give any belief a second thought. But if certain Cartesian views about these matters are correct, these beliefs will always stand up to such reflective scrutiny. Unlike other beliefs, reflective examination of this special class – certain beliefs about the self and about good reason – have an epistemic standing which assures that they need no correction, and so need no further reflective scrutiny. My focus in this volume is on just such beliefs. Where others would exempt such beliefs from doubt, or severely limit the scope of such doubt, I argue that such beliefs should be a matter of special concern.

My doubts do not come from thinking about the possibility of evil demons plotting against us, seeking to deceive us at every turn, nor from considering the possibility that we might be brains in vats, subject to direct electrical stimulation of our gray matter supervised by the twenty-first-century descendants of those evil demons. Such wild hypothetical cases raise concerns about global skepticism, and, for better or for worse, global skepticism is not my concern here. My worries stem from the more homely results of work in social psychology. These doubts, while far less global, are also, in my view, far more pressing. And they target the very kinds of beliefs which more traditional epistemologists have thought most immune to the kind of correction which second thought might provide.

As I see it, there is a great deal of work in epistemology which is prompted by the phenomenon of reflective scrutiny, regarding it as emblematic of the epistemological enterprise. The phenomenon of deliberation about what to believe, or whether to believe, is seen as offering an entree to understanding epistemological matters generally: the nature of

inference, rationality, and justification. More than this, deliberation, and thus the epistemological matters which it is thought to illuminate, is then investigated from a first-person point of view. We understand deliberation, it is thought, when we examine it from the point of view of the deliberator.

My own view is that this is a mistake, and that we can only understand epistemological matters, and even deliberation itself, when we engage in the kind of third-person examination of such phenomena which is found in experimental psychology. The first-person point of view deeply distorts our understanding of ourselves and of deliberation, but we can only come to see that this is so if we engage with the psychological literature on these matters. Far too much epistemological theorizing, especially about deliberation and reflective thought generally, never departs from the first-person point of view.

The result of these essays, taken collectively, is a kind of reversal of the Cartesian picture: where others see beliefs about oneself and about the a priori as especially resistant to doubt, and beliefs about the physical world as especially precarious, my own view is just the mirror image of this picture: I see our beliefs about much of the physical world as especially secure, and beliefs about the self, and the nature of reason, as far more precarious. I hope that bringing these essays together helps to make the case for this way of thinking.

Chapter 1, "Introspection and Misdirection," lays out a central strand in my challenge to the Cartesian picture. Descartes viewed a special kind of reflective self-evaluation as absolutely essential for achieving *scientia*, a kind of high-grade knowledge. More recent analytic epistemologists sympathetic with the big Cartesian picture have translated this into a contemporary idiom. On this way of viewing things, reflective examination of the epistemic credentials of a belief is a prerequisite for justified belief, and a necessary condition for knowledge. The prescribed process of reflective self-examination achieves two things simultaneously: it lays out standards for justified belief, and it gives guidance to the concerned epistemic agent. In this essay, I look at the advice offered by two paradigmatic Cartesian epistemologists: Roderick Chisholm's foundationalist prescriptions, and the coherentist demands laid out by Laurence BonJour. Although the advice presented by these philosophers differs from one another quite dramatically, they betray a common faith in the powers of reflective self-examination to improve an agent's epistemic position. This raises the question of whether the kind of introspective self-scrutiny which Cartesian philosophers endorse does, in fact, achieve the desired result.

There is a wealth of material available in the social psychology literature which bears on this question, and it does not bode well for Cartesian

epistemology. As I argue here, when we stop to reflect on the epistemic credentials of beliefs we hold in order to determine whether we hold those beliefs for good reasons, we are subject to a powerful illusion. We seem to have direct access to the reasons for which we hold our beliefs. In fact, however, our understanding of these reasons is a product of confabulation: we engage in a bit of sub-conscious inference in an attempt to explain how it is that we must have arrived at the beliefs we have. As I argue, this does not make our beliefs about our reasons unreliable: inference to an explanation, at least given suitable background information, can be quite reliable. But when we understand how this inference works, we see that it is especially ill suited to the task which Cartesian epistemologists wish to put it. We engage in reflective self-examination in an attempt to locate and correct our errors. The thought is that unreflective belief may easily go wrong, and it would thus be irresponsible simply to trust to luck that such believing operates as it should. We therefore engage in reflection to provide a check on our beliefs: reflection is recruited to serve the purpose of error-detection. Unfortunately, the way in which confabulation works makes it especially unsuited to this role. As I argue, the errors we make in arriving at our beliefs are not likely to be corrected by way of reflection. Instead, we are liable to emerge more confident than we were before we stopped to reflect, but no more reliable. The process of reflective self-examination gives us the illusion that we have subjected our beliefs to a rigorous screening that will improve our epistemic position. In fact, however, it achieves no such thing.

Chapter 2, "What Is It like to Be Me?," further develops the anti-Cartesian views of Chapter 1. On Descartes' view, we have infallible introspective access to our current mental states. More than this, our infallibility is not, according to Descartes, merely a matter of contingent psychological law. In order to do the work it does in his foundationalist epistemology, the infallibility of introspection must be something Descartes can be assured of prior to and independent of any experimental investigations. These extremely strong claims about the epistemic status of beliefs about our mental states are rarely held anymore, but they provide a useful starting point for challenging even the more qualified Cartesian views which are still very much a part of the current philosophical landscape.

I begin by presenting an extended discussion of an example of an individual suffering from paranoid personality disorder. This is not, of course, the kind of case which inspires Cartesian intuitions about self-knowledge, but, as I argue, it has implications even for far more typical cases. The paranoid individual regularly makes judgments about his own mental states – about the extent to which he is defensive, about whether he

is angry, jealous, or insecure – which are horribly inaccurate. It is not that these beliefs are inaccurate because they are insufficiently reflective. Further reflection on these issues would not lead to greater accuracy; indeed, quite the opposite. The belief-forming processes which go to work when the paranoid individual introspectively examines the various features of his mental life are extremely unreliable, and they are unreliable because of the ways in which his paranoia distorts his judgment. Our mental states form a complex causally interconnected system, and reflective processes of self-examination may be facilitated, or, as they are in this case, distorted, by the ways in which our various mental states interact.

When the paranoid individual introspects and examines his emotional state, he may come to believe that he is not angry even though, in fact, he is. His introspective evidence does not alert him to the fact that he is angry. But now you and I may introspect and come to believe, on that basis, that we ourselves are not angry. If we do so because our introspective evidence provides no reason to think that we are angry, then, in that respect, we are no different from the paranoid individual. For this reason, it seems that our introspective evidence, by itself, cannot give us adequate reason to believe that we are not angry.

Now paranoid personality disorder is not a common condition, and one may think that this is sufficient reason to justify ignoring this worry. That may well be. But, as I argue, there are many mental conditions which are extremely common, and which lead to similar sorts of distortions: they lead individuals who suffer from these conditions to misjudge their mental states on the basis of their introspective evidence. The fact that these conditions are very common suggests that you and I cannot reasonably ignore the possibility that we are subject to them, absent specific evidence that speaks to that issue. But then our introspective evidence of our current mental state, which is no different from that of those who suffer from these common mental conditions, and which distorts their view of what their mental state really is, cannot, by itself, justify our beliefs about our current mental states.

I therefore argue for a mitigated skepticism about self-knowledge. The skepticism is empirically based, and the breadth of this skepticism is entirely dependent on the extent to which there are common mental conditions which distort the introspective view of those who suffer from these conditions on particular subjects. Some mental states may elude this argument if there are no common conditions which distort subjects' views of these particular states. But, as I argue, there are many very common mental conditions which distort subjects' views about many mental states about which all of us have a great many opinions. This is compatible with

the suggestion that we may, nevertheless, come to know that we are, or are not, in these mental states, as long as we have the relevant non-introspective evidence that we are not subject to the distorting mental condition. But it does mean that we cannot have knowledge or justified belief about whether we are in these mental states by introspection alone. And it means as well that since many people are in no position to know whether they suffer from the relevant mental condition, that they may not be in a position to know, or even be justified in believing, whether they are in a great many different mental states about which their introspective evidence seems, on its face, to be decisive.

Chapter 3, "Distrusting Reason," addresses a different part of the Cartesian picture. On Descartes' view, it is not just our view of our own current mental states which occupies a privileged epistemic position; our view of the truths of reason are epistemically privileged as well. More recent epistemological views put this point in slightly different terms: it is widely held that the manner in which we should reason is knowable a priori, and thus resistant to many of the sources of challenge which our empirical beliefs are subject to. In this essay, I examine one particular sort of challenge to this view, one according to which it might be wise to be especially distrustful of appeals to reason.

Certain Marxists, Freudians, feminists, and deconstructionists have urged that we should be especially distrustful of explicit appeals to reason. Without seeking to align myself with any of these views, or attempting to reconstruct the arguments such philosophers give for this distrust, I argue that there is a view here which is not only worthy of serious consideration, but which offers an account of explicit appeals to reasons which is superior to the traditional a priorist account. I approach this by considering the ways in which appeals to reason and argumentation play a role in public discussion.

It is often held that when someone publicly defends a claim with which we disagree, we should consider the reasons which are offered in favor of the claim and carefully evaluate whether the reasons offered do, indeed, support the claim at issue. Sincerely engaging with the reasons offered is part of what it is to be open-minded, and it would be unreasonable to reject such a claim out of hand simply because our own belief on the matter is different.[1] It is true, of course, that in attempting to evaluate the claim, we may miss something important, or we may be biased in certain respects so that we fail to see the relevance of certain issues which are, in

---

[1] For a very useful recent discussion of these issues, see Jeremy Fantl, *The Limitations of the Open Mind*, Oxford University Press, 2018.

fact, relevant. But if there is a public discussion of the claim at issue, then it seems that, should we all approach the matter in the manner recommended, these concerns are likely to be of vanishingly little significance. The speaker may be biased or ignorant in some ways, as may you or I. But if we all focus on the reasons as best we can, our biases and areas of ignorance are likely to cancel one another out. You are likely to notice the ways in which I am biased and you are not, or areas in which I am ignorant and you are not, and vice versa. Since we can each recognize one another's shortcomings in these respects precisely by focusing on the reasons for and against a particular claim, the public discussion of reason is likely to be especially epistemically useful. We should, as a result, be especially trusting of the public discussion of reason.

I think that considerations of this sort are very much at work in explaining the confidence that is often placed in the public discussion of reason, and I believe as well that, in certain sorts of situations, considerations such as these count in favor of such confidence. But this defense of the practice of trusting the public discussion of reason makes certain presuppositions about the distribution of bias which cannot always be counted upon. If the sources of bias are many and idiosyncratic, then the kind of argument just given provides reason for thinking that the public discussion of reason will allow us to transcend our individual biases, since the group is likely to be able to correct the various idiosyncratic biases of the individuals which compose it. But if there is some common source of bias which the members of the group are all, or almost all, subject to, then the public discussion of reasons will not uncover it, but, instead, merely entrench it further. As I argue, the idea that there should be such widespread biases is not at all far-fetched. There are certain subjects on which, indeed, this is exactly what we should expect. And here, the public discussion of reasons will be epistemically counterproductive.

This does not, I argue, count in favor of a blanket skepticism about reason, or simply opting out of the public discussion of reasons across the board. But neither is it a reason to simply ignore the problem and hope for the best. We must, instead, take seriously various hypotheses about the distribution of bias, and investigate the extent to which these various hypotheses are likely to be true. We need, as well, to undertake various steps necessary to ameliorate the common sources of bias, biases which are, epistemically, far more pernicious than the idiosyncratic biases which are often a focus of our attention.

Chapter 4, "The Impurity of Reason," takes on the issue of the a priori far more directly, and whether the manner in which we ought to reason is a

priori knowable. Laurence BonJour has provided one of the most detailed and compelling defenses of the a priori in his book *In Defense of Pure Reason*.[2] He famously argues that a complete rejection of the a priori amounts to "intellectual suicide," for it would undermine all legitimate inference.[3] I examine BonJour's arguments in detail here, and defend a view on which the issue of whether we have a priori knowledge turns on certain empirical questions about human psychology, questions which cannot be settled by way of introspection or reflection.

I defend a largely Quinean view of the a priori, although my own presentation of the Quinean argument does not depend on any sort of skepticism about intentional notions, or the alleged indeterminacy of translation or inscrutability of reference. Instead, I focus exclusively on Quine's holistic view of confirmation, and his invocation of various examples from the history of science to illustrate the claim that some of the most paradigmatic examples of apparently a priori knowledge turned out to be false. Claims which at one time seemed not only obviously true, but which seemed immune from the possibility of empirical refutation, gave way in the face of large changes in theory, changes which themselves were justified by surprising empirical evidence of a sort not previously conceived of.[4] This argument, as I see it, undermines one of the central motivations for the a priori.

Unlike Quine, however, I argue that these considerations still leave room for a priori knowledge, but whether we have such knowledge depends on the fate of a controversial empirical claim. Here I follow Georges Rey,[5] Alvin Goldman,[6] and Louise Antony.[7] Should it turn out that we have certain native psychological processes which take no empirical input, and yet reliably produce true beliefs – for example, in the way that would occur were certain theorem-proving inferential tendencies built right in to our cognitive systems – then we should allow that there is, indeed, a priori knowledge. This is just what certain mental logicians, such as Lance Rips, have argued is in fact the case.[8] Even if this were true, however, a priori knowledge so understood could not do many of the

---

[2] Cambridge University Press, 1998.   [3] Ibid., 5.
[4] I take this to be the force of the very brief and suggestive remarks in section 6 of "Two Dogmas of Empiricism," reprinted in Quine's *From a Logical Point of View*, Harper Torchbooks, 1961.
[5] "The Unavailability of What We Mean I: A Reply to Quine, Fodor and LePore," in Jerry Fodor and Ernest LePore, eds., *Holism: A Consumer Update. Grazer Philosophische Studien* 46(1993), 61–101; and "A Naturalistic A Priori," *Philosophical Studies* 92(1998), 25–43.
[6] "A Priori Warrant and Naturalistic Epistemology," *Philosophical Perspectives* 13(1999), 1–28.
[7] "A Naturalized Approach to the A Priori," *Philosophical Issues* 14(2004), 1–17.
[8] *The Psychology of Proof: Deductive Reasoning in Human Thinking*, MIT Press, 1994.

things which it has traditionally been called upon to do. More than this, there is good reason, as I argue here, for thinking that even on this controversial empirical supposition, the a priori would not play anything like the very fundamental role in grounding inference generally which BonJour, and many other epistemologists in the Cartesian tradition, have supposed.

Chapter 5, "What Reflective Endorsement Cannot Do," examines the role which reflection has played in three different philosophical projects. Ernest Sosa has argued that there are two fundamentally different sorts of knowledge, animal knowledge and reflective knowledge, and reflective knowledge is, in important ways, epistemically superior to mere animal knowledge.[9] Reflective knowledge, on Sosa's view, requires not only that we hold a belief as a result of a reliable competence, but that we reflectively endorse holding the belief. A number of philosophers have argued that the nature of inference is best understood by appealing to the role that reflection plays in deliberating about what to believe.[10] On this view, what it is to infer q from p crucially involves believing that p provides reason to believe that q, a mediating belief which reflection provides in the course of deliberation. On this view, inferring q from p requires not only that one's believing p brought about one's belief that q; it requires that one reflectively endorse the transition from p to q. Harry Frankfurt has argued that reflection provides the key to understanding the nature of freedom of the will.[11] Freedom of the will, on Frankfurt's account, requires not only that one's behavior be caused by the interaction between one's beliefs and desires, but, in addition, that one reflectively endorse acting on the basis of the desire which moves one. And finally, Christine Korsgaard has argued that reflective endorsement holds the key to understanding the source of normativity.[12] The normative standards which govern both belief and action carry the force they do as a result of our reflective endorsements.

I argue that in each of these cases, reflection not only does not do the kind of work which these philosophers require it to do, but that it could

[9] See, among others, "Knowledge and Intellectual Virtue," reprinted in Sosa's *Knowledge in Perspective: Selected Essays in Epistemology*, Cambridge University Press, 1991, 225–244; "Reflective Knowledge in the Best Circles," *Journal of Philosophy* 94(1997), 410–430; "Two False Dichotomies: Foundationalism/Coherentism and Internalism/Externalism," in W. Sinnott-Armstrong, ed., *Pyrrhonian Scepticism*, Oxford University Press, 2004, 146–160.

[10] See, for example, Sydney Shoemaker, "On Knowing One's Own Mind," reprinted in his *The First-Person Perspective and Other Essays*, Cambridge University Press, 1996, 25–49.

[11] "Freedom of the Will and the Concept of a Person," reprinted in his *The Importance of What We Care About: Philosophical Essays*, Cambridge University Press, 1988, 11–25.

[12] *The Source of Normativity*, Cambridge University Press, 1996.

not possibly do the work they assign to it. Each of these philosophers, as
I argue, has a view of first-order psychological processes which underplay
their importance, while pairing that view with an account of reflective
processes which give them powers which no psychological process could
possibly have. What is needed for understanding the phenomena which
these philosophers seek to understand is an account of reflection which
demystifies it.[13]

Chapter 6, "Belief in the Face of Controversy," argues that we should
dramatically reduce the confidence we have in our own beliefs when we
find that our epistemic peers hold views which conflict with our own.
Here, too, this is part and parcel of the anti-Cartesian theme which is the
backbone of this volume. As the Cartesian sees it, when we hold beliefs
which, by our own lights, are well supported by reasons, the fact that some
other party sees matters differently should not weigh heavily with us. Our
private reasons carry a kind of epistemic weight that the contrary opinions
of others cannot significantly diminish. The defense of a kind of epistemic
humility which emerges from this essay is an important part of both the
attack on Cartesian ways of thinking about epistemological issues and the
alternative I offer, which are the central theme of this collection of essays.

Chapter 7, "Naturalism vs. the First-Person Perspective," attempts to
provide at least the beginnings of the demystified view of reflection which
I argued, in Chapter 5, is so badly needed. A great deal of work not only in
epistemology, but in philosophy generally, takes the first-person perspec-
tive at face value. Internalists in epistemology, following in the footsteps of
Descartes, see the first-person perspective on belief acquisition and revision
as the key to understanding the nature of justified belief. Accordingly, they
focus on deliberation about what to believe as the natural route to
understanding fundamental epistemological issues, and they examine such
deliberation from the perspective of the deliberator, that is, from the first-
person perspective. Surprisingly, as I show in this essay, even a number of
self-described externalists in epistemology have focused much of their
attention on the perspective of the deliberator, seeing this perspective as
revealing of important features of our epistemic situation. But important as
an investigation of deliberation is for understanding our epistemic lives, it
need not be conducted from the perspective of the deliberator. Social
psychologists have examined the deliberative process in detail, and any
attempt to understand deliberation and reflective self-examination cannot

---

[13] I examine all four of these philosophical projects in far greater detail in *On Reflection*, Oxford
University Press, 2012.