

# INTRODUCTION

## ON HILARY KORNBLITH'S EPISTEMOLOGY<sup>1</sup>

JOSHUA DIPAULO AND  
LUIS R.G. OLIVEIRA

Hilary Kornblith is one of the world's leading epistemologists, a champion of an innovative philosophical research program that is at once traditional and revisionary. In viewing the study of knowledge as inseparable from the empirical study of the mind, Kornblith aligns himself closely with the approach of the traditional empiricists of the seventeenth and eighteenth centuries. Yet in taking contemporary empirical work seriously, Kornblith has developed views and arguments that shift the epistemological focus away from what is available first-personally within the mind—as the early empiricists would have it—towards what is revealed third-personally of the mind instead.

This shift is not a common symptom of his time. When Kornblith entered the philosophical scene in the 1980s, epistemology was dominated by the Gettier problem and its aftermath. Mainstream epistemology, although variegated, embodied a decidedly a priori research program detached from empirical questions of almost any kind. Since then, few have done more than Kornblith to resist this trend by anchoring epistemological inquiry in the contemporary sciences. Indeed, anyone looking seriously at the history of twentieth-century epistemology can draw a straight line from W.V.O. Quine through Alvin Goldman to Hilary Kornblith—all three, in their own ways, champions of a science-centered epistemology. On the innovative leading edge of that tradition, however, Kornblith has developed a systematic rejection of the traditional philosophical method of conceptual analysis (pace Goldman) while establishing a thoroughly empirical and yet unambiguously philosophical approach to the study of knowledge (pace Quine). As such, Kornblith's work is required reading for anyone interested in the structure and nature of knowledge, the structure and grounds for justification, the sources of epistemic normativity, the history and the prospects for a naturalized epistemology, the legitimacy of intuition-based conceptual analysis as a philosophical method, and the significance and reliability of reflection and reasoning.

This volume engages critically and substantively with Kornblith's decades-long development of a scientific epistemology. It collects 15 critical essays written specifically for this volume along with an essay by Kornblith himself in which he replies to that sustained criticism. In this brief introduction, we canvas some of the main themes of Kornblith's work, and then preview some of the content of the individual chapters.

## 1 Knowledge and Justification

Kornblith's naturalism has led him to develop and defend a systematic externalism about knowledge and justification. Briefly, internalism about justification is the claim that only factors internal to, or available from, a person's perspective can make a difference to whether their beliefs are justified. Externalism about justification denies this: it states that justification can depend on factors external to a person's point of view as well. Internalism about knowledge, in turn, is the claim that knowledge requires justification of some internalist kind. And externalism about knowledge is either the claim that knowledge requires no justification or the claim that it requires at most externalist justification instead.

The contemporary externalist tradition in epistemology gathered momentum in the late 1960s and early 1970s with work from Goldman (1967), Skyrms (1967), Unger (1968), Dretske (1970), Armstrong (1973), and Goldman (1979). In different ways, all of these authors contributed to a rapprochement between traditional epistemology and the empirical sciences. Kornblith's critiques of internalism about justification, and his particular brand of externalism about both knowledge and justification, are nonetheless unique in making contemporary science the central guide behind epistemological theory.

### 1.1 On Knowledge

Kornblith's work on knowledge centers on three interrelated claims. First, that knowledge is reliably produced true belief, or belief formed out of an "attunement" to one's environment; second, that knowledge is a natural kind; third, that there is nothing special about the kind of knowledge available to humans, compared to the kind of knowledge available to non-human animals. In characteristic fashion, Kornblith's defense of these claims is premised on his commitment to taking successful empirical science as a philosophical guide.

Kornblith's argument for the first two claims focuses on work by cognitive ethologists: scientists studying the role of mental states in the explanation of non-human animal behavior. He begins by noting that cognitive ethologists theorize about non-human animals with intentional language that implies a belief-desire psychology. They talk of predators *distracting* potential prey in coordinated attacks, of potential prey *hiding from* predators, and so on. He then notes that this is not a bit of informal talk in otherwise careful scientific work. As he puts it:

The attribution of beliefs and desires to non-human animals is neither less explanatory nor less well motivated than it is in the case of humans. In both cases, the attribution of beliefs

and desires demonstrates its legitimacy by way of its role in a theory that provides successful prediction and explanation. (2002, 38)

On this basis—i.e., given that it plays a role in successful scientific prediction and explanation—Kornblith then suggests that we have good reason to accept that non-human animals do indeed have beliefs and desires. (See the section on *naturalism* below for more on the philosophical underpinnings of this inference).

What Kornblith notes next, however, is that the same kind of reasoning should lead us to recognize that non-human animals are capable of knowing too. While talk of beliefs and desires is enough for adequate scientific explanations of individual behavior, that kind of talk is insufficient for adequate explanations of *successful group behavior*:

If we want to explain why a particular plover left its nest and thrashed about in the open while moving away from the nest, we need only appeal to the plover's belief that a predator was nearby and approaching more closely, together with the plover's desire to protect its eggs. In explaining this behavior, it is irrelevant that the plover's beliefs happen to be true ... If we are to explain why it is that plovers are able to protect their nests, we must appeal to a capacity to recognize features of the environment, and thus the true beliefs that particular plovers acquire will be the product of a stable capacity for the production of true beliefs. The resulting true beliefs are not merely accidentally true; they are produced by a cognitive capacity which is attuned to its environment. In a word, they are reliably produced. (Kornblith 1999a, 330–31)

Cognitive ethologists are interested in animal knowledge precisely because it defines such a well-behaved category, a category that features prominently in causal explanations, and thus in successful inductive predictions. If we wish to explain why it is that members of a species have survived, we need to appeal to the causal role of the animals' knowledge of their environment in producing behavior which allows them to succeed in fulfilling their biological needs. (Kornblith 2002, 62)

Part of Kornblith's argument here is much the same as the one above: since talk of animal knowledge by cognitive ethologists plays a role in successful scientific prediction and explanation, we have good reason to accept that non-human animals are capable of *knowing* besides being capable of mere *believing* (see also Kornblith 2021, Sections 2.2 and 6.2). But part of Kornblith's argument is also a naturalistic defense of reliabilism about knowledge: when cognitive ethologists talk about non-human animal knowledge (which we have good reasons to accept as correct) they are talking specifically about reliably produced true belief. So we also have good reason to accept, on the same basis, that knowledge is reliably produced true belief.

This conclusion plays an important role in Kornblith's argument for the second claim characteristic of his work on knowledge. Following Richard Boyd (1988), Kornblith sees natural kinds as “homeostatically clustered properties, properties that are mutually supporting and reinforcing in the face of external change” (2002, 61). More importantly, he sees natural kinds, thus conceived, as essential for explaining the success of our inductive practices: part of why we are able to explain and predict so much in science is the stability of the internal and external characteristics of the properties clustered together into the natural kind categories we use when formulating our explanations and

predictions (see the section below on *inference* for more details here). As we have just seen, however, cognitive ethologists' talk of knowledge is a reference to the *type* "reliably produced true belief" whose *tokens* could not be successfully featured in similarly fruitful hypotheses. What features in the cognitive ethologist's group-level explanation, that is, is not the mere fact that some beliefs happen to be true, but rather the fact that certain true beliefs have the unifying feature of being instances of knowledge. In other words, knowledge is a category cognitive ethologists use when formulating successful explanations and predictions, a category that picks out a cluster of properties (reliably produced true beliefs) with stable internal and external characteristics. So by taking seriously the indispensability of talk of non-human animal knowledge in scientific theories—theories that earn their keep in explanation and prediction—Kornblith concludes not only that knowledge is reliably produced true belief, but that reliably produced true belief is a natural kind in the same way, and for the same reason, that *water* is a natural kind: they both refer to homeostatically clustered properties that support successful inductive inferences (cf., Kornblith 2002, 159).

Kornblith's third central claim about knowledge comes in response to the natural concern that his arguments so far have relied too heavily on the kind of knowledge available to *non-human* animals. There would be no inconsistency, after all, in granting all of Kornblith's claims so far and still insisting that humans are nonetheless characterized by a higher mode of knowing. One way of formulating this idea, following Ernest Sosa (1991, 2007, 2015), is to suggest that while non-human animals have access only to the "attunement" kind of knowledge cognitive ethologists talk about, humans have access to a more "reflective" kind of knowledge, a kind of knowledge that is related more centrally to our unique capacities for self-conscious monitoring and reasoning. On this way of seeing things, one could very much grant the points we have explored above but resist the suggestion, which is central to the significance of Kornblith's position, that those points are relevant to the kind of human knowledge epistemologists are primarily interested in understanding.

Kornblith of course resists these claims. On his view, there is no good reason to insist on a sharp division between "kinds" of knowledge, and no good reason to insist that so-called human knowledge is a "higher" mode of knowing. On the one hand, Kornblith argues that there is no good way to draw a distinction between "attunement" and "reflective" knowing where the latter is both characteristic of humans (as opposed to an ideal rarely reached) and also unavailable to non-human animals (cf., Kornblith 2004, 127–30; Kornblith 2021, §2.3). He also argues, further, that no way of drawing this distinction supports a division between genuine "kinds" of knowledge (cf., Kornblith 2012, 18–20). On the other hand, Kornblith challenges the idea that adding reflection to attunement increases the epistemic value of some mode of knowing, the value it has from the perspective of truth or accuracy:

The idea, then, that by reflecting on the source of our beliefs, we may thereby subject them to some sort of proper screening, and thereby improve on the accuracy of the resulting beliefs, is simply misguided. When we reflect in this way, we get the impression that we are actually providing some sort of extra screening of our beliefs, and we thus have the very

strong impression that we are actually doing something to assure that our beliefs are, indeed, reliably arrived at. But this is not what we are doing at all. Instead, we are engaged in a process which, in a very wide range of cases, makes us more confident that we are right than before we began, but a process, nevertheless, which is almost useless for improving our reliability. In a large class of cases, the process of reflection is an exercise in self-congratulation. It does nothing, however, in these important cases, to improve on the accuracy of our first-order beliefs. (Kornblith 2012, 24–5)

So not only is it not true that there is some kind of reflective knowledge that is more characteristic of humans than the ordinary kind of attunement knowledge available to many non-human animals, it is also not true that would-be reflective knowledge is a “higher” mode of knowing (see the section on *reflection* below for more on this point).

## 1.2 On Justification

Traditionally, as Kornblith puts it, internalism about justification has taken the following general form:

What is required for a person to be justified in holding a belief is for that person to have a certain justification for the belief, where having justification is typically identified with being in a position, in some relevant sense, to produce an appropriate argument for the belief in question. (Kornblith 2001, 2).

This kind of internalism has deep Cartesian roots, and it answers to a familiar and sensible concern: since we wish to have true beliefs but recognize that we are often wrong, responsible epistemic conduct requires the identification of epistemic practices that allow us to do the best we can to avoid mistakes, and being in a position to produce appropriate arguments for our beliefs seems just the right thing for that kind of job.

In his first published paper in epistemology, Kornblith (1980, 599) dubbed this the “arguments-on-paper thesis.” More fully, and in today’s terms, we can say that the arguments-on-paper thesis is equivalent to the following widely held conjunction of commitments: (1) propositional justification (the state of “having justification for believing that *p*”) is a matter of the logical and/or probabilistic relations that hold between *p* and other propositions internal to (or available from) *S*’s perspective, and (2) doxastic justification (the state of “being justified in believing that *p*”) is a matter of having propositional justification for *p* as well as basing the belief that *p*, in some appropriate way, on that propositional justification. On this kind of view, propositional justification is both independent from, and more fundamental than, doxastic justification. Kornblith’s criticisms of internalism and the arguments-on-paper thesis—now spanning more than 40 years—amount to a sustained attack on these commitments, as well as on the impression that they respond to the familiar and sensible Cartesian concerns.

In some of his earlier work, for example, Kornblith has argued that a responsibilist conception of justification is in tension with the picture of doxastic justification that we find

in the arguments-on-paper thesis. On the one hand, the kind of responsibility connected to epistemic justification has to do with the management of our epistemic life: “an epistemically responsible agent desires to have true beliefs, and thus desires to have his beliefs produced by processes which lead to true beliefs” (1983, 34). On the other hand, responsible epistemic management—much like responsible management in general—is not a matter of piecemeal quality-checks, but rather a matter of developing and maintaining good habits:

People are thus responsible for acquiring good epistemic habits just as much as they are responsible for acquiring other good habits. Of course, in acquiring a habit, one acquires a tendency to move from mental state to mental state, and these transitions are not always accessible, much less known, to the person who goes through them. (Kornblith 1982, 246)

So if being justified is a matter of believing responsibly, and believing responsibly at least sometimes has nothing to do with basing beliefs on good available arguments, then the arguments-on-paper thesis’ view of doxastic justification must be incorrect. (See Kornblith 1986 for a naturalized account of the “epistemically responsible agent”.)

In his later work, Kornblith has extended this concern with the relevance of propositional justification for doxastic justification into an argument against the picture of propositional justification that we find in the arguments-on-paper thesis as well. The problem here is that the role that an apychological notion of propositional justification is supposed to play on that picture (being the basis for doxastically justified belief) would render most of our beliefs unjustified, since empirical evidence shows that “the kinds of inductive inferences humans tend to make do not answer to the standards of traditional a priorist ideas about propositional justification, nor do they often approximate them” (2017, 74–5). So insofar as we wish to avoid generalized skepticism while taking propositional justification as a necessary condition for knowledge—as we must, if we take doxastic justification as a requirement for knowledge and take propositional justification as a requirement for doxastic justification—we must in turn reformulate our understanding of propositional justification in a way that matches our best empirically informed understanding of the psychology of belief.

With this concern in mind, Kornblith reverses the underlying assumption behind the arguments-on-paper thesis that propositional justification is both independent from, and more fundamental than, doxastic justification. Following Goldman (1979), Kornblith (2022, 55) suggests that “a proposition is propositionally justified for a person just in case that person has available psychological processes which could reliably produce belief in that proposition absent any additional sensory input.” In response to the worry that such a reversal—by doing away with the fundamentality of logical and/or probabilistic relations that hold between  $p$  and other propositions internal to (or available from)  $S$ ’s perspective—has disastrous consequences to talk of evidence and epistemic reasons in epistemology, Kornblith (2015) argues that careful attention to the psychology of belief shows that this concern is overblown.

Finally, even if the arguments-on-paper thesis managed to survive as a workable account of epistemic justification, Kornblith has argued that it would fail to measure up

to its internalist motivations. The piecemeal quality-check kind of responsibility for individual beliefs that motivates internalism—in contrast to the habit management kind of responsibility for epistemic conduct that motivates Kornblith's externalism—cannot be satisfied by the arguments-on-paper thesis. According to Kornblith (1986, 1988, 2003), the internalist view of epistemic responsibility is untenable without presuming some kind of Cartesian incorrigible access to the contents of our minds. Once this kind of incorrigibility is correctly rejected as illusory—as even most internalists now do—the dream of piecemeal quality-check responsibility for individual beliefs simply falls apart, and with it the central motivation for internalism about justification.

## 2 Inference and Reflection

Internalism in epistemology is partly motivated by a much deeper philosophical sensibility regarding the central value of human reflection, deliberation, and conscious inference. Aside from its epistemological roles, reflection, for example, has been the basis of claims about the nature of the self and the source of human worth. Similarly, aside from its epistemological roles, deliberation has been the basis of claims about the nature of morality and the character of a just society. (Some of Kornblith's targets on these themes include Sosa 1991, Shoemaker 1996, Frankfurt 1988, Korsgaard 1996, and Velleman 2005.) Part of Kornblith's case for externalism, therefore, consists in a host of arguments that aim to show how reflection and deliberation simply cannot play some of the exalted roles they've been assigned both by philosophical tradition in general and by internalism in particular. Kornblith's work here (e.g., Kornblith 2012) challenges a broad philosophical orthodoxy well beyond the narrow boundaries of traditional epistemology.

Favoring the third-person perspective over the first-person perspective in epistemological theorizing, Kornblith rejects the traditional Cartesian picture that treats reflection and deliberation as the epistemically safest starting points for the responsible expansion of our knowledge. Once again, Kornblith is here reversing the traditional order of things, this time by treating forms of self-knowledge as much less secure than knowledge of the external world. For Kornblith, it is our accumulated knowledge of the external world that expands responsibly, and *inwardly*, by taking reflection and deliberation as objects of scientific investigation. This reversal is supported, according to Kornblith, by the comparative success of the empirical sciences. Influential brands of internalism tell us that our unreflective beliefs are not justified if they cannot withstand critical reflection and the private or public deliberative process of giving and taking reasons. In reply, Kornblith asks what's so special about reflection and deliberation, when it comes to the goal of truth? After all, they themselves may be unreliable belief-forming processes. And, indeed, Kornblith has spent a good part of his career examining empirical evidence from social psychology and behavioral economics and arguing, on that basis, that we have strong reason to distrust the outputs of these processes, as they are susceptible to predictable distortions that cannot be detected from the inside.

## 2.1 On Inference

In his first book, Kornblith (1993a) develops a third-personal account of how inferential, especially inductive, knowledge is possible if it is not to be gained simply by conforming to internally accessible canons of logic or probability theory and if it is not to be certified by conscious reflection. In short, Kornblith argues that inductive knowledge is possible because the world is made up of natural kinds and humans reliably tend to (1) detect these kinds and their essential features and then (2) make inferences (e.g., predictions and explanations) that project these features.

He begins his defense of this view by articulating his fundamental conception of epistemology:

I see epistemology in general, and Quinean epistemology in particular, as addressed to two questions: (1) What is the world that we may know it?; and (2) What are we that we may know the world? (1993a, 2)

Blatantly empirical questions, Kornblith turns to science to answer them. Of course, these two questions presuppose that empirical knowledge *is* possible, that the world is knowable and that *we* may know it. Why grant this?

The standard Quinean line here appeals to evolutionary theory, nicely captured in Quine's pithy remark that, "creatures inveterately wrong in their inductions have a pathetic but praiseworthy tendency to die before reproducing" (1969, 126). In other words, had humans *failed* to have reliable inferential tendencies (either native or regularly acquired through human development), they simply would not have survived the process of natural selection; human minds, since they exist, must be adapted to their environments. However, following Stich (1990, 55–74), Kornblith rejects this argument for the simple reason that it is doubtful that evolutionary processes must provide us with an accurate understanding of the world.

Instead, Kornblith (1993a, 3) directly cites the success of science itself as the primary reason for thinking that our psychologies enable us to know the world. The fact of science's success, it should be noted, features as a major premise in many of Kornblith's arguments. In this case, Kornblith takes the massive success of science to show that we *do* have empirical knowledge, and hence that we *can* have empirical knowledge. What then of evolution? Pace Quine, it does not *demonstrate* that we can know the world; rather, it *explains why* we possess the ability to know the world.

The striking fit between our psychology and the structure of the world stands in need of explanation. It is surely no coincidence. The best explanation for the extent of fit, to my mind, is that it is a product of the evolutionary process. Evolution should not be called upon as evidence that our psychology fits well with the causal structure of the world, for the fact of good fit is independently established. Evolution is only called upon *after* we establish the fit between our psychology and the world, as an explanation of how that fit came about. This is different from the account typically attributed to Quine, but it is, I believe, far more plausible. (1993a, 3)

Because processes of inductive inference are largely automatic and unconscious, Kornblith turns to experimental psychology and cognitive science, rather than more traditional internalistic means of reflection, to answer questions about how inductive inferential knowledge is possible (cf., 2021, 56). Much of Kornblith's work on this issue treats the early conclusions of Tversky and Kahneman's (1971, 1973) famous work on the psychology of judgment as a foil. On the basis of their findings, Tversky and Kahneman articulated a decidedly pessimistic and skeptical view of human inference; they argued that our "intuitive expectations are governed by consistent misperception of the world" (1971, 31). In Kornblith's words, Tversky and Kahneman's perspective entails that:

There seemed to be a remarkable uniformity in the kinds of inferences people make, and those inferences were not, by any reasonable standards good ones. People reason badly; the inferences we make regularly lead us to false beliefs; we're deeply irrational by our very nature. Inferences of this sort, quite clearly, are not a source of knowledge. (2021, 58)

However, this conclusion is at odds, Kornblith argues, not only with common sense and with every day experience, but also with the success of science. Laypeople have plenty of inferential knowledge, but more than this, the prevalence of scientific knowledge would make no sense if inductive inference were so unreliable.

The problem with Tversky and Kahneman's picture, according to Kornblith, is that it measures the success of our reasoning against the standards of probability theory. When subjects in their experiments fail to conform to these standards—not just laypersons, it should be noted; plenty of people with *training* in probability theory as well—Tversky and Kahneman conclude that their subjects' reasoning is therefore flawed. Moreover, their picture assumes, wrongly Kornblith argues, that regularly violating these standards will cause one to end up with more false beliefs than true ones.

In contrast, rather than searching for necessarily true canons of logical inference, Kornblith instead argues that our inferential knowledge is in a sense, unabashedly, contingent. As a matter of fact, humans reason in ways that *tend* to lead to true conclusions in *our* world. One of the great sins Tversky and Kahneman accuse humans of committing is violating the Law of Large Numbers by making predictions on the basis of very small samples. But if we tend to project the "right" features on the basis of these small samples, then we can reliably end up with true beliefs. In other words, if we have inferential tendencies that make presuppositions that, while not true of every possible environment, tend to be true of environments in our world and we thereby reason from those presuppositions, then we can have inductive knowledge that violates canons of statistical inference. And this is just what Kornblith argues is the case:

As it turns out, I have argued, we are quite adept at detecting the very features of natural kinds which are essential to them, and our conceptual structure places these essential features in the position of driving inductive inference. As a result [when we violate the Law of Large Numbers], we typically project the properties of natural kinds which are universally shared by their members. It is thus that our inductive inferences are tailored to the causal structure of the world, and thus that inductive understanding of the world is possible. (1993a, 107)

What of all the “reasoning errors” that Tversky and Kahneman uncovered in their experiments? On the one hand, Kornblith denies that they are indeed all errors. But, on the other hand, he of course grants that we do make inferential errors, even fairly uniform ones. However, he views these mistakes as on a par with perceptual illusions. Just as no one would conclude that the fact that we succumb to visual illusions would demonstrate that our vision is unreliable, so too no one ought to conclude that the Tversky and Kahneman (and similar subsequent) findings demonstrate that human inference is unreliable. Instead, these findings help illuminate the presuppositions that our minds tend to make, presuppositions that may be false in certain actual environments while nevertheless enabling reliable inference generally (Kornblith 1993a, 85–87).

## 2.2 On Reflection

Uniting a great many otherwise disparate and distinctive perspectives in philosophy is a shared commitment to the significance and power of the human capacity for reflection. We do not merely act and believe; we deliberate about whether we have good reasons to act and believe as we are inclined to, or at least we typically have the capacity to engage in such deliberation. Many familiar philosophical views take the *possession* or the *exercise* of this capacity as characteristic of—if not a full-on necessary condition on—the possession of certain other valuable states, e.g., the states of being a person or being justified in believing a particular proposition or possessing an exalted form of knowledge (see the sections above on *knowledge* and *justification*).

On Kornblith’s view, developed over the last four decades, this shared commitment might be better described as *mere dogma* or an *article of faith* rather than a well-supported, defensible position that coheres with empirical findings. Less pointedly, Kornblith has argued that reflection simply cannot do the work it’s enlisted to do by these theories. First, he identifies certain logical barriers here; regress looms, for instance, if critical scrutiny of beliefs is required for those beliefs to be justified—what of the beliefs about those beliefs, and the beliefs about the beliefs about those beliefs, etc.? Furthermore, characteristic of his naturalistic approach, Kornblith also argues that given how reflection actually operates, as revealed through experimental psychology, it simply cannot do what many philosophical views require it to do if it is to have the significance these views attribute to it.

For the sake of illustration, consider what many, including Kornblith, take to be an attractive picture of the role reflection ought to play in belief management.

We typically arrive at our beliefs unreflectively, and when we do so, the processes by which our beliefs are produced are often quite unreliable. After all, we sometimes engage in wishful thinking; we form conclusions, at least at times, quite hastily and on the basis of remarkably little evidence; we are subject to errors and biases in our inferences. For these reasons, resting content with unreflective belief acquisition seems not only lazy, but irresponsible. Anyone who cares about having an accurate understanding of the world around us should thus stop to reflect, both on the beliefs he or she has, and on the reasons for which they are held. By holding one’s beliefs up to reflective scrutiny in this way, we may substantially improve upon the accuracy of our beliefs. (Kornblith 2012, 1)



It would be no more reasonable to hold [that we should never reflect on our beliefs and their epistemic status] than it would be to hold that reflecting on our beliefs can cure all the ills that might beset belief unreflectively arrived at. What we should value, instead, is reliably acquired belief, whether it is acquired with or without the aid of reflection. And this is just to say that when it comes to belief acquisition and retention, reflection is just one more process among others, to be evaluated in just the same way as any other. (Kornblith 2012, 41)

This may seem odd, though. Why would it ever be a good idea for you to reflect on your beliefs if you're likely to end up biased, overconfident, and self-congratulatory? One obvious response is that it's simply an empirical matter *if* and *when* reflection will tend to have these results for individuals; there is no *a priori* guarantee that reflection will have these bad results in all circumstances. But Kornblith's more recent work develops a more positive answer to this question. While Kornblith (1999c) warned of the negative effects individual biases and unreliable reflection could have on the public use of reasons and argumentation, following work on the *social* function of reason and reflection by social psychologists Mercier and Sperber (2017), Kornblith (2021) now argues that what looks like unreliable reflection at the individual level has the potential to contribute to a reliable process of belief management at the social level.

The central insight arises when we ask *why* human reflection tends to be so "lazy and biased." In other words, why might we have the sorts of apparently bad intellectual habits documented in the psychology literature that Kornblith appeals to in his case against reflection? Recall the analogy Kornblith draws between perceptual illusions and inductive errors (Section 2.1). Just as the study of systematic perceptual error reveals how perception succeeds, and the study of systematic inductive error reveals how induction succeeds, might the study of systematic reflective error reveal how reflection actually succeeds? This is, indeed, where Kornblith's views on reflection have recently landed.

Mercier and Sperber have an hypothesis. What they suggest is that the capacity to reflect was not selected for private reflection. Rather, the capacity to reflect was selected for use in social interactions. In particular, social creatures frequently engage in cooperative behavior, and the capacity to reflect on our reasons aids cooperation. (Kornblith 2021, 111)

The basic thought, here, is that lazy and biased individual reflection has been selected not for its reliability in individual cognition, but for its contribution to group inquiry and deliberation. Kornblith illustrates this thought with the following example:

I propose a plan, and you suggest an alternative. We each need to offer reasons in favor of our proposed plan, and this is where reflection comes in. We each reflect on our reasons for favoring the plan we've proposed, and given how reflection works, we will each offer reasons which we ourselves find convincing. We won't be very good in spotting problems with our own reasoning because the way in which we reflect on our reasons is governed by confirmation bias. But I will be quite good in spotting difficulties with your plan, should there be any, and you will be quite good in spotting difficulties in my plan, should there be any. Confirmation bias, in this social interaction, serves to divide the intellectual labor: I do the work of spotting defects in your plan, something you will not be very good at; and you do the work of spotting defect in mine, defects which I will have been unlikely to notice. The net effect of such a

dialogue, and such a division of intellectual labor, is that we will be able to converge on a judgment about which plan is better, and we will not only reach agreement on this, but the plan we ultimately agree on is likely, in fact, to be better than either of our original plans as a result of being subjected to assessment in this way. (Kornblith 2021, 113)

What goes for practical deliberation and planning, goes for epistemic deliberation as well. Biased individuals, poor at uncovering problems within their thinking *on their own*, can come together to collectively arrive at better beliefs by uncovering problems with each other's thinking.

Of course, optimism here must be measured. Discussion within plenty of groups simply functions to reinforce group members' beliefs, just as individual reflection, according to Kornblith, tends to be self-congratulatory. That being said, depending on the goals and makeup of a particular group, unreliable individual reflection has the potential to contribute to emergent reliable group reflection.

When group members are, by and large, engaged in an enterprise which the majority is committed to, and which they regard as an attempt to figure out a solution to a problem rather than an attempt to get their own way, the Mercier and Sperber hypothesis will apply. Not all groups are like this, but many groups are. The Mercier and Sperber hypothesis is a substantive claim about such groups, and it is a substantive claim about the function of our capacity to think about and evaluate our reasons. (Kornblith 2021, 119)

One hope, of course, is that when there is a net positive increase in accuracy at the group level this will “trickle down” to individuals, leading them to adopt more accurate beliefs themselves. After decades of systematically defending a thoroughly pessimistic stance on the powers of reflection, this slight shift in an otherwise unwavering consistency in Kornblith's position may itself provide a tiny bit of anecdotal evidence to think this hope is not utterly groundless.

### 3 Naturalism, Normativity, and Methodology

In broad terms, Kornblith (1985, 3) takes naturalism in epistemology as the view that the question “How ought we to arrive at our beliefs?” cannot be answered independently from the empirical question “How do we arrive at our beliefs?” While epistemological questions constitute a legitimate kind of inquiry, that kind of inquiry is not—pace traditional epistemology—insulated from empirical findings. Kornblith (1994, 49) is clear about the rationale behind this kind of naturalism:

What does have priority over both metaphysics and epistemology, from the naturalistic perspective, is successful scientific theory, and not because there is some a priori reason to trust science over philosophy, but rather because there is a body of scientific theory which has proven its value in prediction, explanation, and technological application. This gives scientific work a kind of grounding which no philosophical theory has thus far enjoyed.

There are two connected ideas here. First, that success in prediction, explanation, and technological application is a good indication of truth. This is a claim familiar from the extensive literature on the so-called “no miracles argument” (the idea that the best explanation for the success of the sciences is that their claims are approximately true). Second, this kind of success, enjoyed almost exclusively by the mature sciences, is a far better indication of truth than any other kind of success enjoyed by pure philosophy. How Kornblith’s epistemological naturalism takes shape is now clear: since epistemologists are interested in all aspects and dimensions of (and in all phenomena related to) true belief, they do well by taking their lead from the kind of inquiry that has seen the most success in that regard.

We have documented a variety of ways in which Kornblith’s epistemology takes its cue from the empirical sciences. Here we discuss the connections that Kornblith sees between his naturalism and questions about normativity and methodology.

### 3.1 On Naturalism and Normativity

One of the traditional complaints about naturalistic epistemology is that it turns a *properly normative* discipline into a *merely descriptive* one. This was a central part of Jaegwon Kim’s 1988 famous broadside against Quine’s 1969 “Epistemology Naturalized.” After all, the thought goes, since science trades in descriptions and not in norms, epistemic theorizing limited to the resources of a scientific worldview would be unable to engage in the central epistemic project of evaluation. Epistemology done this way, in other words, would lose its relevance for traditional questions about, for example, *reasonable* belief (Descartes) and the *value* of knowledge (Plato).

Despite his naturalism, Kornblith does not deny that epistemology is a fundamentally normative discipline:

If you tell me that a belief of mine is unjustified, this gives me reason to give up that belief. The epistemic claim is something about which I should care, and an account of the source of epistemic norms must explain why it is that I should care about such things. (Kornblith 1993b, 363)

Philosophers are interested in a category that has normative implications. To say that a belief is an item of knowledge is to praise it in a certain way; it is to approve of it as meeting our cognitive ideals; it is to recommend it. (Kornblith 2002, 159)

What Kornblith denies, instead, is that accounting for this normative dimension requires abandoning philosophical naturalism.

Take first the question of reasonable belief. As we have already seen (see section above on *Justification*), Kornblith is interested in the notion of *epistemic responsibility* that underlies many accounts of reasonable or justified belief. The typical connection here is that “reasonable” belief is “justified” belief, and that justified belief is epistemically “responsible” belief: belief that measures up to the norms that are relevant from the epistemic point of view, namely, the point of view that takes *truth* as its central value. Kornblith’s aim is to provide an account of this notion that fits neatly within the naturalistic approach.

Kornblith here suggests a reductive and descriptive account of epistemic responsibility:

An epistemically responsible agent desires to have true beliefs, and thus desires to have his beliefs produced by processes which lead to true beliefs; his actions are guided by these desires. (Kornblith 1983, 34)

An epistemically responsible agent, on this account, certainly cares about both good reasoning and process reliability. But doing the best one can, in light of these desires and concerns, sometimes requires more than just good reason and reliability, and is nonetheless compatible with falling short of both. Understanding epistemic responsibility, for Kornblith, is less a matter of some kind of a priori deference to good reasoning and reliability as criteria to be matched, and more a matter of carefully studying the conditions and considerations that pertain to real human subjects interested in acquiring true beliefs.

Take now Plato's question about the value of knowledge. Kornblith here offers a novel and ingenious desire-based account of epistemic normativity—of the reason-giving force of truth-oriented norms—that doubles as an explanation of the value of knowledge.

Some desire-based accounts of normativity in general begin by determining which values are the correct values to have. And some desire-based accounts make epistemic normativity, in particular, contingent on our actually having a desire for truth. The problem with the first kind of account is that it seems in tension with a thoroughgoing naturalism: it is the *correctness* of the desire, after all, and not the fact of the desire itself, that does the normative heavy-lifting. The problem with the second kind of account, on the other hand, is that it fails to deliver on the universal force of epistemic norms: some of us don't have a desire for truth, or don't have a desire for truth that overrides our competing desires, or don't have an overriding desire for truth regarding everything and anything at all the times.

Kornblith, instead, centers on the fact that successfully satisfying our desires—whatever these may be—requires a decent grasp of the truth about the world around us. Take the full variety of human desires. Forget about whether someone's desires are good or bad, and forget about what exactly they desire. If someone desires X, for any non-trivial X, then successfully acquiring X, or successfully bringing it about that X, requires taking the necessary means for X. In turn, taking the necessary means for X, at least in the non-accidental way that leads to the regular successful promotion of our desires, requires having *true beliefs* about those necessary means. If I have terribly mistaken beliefs about my kitchen, for example, I can't expect to get the milk from the fridge for breakfast every day.

Surprisingly, what this means is that *all of us*, no matter what we each desire, have instrumental reasons to care about having a decent number of true beliefs and a decently reliable cognitive system (one that produces a decent number of true beliefs about the variety of contexts wherein we seek to promote our various desires). And given Kornblith's naturalistic defense of knowledge as reliably produced true belief (see the section above on *knowledge*), what this means is that we *all* have instrumental reasons to care about knowledge:

Knowledge is of extraordinary instrumental value, for it allows us to achieve our biologically given goals, as well as our more idiosyncratic individual goals, whatever those goals may be. This makes it a condition that is universally valuable. (Kornblith 2002, 160)

Knowledge has a call on us, that is, it is in the interest of creatures with beliefs and desires to attain knowledge, whether they otherwise value knowledge or not. Given my account of the source of this demand, this thereby makes for real normativity in nature. (Kornblith 2002, fn. 25, p. 160)

This is different from saying that we all *must* value knowledge, and different from insisting that we all in fact *do* value knowledge. What Kornblith shows instead is that knowledge is *valuable for us*, no matter what else is going on in our psychology (at least so long as those goings-on include desires).

### 3.2 On Naturalism and Methodology

Kornblith's naturalism also has implications for philosophical methodology. The dominant methodology in the last quarter of the twentieth century was without a doubt a mix of conceptual analysis and reflective equilibrium: if you want to know what X is—e.g., *knowledge*, *free will*, *personal identity*—then you should (1) articulate the contours of your concept of X in terms of necessary and sufficient conditions, and (2) test those conditions against your intuitions about possible cases related to X. Even Alvin Goldman (1998, 2010), whose reliabilism and overall science-minded approach were both significant influences on Kornblith, famously championed this method of conceptual analysis and reflective equilibrium. Indeed, in some form or another, this method has been the unquestioned backbone of analytic philosophy since its very inception. Kornblith's work, however, represents a wholesale rejection of this approach.

On one way of thinking, this traditional method seems to mirror the methodology of science: it tests definitions against our intuitive verdicts regarding various thought experiments in a way that is analogous to how scientists test theories against experimental observations. On this way of thinking, our intuitions seem to be evidence for or against some philosophical theory in the same way that observations are evidence for or against some scientific theory.

The problem with this way of thinking is that the analogy breaks down where it matters the most: both scientists and philosophers are primarily interested in real phenomena in the world, and not primarily interested in our current beliefs about it. There is room for attention to the latter in our investigations about the former, but the former, not the latter, is the real target. The method of conceptual analysis, however, turns our beliefs about real phenomena (our concepts) into the primary interest and final target of philosophical investigation. This is a mistake. As Kornblith (2002, 9–12) puts it, what matters is *knowledge itself*, the actual phenomena in the world, and not merely our *concept of knowledge*.

This is not to say that our conceptions and intuitive verdicts have no place in philosophical methodology. It is rather to say that they have the same place in philosophy that they have in science: they are not primary evidence; they merely give us our initial bearings towards a subject matter. This, then, is the proper way to make philosophical methodology analogous with the sciences:

Proper method in philosophy is not so different from proper method in the sciences. We find some phenomenon which catches our attention and which we seek to understand better. We begin by looking at what seem to be clear cases of that phenomenon and we try to figure out, by examining the phenomenon itself, what it is that these cases have in common. We don't consider wild hypothetical cases which are utterly unlike anything that might arise in nature; rather, we look at the kinds of cases which actually occur and subject them to scrutiny. This is how chemists and physicists were able to come to an understanding of what water is, and what gold is, and what heat is. What I have been suggesting is that we may proceed in just the same way in philosophy. (Kornblith 2021, 140)

And since science is our best known guide to truth, once there is empirical work available on some particular topic, philosophers best privilege it even over time-honored intuitions.

We have seen this “science-first” approach above while canvassing Kornblith's work on knowledge, reflection, and inference. It should then be clear that it is far from the bogeyman naturalism suggesting that proper epistemology “falls into place as a chapter of psychology” (Quine 1969, 82). It is rather the promising suggestion that philosophy is “continuous” with the sciences, continuous with the most successful intellectual enterprise known to humans:

In recognizing that philosophy is continuous with the sciences, we need not fear that philosophy will thereby be 'eclipsed' by science. The constraints that science presents for philosophical theorizing should be welcomed, for philosophical theorizing unconstrained by empirical fact loses its connection with the very phenomena which we, as philosophers, seek to understand. Philosophy is an autonomous discipline, in the sense that it addresses a distinctive set of questions and concerns, and in this respect it is no more nor less autonomous than physics or chemistry or biology. This is surely all the autonomy we should want. It is, in any case, all the autonomy that we may have. (Kornblith 2002, 27)

It is hard to see what more one could have in terms of aspiration, and how one could do better in terms of self-conception.

Kornblith's naturalism, and his science-first methodology, also undermine two familiar features of traditional epistemology. First is the over-reliance on everyday examples. Second is the over-emphasis on skeptical possibilities.

Much of traditional epistemology starts by considering everyday cases of perceptual and testimonial beliefs, cases where we are confronted with a direct perception of a medium-sized material object (e.g., a tree in the yard) or a with an ordinary testimonial report about a clear-cut topic (e.g., directions to some city's landmark). These cases are then examined in an attempt to reveal the nature of epistemic justification and knowledge in a way intended as generalizable to cases of any other kind.

On Kornblith's view, however, this approach mistakenly places emphasis on cases where epistemological issues are not as central as in cases of scientific inquiry, and therefore not as instructive:

One of the distinctive features of scientific theorizing is the self-conscious consideration of evidence and its bearing on theory. Theorists are deeply concerned with the adequacy of the

evidence in favor of candidate theories and the comparison of different candidate theories against one another. Unlike everyday situations, where self-conscious theorizing is rare, the weighing of evidence is arguably routine in scientific contexts. The act of justifying beliefs is thus more a feature of scientific practice than it is a feature of everyday perceptual situations. One might therefore think that the notion of justification finds its natural home in scientific contexts, that this is where it is to be examined and best understood. (Kornblith 1999b, 30–31)

What we see here, once again, is part of the case for favoring a focus on science as the locus of philosophical investigation. The point here is not so much that philosophy should take into account the results of scientific inquiry (a point that we have seen Kornblith defend and put to good use above), but rather that scientific inquiry should itself be a topic of investigation for epistemologists.

Similarly, much of traditional epistemology starts by considering “skeptical challenges,” those concerning suggestions that even our everyday beliefs might fail as cases of knowledge or justified belief. The work of epistemology, on this approach, consists in showing how knowledge or justified belief are even possible, in light of the possibility that we are being deceived by some evil demon, or some evil scientist, or some computer programmer.

Kornblith has two reasons to minimize the value of this approach. First, because it is a lost cause:

it is a mistake to try to respond to the skeptic on his own terms. That is, it is a mistake to try to convince someone who denies knowing anything at all, and denies that anyone else knows anything at all, that they are wrong about this. We can enter into rational dialogue with others, trying to convince them that they have made some mistake, only if they have some beliefs already and some forms of inference which they accept as legitimate. Without any such starting place for rational discussion, it is impossible to give someone who disagrees with you reasons for changing their mind. The skeptic offers us no such starting place, no point of entry for rational engagement. This makes it pointless to try to convince the skeptic, as has often been recognized. It also means that the fact that we cannot rationally convince the skeptic that knowledge is possible is entirely unrevealing about the nature of knowledge. (Kornblith 2021, 130–31)

Second, because it is bad methodology:

We may best make progress in understanding the nature of knowledge by looking at the phenomenon of knowledge itself, rather than focusing on skeptical scenarios in which knowledge is not only absent but in which it cannot be attained. (Kornblith 2021, 131)

Just as scientists examine some natural phenomena by primarily considering instances of it (and not instances of their absence), epistemologists should instead begin by considering cases where knowledge is present.

Importantly, of course, what we learn from examining the phenomenon of knowledge itself is that there is much knowledge around us: reflective and unreflective knowledge in adults, and unreflective knowledge in prelinguistic infants and non-human animals (cf., Kornblith 2021, ch. 2). None of this means that Kornblith is some kind of extreme Moorean or anti-skeptic. As the case of reflection makes clear, it would be fair to describe

some of Kornblith's views as at least moderately skeptical. It's just that in his view the wild hypothetical scenarios found in much traditional philosophy in general, and epistemology in particular, do little to nothing to ground reasonable forms of skepticism—or reasonable forms of knowledge, for that matter. Knowledge of the extent and limits of knowledge is to be found where knowledge is most secure: science. And it is hard to think of someone other than Hilary Kornblith who has better brought together our best philosophical and scientific insights on these and related topics.

## 4 Summary of Chapters

Kornblith's work spans more than four decades of philosophical contributions. It represents a unique and influential voice in contemporary epistemology. In Sections 1–3, we have presented and explained some of the main features of his work not only in hopes of enticing readers into newfound interest in Kornblith's naturalistic epistemology, but also to prepare them to engage with the critical essays contained in this book. In this final section, we provide a brief description of each of those essays.

### 4.1 Knowledge and Justification

In Chapter 1, “Knowledge of Knowledge and Its Place in Nature,” Daniel Greco argues that Kornblith's own methodology undermines some of his central commitments. Drawing on formal models from economics and evolutionary game theory, Greco argues that these models support structural principles about knowledge that externalists like Kornblith typically reject—especially strong connections between first-order and higher-order epistemic states. Furthermore, the pluralism about modeling in these disciplines complicates the idea that knowledge is a natural kind with fixed, interest-independent boundaries. Instead, Greco suggests that what it is useful to count as knowledge depends on our explanatory aims. This pluralistic, model-based perspective, Greco argues, both extends and subverts Kornblith's naturalism, showing that the most empirically productive conceptions of knowledge are more flexible and context-sensitive than his own framework allows.

In Chapter 2, “Kornblith, Naturalism, Relativism,” Martin Kusch and Robin McKenna examine three core commitments that run through Kornblith's work in epistemology. First, epistemologists should investigate knowledge itself rather than the concept of knowledge. Second, knowledge is a natural kind. Third, knowledge is reliably produced true belief. These commitments are related in several ways, and they are intended to provide mutual support for each other. Kusch and McKenna argue that, when these commitments are subjected to detailed scrutiny, they have some quite surprising results. In particular, they argue that Kornblith should be open to a promiscuous pluralism about knowledge that, among other things, threatens his first core commitment and opens the door to non-factive and relativistic conceptions of knowledge. Moreover, it is not clear, Kusch and McKenna argue, that Kornblith can preserve his reliabilist picture of knowledge, and he may be forced to accept a form of relativism that goes beyond what

he himself seems willing to accept. On their view, these commitments lend support to the view that knowledge is not factive, to the view that knowledge is a social kind as well as a natural kind, to a promiscuous sort of pluralism about knowledge, and even to a form of relativism that Kornblith himself is at pains to reject.

In Chapter 3, “On the ‘Arguments-on-Paper’ View of Epistemic Justification,” Juan Comesaña criticizes Kornblith’s take on the “Arguments-on-Paper” view of epistemic justification. Comesaña argues that what bothers Kornblith about the Arguments-on-Paper view is, deep down, its reliance on objective and necessary epistemic principles. According to Comesaña, once we understand why Kornblith is opposed to this view, an illuminating contrast emerges between four different ways of thinking about epistemic justification: the Arguments-On-Paper view, Mad Dog Externalism, Fancy Externalism, and Ostrich Externalism. Comesaña argues that Kornblith’s view is closer to Ostrich Externalism.

In Chapter 4, “Logic: It’s Not Just a Good Idea; It’s the Law,” David Christensen questions Kornblith’s rejection of logic and probability theory as proper standards for good belief. In particular, Christensen targets the following pair of claims defended by Kornblith: first, that logic does not give us “laws of thought” that have implications for a theory of justified belief; and second, that idealization in accounts of epistemic rationality—at least of the sort that figures in Bayesian accounts of rational belief—robs these accounts of their connection to the epistemic phenomena, and thereby robs them of any interest. Christensen responds to the worries Kornblith presents here, arguing that logic-based constraints of the sort that figure in Bayesian accounts of ideal rationality need not have the problematic implications that Kornblith wants to avoid. On Christensen’s view, there is a place for theorizing about rationality using idealizations that aren’t constrained by the cognitive capacities of higher primates.

In Chapter 5, “Epistemological Detente?,” Catherine Z. Elgin examines the standard construal of the debate between evidentialism and reliabilism. These theories are typically construed as rival answers to the same epistemological question: What secures the tenability of a belief? Kornblith, as a process reliabilist, maintains that the answer is provided by cognitive psychology, as it lies in the subliminal psychological processes that underlie belief. Elgin argues that such an answer does not reflect the way that beliefs are answerable to reasons or evidence. But rather than concluding that evidentialism is right and reliabilism is wrong, Elgin suggests that the original question is ambiguous. Under a reliabilist interpretation it means “What are the underlying sources of a belief and how reliable are they?” Under an evidentialist interpretation it means “What intersubjectively accessible and acceptable reasons are there to support a belief and how good are they?” According to Elgin, both are legitimate questions, and answers to both contribute to epistemology.

## 4.2 Inference and Reflection

In Chapter 6, “Reflection, Confabulation, and Reasoning,” Jennifer Nagel examines the epistemology of reason-giving in the context of recent breakthroughs in the development and deployment of large language models. Humans have distinctive powers of reflection:

no other animal seems to have anything like our capacity for self-examination. Kornblith argues against philosophical views which hold that there is something “magical” about second-order mental states, situating them in the same causal net as ordinary first-order mental states. As we’ve seen, sometimes he goes further, suggesting that there is something deeply misleading about reflection’s first-person view of cognition. Nagel takes a closer look at the conditions under which reflection is misleading. Even if humans are the only reflective animals, we are no longer the only creatures that can be questioned about their claims. Chatbots can also be asked to explain themselves, and Nagel argues that by studying the conditions under which they produce clarification as opposed to confabulation, we can develop a non-skeptical account of the epistemic value of reflection. In particular, Nagel argues that this capacity naturally supports the distribution of complex problem solving among multiple agents.

In Chapter 7, “Reflection without Reification,” Berislav Marušić challenges Kornblith’s views on reflection, arguing that the problems Kornblith identifies with reflection point, instead, to problems with his conception of reflection itself. It is common to think of reflection as giving rise to second-order belief. Yet then, it is open to Kornblith, as we’ve seen, to ask: Why wouldn’t the problems that arise at the first order not arise at the second order as well? And why should we think of reflection as reliable—especially since empirical research shows it not to be? Marušić argues that Kornblith’s conception is distorting because it reifies the mind: it conceptualizes our believing something as being the kind of thing that persists as it is, independently of our attitudes. Yet, according to Marušić, our beliefs do not persist as they are, independently of our attitudes, precisely because they form part of our attitudes. Instead, Marušić conjectures, reflection brings into view the way in which our beliefs and other attitudes are essentially social.

In Chapter 8, “Easy Knowledge, Junk Knowledge; Knowledge Nonetheless?,” Katia Vavova revisits the now-classic bootstrapping objection to reliabilism and Kornblith’s well-known response to it. According to this objection, reliabilism licenses objectionably circular reasoning; in particular, it entails that believers are justified in concluding that a process is reliable simply by doing nothing else aside from (iteratively) exercising that very process and reflecting on its products. Kornblith replies, in true reliabilist fashion, by denying that reliabilism *generally* licenses anything of the sort. In cases where agents do this *reliably*, then reliabilism would entail that their beliefs are justified; in other cases, reliabilism would instead entail the opposite. Against Kornblith, Vavova maintains that even reliable bootstrapping is objectionable, exploring the prospects of a response on behalf of reliabilism that blocks the bootstrapping process by arguing that knowledge doesn’t always suffice for reasoning.

In Chapter 9, “Reasons and their Place in Nature,” Tricia Magalotti questions Kornblith’s arguments against the role of reasons in a naturalistic epistemology. Kornblith has argued that epistemologists, and especially reliabilists about epistemic justification, would be better off abandoning the concept of reasons in their epistemic theorizing. The thrust of his argument is that it is plausible that belief explanations that invoke the concept of epistemic reasons are likely to either conflict with or be made obsolete by more sophisticated explanations from cognitive science. Magalotti resists this conclusion, arguing that there is no reason to think that the concept of epistemic reasons is incompatible with evidence from cognitive science and that epistemic reasons play a crucial role in

explaining various doxastic phenomena. Magolotti argues that her conclusions are true even if reliabilism is true.

In Chapter 10, “Inference, Taking, and Reason-Responsiveness,” Anna-Sara Malmgren takes up the question: What is it to make an inference? Malmgren seeks to make progress towards answering this question by looking at some illuminating points of disagreement between Kornblith and other participants in the recent debate on the nature of inference. Special attention is paid to the motivation for, and viability of, higher-order requirements: requirements to the effect that the inferring agent must somehow endorse their inference—e.g., know or judge or “take” the inference to be good, or rational, or warrant preserving. Kornblith has argued at length against accounts of inference that incorporate such requirements, defending what Malmgren considers an interesting—but somewhat programmatic—rival account. Some of Kornblith’s criticism is compelling, Malmgren admits, but she worries that the rival view he defends lacks the resources to address a challenge at the heart of that criticism: namely, how we ought to distinguish inference from certain deviant causal-psychological explanations.

### 4.3 Naturalism, Normativity, and Methodology

In Chapter 11, “Naturalized Epistemic Oughts,” Mona Simion argues against Kornblith’s naturalistic account of normativity. For a long time, Simion begins, epistemology has been negative epistemology: we cared about identifying restrictions on our permissions to believe, judge, use a proposition as a premise in reasoning, and so on. Kornblith is a pioneer of positive epistemology: he was among the first and most prominent champions of epistemic obligations and their impact on epistemic justification. As we’ve seen, for Kornblith, epistemic oughts are naturalistic in that they are desire-generated norms: they are hypothetical imperatives sourced in a desire for truth that we all ought to have, in virtue of its conduciveness to the satisfaction of any desire. Simion does two main things in her chapter. First, she argues that Kornblith’s desire-generated epistemic oughts don’t deliver the naturalistic positive epistemology that he strives for. Second, she suggests that a better way to be a naturalist about epistemic oughts is to be a functionalist about epistemic oughts.

In Chapter 12, “Kornblith on Epistemic Normativity,” Matthew McGrath argues that Kornblith’s classic (1993b) paper on epistemic normativity does not succeed in answering the questions it raises. Kornblith’s account is as simple as it is bold, McGrath claims: the source of epistemic normativity is desire, not a desire for true belief, or knowledge, but any set of desires. As we’ve seen, Kornblith’s view is that no matter what desires you have, so long as you are a being of a kind that employs beliefs in instrumental reasoning and/or cost-benefit analysis, certain sorts of truth-centered epistemic norms will have normative force for you. McGrath distinguishes two questions about epistemic normativity, both seemingly under discussion in Kornblith’s paper, but which McGrath thinks he does not clearly distinguish: (1) why should we care about having beliefs that satisfy epistemic norms? (2) how do epistemic considerations have normative force with respect to particular propositions? McGrath argues that Kornblith’s proposal goes some distance toward answering the first question but is less helpful in answering the second.

In Chapter 13, “Armchair Philosophy and Naturalized Epistemology,” Ernest Sosa defends an account of armchair philosophy in light of Kornblith’s opposition to it. Without dismissing the importance of scientific input, Sosa defends armchair thought and sustained dialectic. But Sosa’s defense is combined with substantial irenic agreement on the project of naturalized epistemology.

In Chapter 14, “Naturalistic Function–First Epistemology,” Georgi Gardiner aims to build a bridge between Edward Craig’s function–first methodology and Kornblith’s naturalism. Craig’s function–first methodology claims that we can illuminate the concept of knowledge by asking what functions the concept evolved to fulfill. To do this, Craig imagines a fictional state of nature in which humans lacked this concept. Craig claims that what those hypothetical humans cannot do reveals the function, and thereby the nature, of knowledge. Kornblith rejects every part of Craig’s methodology, insisting as we’ve seen that we should study knowledge—not its concept—through the scientific study of animal cognition. Gardiner aims to broker peace between these two apparently rival approaches in a defense of methodological pluralism.

In Chapter 15, “How Much Does Knowledge Matter?: Reality, Representation, and the Aims of Epistemology,” Tristram McPherson and David Plunkett engage with Kornblith’s methodology as an “object-level” epistemologist. “Object-level” epistemologists aim ultimately to understand epistemic properties, relations, and facts themselves, rather than our thought and talk about them. For example, as we’ve seen, Kornblith (2002) aims to investigate the nature of knowledge itself rather than the concept of knowledge. A natural motivation for object-level epistemology is the idea that epistemologists should investigate those epistemic entities that are explanatorily and normatively important. McPherson and Plunkett argue that, even if Kornblith has offered a correct account of the knowledge relation, he could have failed to locate and study the most explanatorily and normatively important epistemic entities. On this basis, McPherson and Plunkett argue that thoroughgoing object-level epistemologist (like Kornblith) should not only question the significance of our epistemic “intuitions,” but also the idea that folk words like “knowledge” pick out the apt targets for object-level epistemic inquiry.

In Chapter 16, Kornblith briefly replies to the main themes in these chapters.

## Note

- 1 We are grateful to Hilary Kornblith for comments on this chapter, and for comments on pretty much *everything* that we have written since our time in graduate school.

## References

- Armstrong, D. (1973) *Belief, Truth, and Knowledge*. Cambridge University Press, Cambridge.
- Boyd, R. (1988) How to be a moral realist. In: Sayre-McCord, G. (ed.) *Essays on moral realism*. Cornell University Press, Ithaca, NY, pp.181–228.
- Dretske, F. (1970) Epistemic operators. *The Journal of Philosophy* 67 (24), pp.1007–23.

- Frankfurt, H. (1988) *The Importance of What We Care About*. Cambridge University Press, Cambridge.
- Goldman, A. (1967) A causal theory of knowing. *The Journal of Philosophy* **64** (12), pp.357–72.
- Goldman, A. (1979) What is justified belief? In: Pappas, G. (ed.) *Justification and Knowledge: New Studies in Epistemology*. D. Reidel, Dordrecht, pp.1–25.
- Goldman, A. (1998) Philosophical theory and intuitional evidence. In: Raymond DePaul, M., and Ramsey, W. M. (eds) *Rethinking Intuition: The Psychology of Intuition and its Role in Philosophical Inquiry*. Rowman & Littlefield Publishers, Lanham, MD, pp.179–200.
- Goldman, A. (2010) Epistemic relativism and reasonable disagreement. In: Feldman, R., and Warfield, T. A. (eds) *Disagreement*. Oxford University Press, Oxford, pp.187–215.
- Kim, J. (1988) What is naturalized epistemology? *Philosophical Perspectives* **2**, pp.381–405.
- Kornblith, H. (1980) Beyond foundationalism and the coherence theory. *The Journal of Philosophy* **77** (10), pp.597–612.
- Kornblith, H. (1982) The psychological turn. *Australasian Journal of Philosophy* **60**, pp.238–53.
- Kornblith, H. (1983) Justified belief and epistemically responsible action. *Philosophical Review* **92** (1), pp.33–48.
- Kornblith, H. (1985) What is naturalistic epistemology? In: Kornblith, H. (ed.) *Naturalizing Epistemology*. MIT Press, Cambridge, MA, pp.1–13.
- Kornblith, H. (1986) Naturalizing rationality. In: Hare, P., and Garver, N. (eds) *Naturalism and Rationality*. Prometheus Books, Buffalo, NY, pp.115–33.
- Kornblith, H. (1988) How internal can you get? *Synthese* **74**, pp.313–27.
- Kornblith, H. (1993a) *Inductive Inference and Its Natural Ground: An Essay in Naturalistic Epistemology*. MIT Press, Cambridge, MA.
- Kornblith, H. (1993b) Epistemic normativity. *Synthese* **94**, pp.357–76.
- Kornblith, H. (1994) Naturalism: both metaphysical and epistemological. *Midwest Studies in Philosophy* **19**, pp.39–52.
- Kornblith, H. (1999a) Knowledge in humans and other animals. *Philosophical Perspectives* **13**, pp.327–46.
- Kornblith, H. (1999b) The epistemology of science and the epistemology of everyday life. *Facta Philosophica* **1**, pp.21–37.
- Kornblith, H. (1999c) Distrusting reason. *Midwest Studies in Philosophy* **23**, pp.181–96.
- Kornblith, H. (2001) Internalism and externalism: a brief historical introduction. In: Kornblith, H. (ed.) *Epistemology: Internalism and Externalism*. Blackwell, Oxford, pp.1–9.
- Kornblith, H. (2002) *Knowledge and its Place in Nature*. Oxford University Press, Oxford.
- Kornblith, H. (2003) Can internalism be saved? *Metaphilosophy* **34**, pp.621–69.
- Kornblith, H. (2004) Sosa on human and animal knowledge. In: Greco, J. (ed.) *Sosa and His Critics*. Blackwell, Oxford, pp.126–34.
- Kornblith, H. (2012) *On Reflection*. Oxford University Press, Oxford.
- Kornblith, H. (2015) The role of reasons in epistemology. *Episteme* **12**, pp.225–39.
- Kornblith, H. (2017) Doxastic justification is fundamental. *Philosophical Topics* **45**, pp.63–80.
- Kornblith, H. (2019) *Second Thoughts and the Epistemological Enterprise*. Cambridge University Press, Cambridge.
- Kornblith, H. (2021) *Scientific Epistemology: An Introduction*. Oxford University Press, Oxford.
- Kornblith, H. (2022) What does logic have to do with justified belief? Why doxastic justification is fundamental. In: Silva, P., Jr., and Oliveira, L. R. G. (eds) *Propositional and Doxastic Justification: New Essays on Their Nature and Significance*. Routledge, Abingdon, pp.40–58.
- Korsgaard, C. (1996) *The Sources of Normativity*. Cambridge University Press, Cambridge, MA.
- Mercier, H., and Sperber, D. (2017) *The Enigma of Reason*. Harvard University Press, Cambridge, MA.
- Quine, W. V. O. (1969) Epistemology naturalized. In: *Ontological Relativity and Other Essays*. Columbia University Press, New York, pp.69–90.

