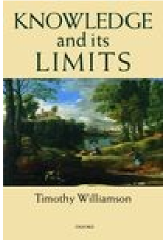

Knowledge and its Limits
Timothy Williamson
https://doi.org/10.1093/019925656X.001.0001
Published: 2002 **Online ISBN:** 9780191598678 **Print ISBN:** 9780199256563

Search in this book

CHAPTER

Introduction

Timothy Williamson

https://doi.org/10.1093/019925656X.003.0001 Pages 1–20
Published: October 2002

Subject: Epistemology, Metaphysics
Collection: Oxford Scholarship Online

> Everyone by nature desires to know.
>
> Aristotle, *Metaphysics* A1 980a21 (modernized)

1 Knowing and Acting

Knowledge and action are the central relations between mind and world. In action, world is adapted to mind. In knowledge, mind is adapted to world. When world is maladapted to mind, there is a residue of desire. When mind is maladapted to world, there is a residue of belief. Desire aspires to action; belief aspires to knowledge. The point of desire is action; the point of belief is knowledge.

Those slogans are not platitudes—unless platitudes can be generally contested. According to many philosophers, desire aspires only to satisfaction, and belief only to truth. Action is a systematic way to satisfied desire, and knowledge to true belief, but desires can also be satisfied and beliefs true by chance. There is satisfied desire without action and true belief without knowledge. Why ask for more? Satisfaction and truth already constitute the required match between mind and world, with the appropriate directions of fit. Of course, we sometimes desire to act; those desires are satisfied only if there is action. We sometimes believe ourselves to know; those beliefs are true only if there is knowledge. But such cases are special; our desires and beliefs frequently concern states of the world of which actions and beliefs are not themselves constituents.

Although desires can be satisfied as well by chance as by action, that is no reason to marginalize the category of action in the understanding of mind. The place of desire in the economy of mental life depends on its potential connection with action. Similarly, although beliefs can be true as well by chance as by knowledge, that is no reason to marginalize the category of knowledge in the understanding of mind. This book develops a conception on which the place of belief in the economy of mental life depends on its potential connection with knowledge.

p. 2 The foregoing vague phrases will later be partially replaced by something ↳ more precise. But that is not the purpose of this introduction, which is painted with a broad brush. Its aim is to give the reader a rough overall picture in which the layout of the main parts is visible. Subsequent chapters fill in details in the parts. Even they will amount to nothing like a proof that the picture is correct. Epistemological theories are not usually susceptible of proof. This book shows how to understand cognitive phenomena on the basis of some simple but generally overlooked ideas. The reader will judge those ideas by their fruit.

2 Unanalysable Knowledge

Contemporary accounts of mind tend to marginalize the category of knowledge, sometimes not mentioning it at all; they certainly make it less central than the category of action. As a reverse counterpart of the output from mind to world in action, they admit the input from world to mind in perception. The latter is a more restricted category than knowledge; it excludes the products of memory and conscious inference. Perception is the reverse counterpart of action if both are single episodes of causal interaction with the environment. But acting, in the sense of intentionally making something the case, includes far more complex and mediated adaptations of world to mind over extended periods. The reverse counterpart of action in that sense is knowledge. It includes far more complex and mediated adaptations of mind to world over extended periods than perception does.

On contemporary accounts of mind, the general category for states with the mind-to-world direction of fit is belief. The belief is true if it fits the world, false otherwise. Although true and false belief are the same mental state in different worlds, the place of belief in the economy of mental life depends on its potential connection with truth. Knowledge is merely a peculiar kind of true belief. Since Gettier showed that even justified true belief is insufficient for knowledge, epistemologists have expended vast efforts attempting to state exactly what kind of true belief knowledge is, but that programme is assigned no significance for the philosophy of mind. On such a view, knowledge is to be explained in terms of belief, and belief is what matters for the understanding of mind. The converse attempt to explain belief in terms of knowledge sounds eccentric and perverse. To summarize this orthodoxy: belief is conceptually prior to knowledge.

The orthodox claim is frequently taken for granted, rarely supported by argument. Why should we suppose
p. 3 that belief is conceptually prior ↳ to knowledge? One argument is that since knowledge entails belief but not vice versa, the entailment should be explained by the assumption that we conceptualize knowledge as the conjunction of belief with whatever must in fact be added to belief to yield knowledge—truth and other more elusive features. The conjuncts are conceptually prior to the conjunction. Given that knowledge entails belief, it is trivial that one knows *p* if and only if (1) one believes *p*; (2) *p* is true; and (3) if one believes *p* and *p* is true, then one knows *p*. But that equivalence is useless for establishing that belief is conceptually prior to knowledge, for it is circular: 'know' occurs in (3). The received idea is that we can conceptualize the factors whose conjunction with belief is necessary and sufficient for knowledge independently of knowledge; we can think of the former without already thinking of the latter, even implicitly. But the argument does not show that such independent conceptualization is possible, for a necessary but insufficient condition need not be a conjunct of a non-circular necessary and sufficient condition. Although being coloured is a necessary but insufficient condition for being red, we cannot state a necessary and sufficient condition for being red by conjoining being coloured with other properties specified without reference to red. Neither the equation 'Red = coloured + X' nor the equation 'Knowledge = true belief + X' need have a non-circular solution.

Thus belief can be a necessary but insufficient condition of knowledge even if we do not implicitly conceptualize knowledge as the conjunction of belief with that which must be added to belief to yield knowledge. Perhaps the inference from knowledge to belief derives from a conceptualization of belief in terms of knowledge rather than from a conceptualization of knowledge in terms of belief. If believing *p* is conceptualized as being in a state sufficiently like knowing *p* 'from the inside' in the relevant respects, then belief is necessary for knowledge, since knowing *p* is sufficiently like itself in every respect, even though knowledge is conceptually prior to belief. Indeed, the inference from knowledge to belief does not even require knowledge and belief to be conceptually ordered. We might master 'know' and 'believe' independently, from examples, and then realize on that basis that believing is necessary but insufficient for knowing, just as we might master the terms 'red' and 'scarlet' independently, from examples, and then realize on that basis that being red is necessary but insufficient for being scarlet. That belief is necessary but insufficient for knowledge does not show that belief is conceptually prior to knowledge. The orthodox claim would require a deeper defence.

Some epistemologists defend the conceptual priority of belief over knowledge by citing their favoured p. 4 analyses of knowledge in terms of ↳ belief. Just what kind of conceptual priority such an analysis might support would depend on its status: for instance, on whether it was analytic or knowable a priori in some sense. But those issues become notional when, as usually happens, a counterexample is found to show that the proposed condition is not even necessary and sufficient for knowledge. Other analyses are circular rather than false; if someone insists that knowledge *is* justified true belief on an understanding of 'justified' strong enough to exclude Gettier cases but weak enough to include everyday empirical knowledge, the problem is likely to be that no standard of justification is supplied independent of knowledge itself. This book makes no attempt to survey even the most salient analyses of knowledge proposed in recent decades and the counterexamples to which they succumb; many other authors have already done that adequately. It will be assumed, not quite uncontroversially, that the upshot of that debate is that no currently available analysis of knowledge in terms of belief is adequate (not all parts of the book depend on that assumption). Consequently, the supposed conceptual priority of knowledge over belief is not to be defended by appeal to a particular analysis of knowledge in terms of belief.

A more cautious argument for the conceptual priority of belief over knowledge is that, even if all currently available analyses of knowledge in terms of belief are circular or fall to counterexamples, some of them are sufficiently good approximations to indicate strongly that a further refinement on similar lines will eventually succeed. But the possibility of approximating one concept with others is not good evidence that the former can be analysed in terms of the latter. For instance, to a very good approximation, *x* is a parent of *y* if and only if *x* is an ancestor of *y* and *x* is not an ancestor of an ancestor of *y*. The only counterexamples are *recherché* cases of incest: if a father incestuously begets a son on his daughter, the father is an ancestor of an ancestor of his son. But no more refined definition of parenthood in terms of ancestry alone avoids the problem. Since the father and the mother of his daughter are symmetrically related to the daughter and son in terms of ancestry but not in terms of parenthood, parenthood cannot be defined in terms of ancestry without extra conceptual resources. Moreover, the approximate definition of parenthood in terms of ancestry plays no significant role in our understanding of 'parent'. We can approximate a circle as closely as we like with sufficiently many sufficiently small triangles; it does not follow that we should think of the circle as made up out of triangles. The possibility of approximating knowledge in terms of belief and other concepts is not good evidence for the conceptual priority of belief over knowledge (section 1.3). Section 1.4 p. 5 shows how one might ↳ characterize knowledge without reference to belief. Section 1.5 briefly discusses how one might characterize belief by reference to knowledge.

A chief aim of this book is to develop a rigorous way of doing epistemology in which knowledge is central, and not subordinate to belief. It enables us to abandon the attempt to state necessary and sufficient conditions for knowledge in terms of belief without abandoning epistemology itself. Indeed, by abandoning that fruitless search we can gain insight into epistemological problems, because we are freed to use the notion of knowledge as an instrument of understanding in ways that its subordination to belief would not permit (see, in particular, Chapter 9).

3 Factive Mental States

The idea that belief is conceptually prior to knowledge has another source: the internalist conception of mind, and world external to mind, as two independent variables. Belief is simply a function of the mind variable. Truth is simply a function of the external world variable, at least when the given proposition is about the external world. For the internalist, knowledge is a function of the two variables, not of either one alone; whether one knows that it is raining does not depend solely on one's mental state, a state which is the same for those who perceive the rain and those who hallucinate it, but it also does not depend solely on the state of the weather, a state which is the same for those who believe the appearances and those who doubt them. The internalist therefore conceives knowledge as a complex hybrid crying out for analysis into its internal and external components, of which belief and truth respectively are the most salient. The analysis is expected on general metaphysical grounds.

Recent developments in the philosophy of mind have called the metaphysics of internalism into question by indicating ways in which the content of a mental state can constitutively depend on the environment. I believe that tigers growl; an exact physical replica of me lacks that belief if his contact has been not with tigers but with schmigers, beasts of a similar appearance belonging to a different species; his belief is that schmigers growl. Some internalists conclude that not even belief as attributed in ordinary language is simply a function of mind, and try in theory to isolate a core of purely mental states. Such attempts have not succeeded. Rather, we may conceive mind and external world as dependent variables, and reject the metaphysics that led us to expect analysis into purely internal and purely external components. On this view, p. 6 ↳ belief as attributed in ordinary language is a genuine mental state constitutively dependent on the external world.

If the content of a mental state can depend on the external world, so can the attitude to that content. Knowledge is one such attitude. One's knowledge that it is raining depends on the weather; it does not follow that knowing that it is raining is not a mental state. The natural assumption is that sentences of the form ‘S knows *p*’ attribute mental states just as sentences of the forms ‘S believes *p*’ and ‘S desires *p*’ do. Chapters 1 and 2 defend such an externalist conception of knowing as a state of mind. In particular, section 2.3 refutes internalist attempts to isolate a non-factive state as the purely mental component of knowing.

What is at stake is much more than whether we apply the word ‘mental’ to knowing. If we could isolate a core of states which constituted ‘pure mind’ by being mental in some more thoroughgoing way than knowing is, then the term ‘mental’ might be extended to knowing as a mere courtesy title. On the conception defended here, there is no such core of mental states exclusive of knowing. If we want to illustrate the nature of mentality, knowing is as good an example as believing. The philosophy of mind cannot afford to neglect knowing, for that state is part of its core subject matter. For similar reasons, other truth-entailing attitudes such as perceiving and remembering that something is the case may also be classified as mental states. Knowing can be understood as the most general of such truth-entailing mental states (section 1.4).

Sceptics and their fellow-travellers characteristically suppose that the truth-values of one's beliefs can vary independently of those beliefs and of all one's other mental states: one's total mental state is exactly the same in a sufficiently radical sceptical scenario as it is in a common-sense scenario, yet most of one's beliefs about the external world are true in the common-sense scenario and false in the sceptical scenario. But if knowing is itself a mental state, that supposition is tantamount to the sceptical conclusion that in the common-sense scenario one's beliefs do not constitute knowledge, even though they are true. For, since false beliefs never constitute knowledge, one certainly does not know in the sceptical scenario; the supposition that one is in exactly the same mental state in the two scenarios therefore implies that one does not know in the common-sense scenario either, given that knowing makes a difference to one's total mental state (section 1.2). The anti-sceptic should not accept the supposition. Any mental life in the sceptical scenario is of a radically impoverished kind. Of course it does not *feel* impoverished ‘from the inside’, but that failure of self-knowledge is part of the impoverishment.

p. 7 If action is the reverse counterpart of knowledge, and knowing is a ↳ mental state, should we expect acting to be a mental state too? If so, we might compare the sceptic's denial that we know about the external world to the denial that we act on the external world (perhaps made by those who believe that free will is both an illusion and a precondition of genuine action). But the analogy between knowledge and action is not perfect. Acting is by definition no state of any kind; it is dynamic, not static. Moreover, while knowing that the door is shut may be a mental state, shutting the door is surely not a mental action. Only actions such as inferring are naturally classified as mental. Similar asymmetries arise if we pursue the more restricted analogy between action and perception, for instance, between breaking the window and seeing that the window is broken. One starts seeing that the window is broken after the light rays reach one's retina; we do not make the apparently symmetric claim that one finishes breaking the window before the stone leaves one's hand. Why do we conceptualize the input and output sides so differently? The answer may lie in our tendency to individuate by origins. Effects depend on their causes in a way in which causes do not depend on their effects. Thus early stages in the process leading from a cause in the environment to a perceptual experience typically do not depend on the perceiver's involvement, whereas even late stages in the process leading from an intention to an effect on the environment do depend on the agent's involvement. Thus we naturally group both early and late stages of the output process into something attributable to the agent, while grouping only late stages of the input process into something attributable to the perceiver (this notion of grouping is intended to be neutral between different theories of the ontology of action). We treat early stages of the input process merely as preconditions for what we attribute to the perceiver. We extend this scheme to cases of knowledge and action with a more complex causal structure. Since late stages of the output process which occur without need of continued mental involvement are grouped into the action, we are reluctant to conceive it as mental. By contrast, since early stages of the input process are not grouped into the perception or knowledge, there is no corresponding block to conceiving it as mental.

Knowledge and action are related in another way. We expect genuine mental states to occur significantly in causal explanations of action, for otherwise postulating them looks redundant. Thus if knowing is a genuine mental state, it should occur significantly in such explanations. But many philosophers assume that attributions of knowledge in causal explanations of action can be replaced without explanatory loss by the corresponding attributions of belief. Section 2.4 and, in more depth, Chapter 3 undermine that assumption.
p. 8 Action typically involves complex ↳ interaction with the environment; one needs continual feedback to bring it to a successful conclusion. For example, writing a book involves reading during the process. Attributions of knowledge often explain the success of these interactions better than do the corresponding attributions of belief, even of true belief. One's belief in a proposition p is more robust to evidence if one knows p than if one merely believes p truly; one is less likely to lose belief in p in the course of interacting with the environment by discovering new evidence which lowers the probability of p. Thus one is more likely to complete an extended action that depends on a continuing belief in p if at the start one knows p rather than merely believes p truly. One is better placed to write a mathematical paper if one knows the truth of Goldbach's Conjecture than if one merely believes the conjecture and it is true. The point is not answered by an analysis of actions into series of basic actions, for the causal explanations even of basic actions often cite mental states at a temporal distance. One deliberates, forms an intention and then executes it later or abandons it in the light of new developments. The gap between deliberating and completing the action allows differences between knowledge and mere true belief in the basis of the deliberation to manifest

themselves in action. If the causal explanation of the action cited only mental states immediately preceding the action, it would omit those on which the deliberation was based, and thereby miss the rationality of the action. These considerations can be generalized from attributions of knowledge to attributions of mental content involving reference to the environment; they all play distinctive roles in the causal explanation of temporally distant actions. Chapter 3 uses such considerations to argue that an externalist mental state normally cannot be decomposed as the conjunction of purely internal and purely external components.

4 Knowledge as the Justification of Belief and Assertion

The idea that belief is conceptually prior to knowledge easily leads to the idea that evidence and justification are conceptually prior to knowledge too. Although that is most vivid in the traditional definition of knowledge as justified true belief, Gettier's counterexamples to that definition did not remove the idea that the concept of justification or evidence would occur with the concept of belief in a more complex analysis of the concept of knowledge. Consequently, the concept of knowledge was assumed to be unavailable for use in
p. 9 an elucidation of ↳ the concept of justification or evidence, on pain of circularity. Once we cease to assume that belief is conceptually prior to knowledge, we can experiment with using the concept of knowledge to elucidate the concepts of justification and evidence.

Chapter 9 makes the experiment. It argues that one's total evidence is simply one's total knowledge. Thus a hypothesis is inconsistent with the evidence if and only if it is inconsistent with known truths; it is a good explanation of the evidence if and only if it is a good explanation of known truths. One's evidence justifies belief in the hypothesis if and only if one's knowledge justifies that belief. Knowledge figures in the account primarily as what justifies, not as what gets justified. Knowledge can justify a belief which is not itself knowledge, for the justification relation is not deductive. For example, I may be justified in believing that someone is a murderer by knowing that he emerged stealthily with a bloody knife from the room in which the body was subsequently discovered, even if he is in fact innocent and I therefore do not know that he is a murderer.

The equation of one's evidence with one's knowledge does not imply any particular theory of how a given body of propositional evidence justifies a given belief. Rather, it connects absolute and relative justification. A belief is justified relative to some other beliefs from which it has been derived in some appropriate way (perhaps by deduction), but it is not justified absolutely unless those other beliefs are justified absolutely. Where does the regress end? On the assumption that it ends at evidence, the equation of evidence with knowledge implies that one's belief is justified absolutely if and only if it is justified relative to one's knowledge. The regress of justification ends at knowledge.

The account might be thought to make all knowledge self-justifying in an absurdly trivial way: one's knowledge is justified absolutely if and only if it is justified relative to itself. This objection would be fair if the point of justification were to serve at its best as a condition for knowledge. But on the present account that is not the point of justification. Rather, justification is primarily a status which knowledge can confer on beliefs that look good in its light without themselves amounting to knowledge. Knowledge itself enjoys the status of justification only as a limiting case, just as, trivially, every shade of green counts as similar to a shade of green.

The objector might still point out that non-trivial questions appear to arise about the justification and evidence for much of our knowledge, especially that which is mediated by theory. We miss the specificity of these questions if we treat them merely as general questions about how we know. Nevertheless, they can be
p. 10 understood as non-trivial on the ↳ present approach. For even if one knows *p*, one can call that knowledge into question, provisionally treat *p* as though it did not belong to the body of one's knowledge, and then assess *p* relative to the rest of one's knowledge—one's independent evidence. Non-trivial issues of evidence

and justification will then arise for *p*. This procedure is a good test of some kinds of supposed knowledge, especially those mediated by theory. In such cases, given the purported manner of knowing *p*, one knows *p* only if the rest of one's knowledge justifies *p*. But the test is not universal; it yields poor results if too much of one's knowledge is simultaneously called into question, for then one may easily have that knowledge even if removing it from the body of one's knowledge leaves too little to justify it. Some sceptics go wrong by applying the test in such cases.

The consideration of reduced bodies of evidence serves a different purpose. It enables us to isolate the contribution of specific pieces of evidence to the justification of specific hypotheses by comparing the status of those hypotheses relative to the total body of evidence with their status relative to the result of removing the piece of evidence in question from the total body of evidence This point is related to a form of the so-called problem of old evidence.

Chapter 10 develops these ideas in a more technical direction, by combining them with a theory of evidential probability in a modified objective Bayesian framework (some readers may prefer to skip this chapter). The evidential probability of a hypothesis for one is its probability conditional on one's total evidence; given the equation of one's evidence with one's knowledge, that is its probability conditional on one's total knowledge. Thus knowledge automatically receives evidential probability 1. Even so, knowledge is not treated as indefeasible evidence, for one can lose as well as gain knowledge. Thus the future evidential probability for me of my present knowledge may be less than 1. Together, Chapters 9 and 10 illustrate a way of doing epistemology on which knowledge is taken as the starting point, the unexplained explainer, yet some degree of rigour is maintained.

Chapter 11 extends the approach to the philosophy of language, with an account of the speech act of assertion. In a natural way we can regard assertion as the verbal counterpart of judgement and judgement as the occurrent form of belief. If one assumes that belief is conceptually prior to knowledge, one will therefore expect an account of assertion not to use the concept of knowledge. It might instead use the concepts of truth and justified belief, perhaps independently of each other. But if belief is not conceptually prior to knowledge, and knowledge is what justifies belief, then knowledge should play a key role in an account of ↳ assertion. On the proposal made in Chapter 11, the fundamental rule of assertion is that one should assert *p* only if one knows *p*. Although that knowledge rule might appear to be derivable from the truth rule that one should assert *p* only if *p* is true, by the consideration that in asserting *p* one does not know that one is conforming to the truth rule unless one is in fact conforming to the knowledge rule, the attempted derivation fails because it predicts the wrong epistemology for some examples. Even more incorrect predictions issue from an account of assertion based on the justification rule that one should assert *p* only if one is justified in believing *p*. The concept of knowledge is needed to capture our practice of assertion.

Given the combined conclusions of Chapters 9 and 11, the propositions which one is permitted to assert outright are exactly those which constitute one's evidence. More speculatively, we may project the account of assertion back onto its mental counterpart, judgement (or belief). What results is the rule that one should judge (or believe *p*) only if one knows *p*. That would make some sense of the claim that belief aims at knowledge. It also harmonizes with the account of evidence: to believe *p* without knowing *p* is to exceed one's evidence. Although we may have qualms about applying the notion of a rule to mental acts in addition to speech acts, the idea that belief is governed by a norm of knowledge is at least as intelligible as the idea that it is governed by a norm of truth.

5 The Myth of Epistemic Transparency

An account has been sketched of knowledge as a mental state which constitutes the evidential standard for assertion and belief. The several components of the account face a common epistemological objection. It starts from the observation that one is not always in a position to know whether one knows something. If one knows *p*, it does not follow that one is in a position to know that one knows *p* (section 5.1); if one does not know *p*, it does not follow that one is in a position to know that one does not know *p* (section 8.2). In both cases, the conclusion fails to follow even if we add the extra premise that one is wondering whether one knows *p*; the problem is not confined to subjects who are unconscious, lack the concept of knowledge, or the like. In the simplest examples, one does not know and is not in a position to know that one does not know *p*. Sometimes *p* is false, so one does not know *p*, even though systematically misleading appearances place one in a state which feels just like ↳ knowing *p* 'from the inside', so one is not in a position to know that one does not know *p*. One falsely but justifiably believes oneself to know *p*. Examples in which one knows without being in a position to know that one knows will be discussed later. Why are these limitations on one's ability to know whether one knows supposed to threaten the foregoing account of knowledge? p. 12

Consider first the thesis that knowing is a mental state. We are often said to have special access to our own mental states, so that we can know without observation what mental states we are in. If S is a mental state only if one is always in a position to know whether one is in S (at least when one is in a position to wonder whether one is in S), then knowing is not a mental state.

A similar objection applies to the equation of evidence with knowledge. Rationality requires one to conform one's beliefs to one's evidence. Rationality cannot require one to do the impossible. But how can one conform one's beliefs to one's evidence unless one is in a position to know what it is? If one is always in a position to know what one's evidence is, then one's evidence is not one's knowledge.

Since one is not always in a position to know whether one knows *p*, one is not always in a position to know whether in asserting *p* one is complying with the rule 'Assert only what you know'. In particular, one may fail to meet the knowledge condition even though it feels 'from the inside' just as though one met the condition. A violation of the knowledge rule in such circumstances may look blameless. How can a speech act be governed by a rule that one can blamelessly violate? If one is always in a position to know whether one's assertions comply with the rule for assertion, then the rule is not 'Assert only what you know'.

If the first objection is sound, then every mental state has the property that one is in a position to know whether one is in it (the qualification 'whenever one is in a position to wonder whether one is in it' will henceforth often be left tacit). If the second and third objections are also sound, then mental states are qualified by their possession of that property to be both evidence and the standard for assertion, at least in respect of accessibility. Indeed, it is unclear how anything other than a mental state could be accessible in the required way (we may exclude trivial states which one is always or never in). For suppose that one is always in a position to know whether a condition C obtains. Consider an ordinary case α in which one might be and a sceptical counterpart α^* of α. In α^*, one is not in α but appears to oneself to be in α and for all one knows one is in α. If C obtains in α, then for all one knows in α^* one is in a situation in which C obtains; thus if C does not obtain in α^*, one is not in a position to know in α^* whether C obtains, contrary to ↳ hypothesis; therefore C obtains in α^*. By a parallel argument, if C does not obtain in α then C does not obtain in α^*. Thus C obtains in α if and only if C obtains in α^*. C is insensitive to the difference between ordinary cases and their sceptical counterparts. Mental states are the only obvious candidates for exhibiting such insensitivity. p. 13

We have also uncovered another temptation to scepticism, for if we combine the argument of the previous paragraph with the principle that one is always in a position to know what one's evidence is, the upshot is that one has exactly the same evidence in an ordinary case and its sceptical counterpart. How, then, can one know which case one is in?

The three objections assume that some non-trivial states meet the accessibility requirement; one is always in a position to know whether one is in them. Chapter 4 challenges that assumption. It provides a general form of argument, applicable to almost any condition, to undermine the claim that one is always in a position to know whether it obtains. More specifically, with rather trivial exceptions it undermines the claim that the condition is *luminous*, in the sense that whenever it obtains (and one is in a position to wonder whether it does), one is in a position to know that it obtains. The main idea behind the argument against luminosity is that our powers of discrimination are limited. If we are in a case α, and a case α′ is close enough to α, then for all we know we are in α′. Thus what we are in a position to know in α is still true in α′. Consequently, a luminous condition obtains in α only if it also obtains in α′, for it obtains in α only if we are in a position to know that it obtains in α. In other words, a luminous condition obtains in any case close enough to cases in which it obtains. What counts as close enough depends on our powers of discrimination. Since they are finite, a luminous condition spreads uncontrollably through conceptual space, overflowing all boundaries. It obtains everywhere or nowhere, at least where we are in a position to wonder whether it obtains. For almost any condition of interest, the cases in which it obtains are linked by a series of imperceptible gradations to cases in which it does not obtain, where at every step we are in a position to wonder whether it obtains. The condition is therefore not luminous. The full version of the argument cashes out those spatial metaphors in epistemic terms, to ensure that they do not import unwarranted presuppositions. In particular, the full version is formulated in a way which does not presuppose perfect sharpness in the boundary between the cases in which the condition obtains and the cases in which it does not. The upshot of the argument is that the gap between what is true and what we are in a position to know is not a special feature restricted to some problematic areas of discourse; it is normal throughout discourse.

Some may doubt the applicability of the argument against luminosity to mental states, on the grounds that it relies on a model of discrimination between independently constituted items, whereas one's mental states and one's judgements about them are held to be constitutively interdependent. But no such interdependence makes one's judgements about one's present mental states infallible. For instance, if my guru tells me that I shall feel intense pain at midnight and I am sufficiently gullible, I may judge at midnight that I am feeling intense pain; it does not follow that I am feeling intense pain. Of course, I may be in a position to know that I do not feel intense pain; my failure may be to actualize that potential. But the example still shows a gap between judgement and truth, even if smaller than elsewhere, which the argument can use as the thin end of a wedge against luminosity. The full version proceeds by analysis of the gradually varying degrees of confidence with which one judges, in a way applicable to judgements about mental states.

For virtually no mental state S is the condition that one is in S luminous. The condition that one is *not* in S is equally non-luminous. For example, one can love someone without being in a position to know that one loves them, and one can fail to love someone without being in a position to know that one fails to love them. One can want something without being in a position to know that one wants it, and one can fail to want something without being in a position to know that one fails to want it. Granted that knowing is a mental state, one should therefore not be surprised that one can fail to know something without being in a position to know that one fails to know it. Indeed, one can argue independently against luminosity for many mental states. They involve patterns of causal connections; sometimes one makes a judgement about one's present state which one is subsequently forced to retract, because one's intervening behaviour was in tension with the self-attributed pattern. One's judgements may be subject to systematic distortion. One's self-attributions of mental states are sometimes too unreliable to constitute knowledge. Mental states incompatible with one's self-image may be concealed from one. The difference between remembering an incident in one's early childhood and imagining it is a difference in mental state, but it is also one about which it is easy to be wrong.

None of this is to deny that in favourable cases one can know without observation whether one is in a given mental state. But knowledge meets that condition. You may know without observation whether you know that it rained two days ago, just as you may know without observation whether you believe that it rained two days ago. If you know that it rained two days ago, that knowledge (and belief) may result from p. 15 ↳ past observations, but no *further* observations were needed to know that you know (and believe). Of course, subsequent observations indicating that it did not rain two days ago undermine the self-attribution of past knowledge that it rained two days ago without undermining the self-attribution of past belief that it rained two days ago. But if a judgement can be undermined by reasons of some kind, it does not follow that it was made on the basis of other reasons of the same kind. I can know without further observation that I know *p* even though observation can falsify a claim to know *p*.

Our extensive but not unlimited ability to know without further observation whether we know something is what enables us to use knowledge as evidence. It constitutes an extensive but not unlimited ability to know without further acquisition of evidence whether something is part of our present evidence. To complain that we are not always in a position to know whether we know something is to bankrupt the notion of evidence, for only luminous conditions meet that more stringent constraint, and luminous conditions are trivial. Although the constraint might drive us to suppose that one's evidence consists of appearances to oneself, the discrimination argument shows that not even the condition that things appear to one in a given way is luminous. For example, one may appear to oneself to be seeing a red patch even though one is not in a position to know that one appears to oneself to be seeing a red patch. Once the standard for the epistemic accessibility of evidence is set at an attainable level, knowledge meets the standard.

Chapter 8 traces the way in which excessive demands on the accessibility of evidence invite scepticism by diminishing evidence to an imaginary phenomenal substratum. If we presume to know too much about our evidence, we find ourselves knowing too little about the external world. The best argument for supposing that we have no more evidence in ordinary cases than in their sceptical counterparts trades on the false premise that the condition for being evidence is luminous. Since sceptics have not refuted the equation of evidence with knowledge, they are not entitled to assume that we have no more evidence in ordinary cases than in their sceptical counterparts, for on the view against which they are attempting to argue we do have more knowledge in ordinary cases than in their sceptical counterparts.

Since rationality requires one to conform one's beliefs to one's evidence, and one is not always in a position to know what one's evidence is, we need a conception of rationality on which we are not always in a position to know what it demands. Indeed, the anti-luminosity argument takes us more directly to the conclusion that one may be rationally p. 16 ↳ required to do something even though one is not in a position to know that one is rationally required to do it. If we imagine that some candidate criterion of rationality is perfectly accessible, then we are always likely to prefer that criterion; but once we recognize that perfect accessibility is quite generally an unattainable ideal, we can learn to live with an imperfectly accessible criterion. We have nothing else to live with. Provided that one's evidence is more accessible than the truth-values of the hypotheses under investigation, the former can still serve as a useful guide to the latter. Real life is messy. Section 10.6 explores some unexpected implications of imperfect accessibility for decision theory. Imperfect accessibility has ethical implications too; we are not always in a position to know our duty.

For the same reason, one should expect not always to be in a position to know whether in asserting *p* one conforms to the rule of assertion. That the account of assertion based on the rule ‘Assert only what you know’ has that consequence is therefore no objection. An account based on the rule ‘Assert only what you rationally believe’ would have the same consequence. Our practice of assertion is workable because we often enough know whether we know something.

The imperfect accessibility of rationality casts light on the external individuation of mental content, mentioned earlier. For rationality has some relation to deductive logic, although the relation is not easy to spell out, and the external individuation of content makes the deductive validity of inferences imperfectly accessible. Whether the inference from ‘It is hot here’ and ‘It is wet here’ to ‘It is hot and wet somewhere’ is valid in a given context depends on whether the two occurrences of ‘here’ have the same content in that context. Someone who accepts the premises and rejects the conclusion avoids inconsistency only if the argument is invalid. If the content of ‘here’ is determined at least in part by the environment, one may not be in a position to know whether the inference is valid. Similarly, whether the inference from ‘Everything changes’ to ‘Bourbaki changes’ is valid depends on whether ‘Bourbaki’ has a content. If the content of ‘Bourbaki’ (if any) is determined at least in part by the environment, then one may not be in a position to know whether the inference is valid. These examples are only indicative; the argument from external individuation to imperfect accessibility is much less straightforward if the externally individuated content of a term is not identified with its referent. Nevertheless, if we accept on independent grounds that one is not always in a position to know what rationality demands, we should not object to an account that individuates content externally just on the grounds that it makes validity imperfectly accessible.

p. 17 Chapter 5 articulates the constraints on knowledge implicit in the argument against luminosity. Where one has only a limited capacity to discriminate between cases in which *p* is true and cases in which *p* is false, knowledge requires a margin for error: cases in which one is in a position to know *p* must not be too close to cases in which *p* is false, otherwise one's belief in *p* in the former cases would lack a sufficiently reliable basis to constitute knowledge. The kind and degree of closeness in question depend on the specific limitations of one's powers of discrimination in that context. Thus the area of conceptual space in which one is in a position to know *p* is separated from the surrounding area in which *p* is false by a border zone in which *p* is true but one is not in a position to know *p*. The implications of the model are explored for iterated knowledge. In particular, one has only a limited capacity to discriminate between cases in which one knows *p* and cases in which one does not know *p*, so one can know *p* without being in a position to know that one knows *p*. Further iterations of knowledge are even harder to achieve. Naturally, one has an even more limited capacity to discriminate between cases in which others know *p* and cases in which they do not know *p*, so it is even harder to achieve iterations of shared knowledge (‘We all know that we all know that we all know . . . *p*’), and a fortiori to achieve the infinitely many levels of iteration required for common knowledge. Chapter 6 uses this difficulty to account for the Paradox of the Unexpected Examination and some paradoxical arguments in game theory which assume that it is common knowledge amongst the players that they are all rational.

Because we need margins for error, it is implicit in our practice of assertion that truth outruns warranted assertion. We are warranted in asserting *p* only if we know *p*; we know *p* only if *p* is true in nearby cases. To interpret our assertions as warranted, we must interpret their content as true in some cases in which we are not warranted in asserting it. Our ignorance is a precondition of our knowledge. Contrary to anti-realist theories, the gap between assertibility and truth is built into even the simplest kinds of assertion.

Chapter 7 exploits margins for error in another direction. Sometimes one knows *p* by believing *p* and leaving a large margin for error even though, if *p* were false, one would still believe *p*. For one's judgement concerning *p* may be almost completely accurate but subject to a very slight distortion: when *p* is false but very close to being true, one falsely believes *p*. Since one's belief in *p* leaves a wide margin for error, it would have been false only if things had been very different; it may well be that if *p* had been false, it would still have been very close to being true. For example, if someone is very much less than two metres tall, I may know ↳ by sight that she is less than two metres tall, even though, if she were not less than two metres tall, she would be only very slightly more than two metres tall, and I would falsely judge her by sight to be less than two metres tall. Such cases falsify accounts of knowledge on which a necessary condition for knowing *p* is that if *p* were false one would not believe *p*. Sophisticated modifications of such accounts are refuted by p. 18

cases in which the distortion in judgement is slight but ubiquitous. This result bears on some sceptical arguments, for we can take *p* to be the proposition that I am not in a sceptical scenario α. If *p* were false I would be in α and (by construction of α) would believe that I was not in α, but that does not justify the sceptical claim 'I do not know that I am not in α', for the counterfactual condition is not necessary for knowledge. Although it can still be insisted that a necessary condition for knowing *p* is that if *p* were false one would not believe *p* on the very same evidence, that counterfactual does not justify the sceptical claim, for the sceptic has not shown that one would have the very same evidence in the sceptical scenario. Given the account in chapter 9, in a sceptical scenario one's evidence is so radically impoverished that one is not in a position to know that it is impoverished at all.

Margins for error constitute a kind of epistemic friction. For some purposes it is useful to idealize them away in thought experiments, but a world in which there were no margins for error would be as different from our world as would a world in which there was no friction. We have no more reason to postulate that there really is a kind of knowledge of temporal matters without margins for error than we have to postulate that somewhere in space there really is a frictionless plane.

6 Unknowable Truths

When knowing *p* requires a margin for error, the cases in which *p* is known are separated from the cases in which *p* is false by a buffer zone, a protective belt of cases in which *p* is true but unknown. That belt has the peculiarity that one cannot know that one is in it. For to know that would be to know that *p* is true and unknown; but knowing that involves knowing that *p* is true (since knowing a conjunction involves knowing its conjuncts); then *p* is not unknown, so it is not true that *p* is true and unknown, so it is not known that *p* is true and unknown (since only truths are known). Thus it is impossible to know that *p* is true but unknown. When *p* is in the protective belt, that is an unknowable truth.

The limits on knowledge in question are of a stronger kind than anything established by the anti-luminosity argument of Chapter 4. A proposition requires a margin for error precisely so that it can be known; the point of the anti-luminosity argument is just that the cases in which *p* is available to be known do not exhaust the cases in which *p* is true. By contrast, the point about the conjunctive proposition that *p* is true and unknown is that, in virtue of its structure, it is not available to be known in any case whatsoever. The argument for this conclusion was first published by Fitch in 1963. Contrapositively, he showed that all truths are knowable only if all truths are known. This is sometimes known as the Paradox of Knowability, although why it should be thought to constitute a paradox is unclear. It is the topic of Chapter 12.

Attempts have been made to take the sting out of Fitch's argument. Although you cannot know today the conjunction that *p* is true and you do not know *p* today, you can know the conjunction that *p* is true and you did not know *p* yesterday, and you can know the conjunction that *p* is true and I do not know *p* today. Thus one might distinguish a context in which Fitch's conjunction is true from a context in which its truth in the former context is known. The immediate response is to generalize the second conjunct, as in '*p* is true and no one ever knows *p*'. Fitch's argument shows that no one can ever know the conjunction that *p* and no one ever knows *p*. Many truths are never known by anyone. For example, either it is true that I had an even number of books in my office exactly a year ago or it is true that I had an odd number of books in my room exactly a year ago; no one will ever know which, because they were not counted at the time and it is now too late to find out. Nevertheless, one might try to distinguish a possible world *w* in which the conjunction (that *p* is true and no one ever knows *p*) is true from a possible world *w** in which the truth of the conjunction in *w* is known. The trouble with this move is that it promises only trivial knowledge. We specify merely possible worlds by description; in *w** we can describe a world as one in which *p* is true, and thereby know that *p* is true in such a world, but that is hardly a notable achievement. Section 12.5 argues that this knowledge of other possible worlds does not significantly relax the limits on knowability that Fitch's argument identifies.

Section 12.2 discusses a more direct challenge. Fitch's argument uses the distribution principle that knowledge of a conjunction implies knowledge of its conjuncts. Although the principle sounds compelling, a few accounts of knowledge are inconsistent with it. Probably that indicates something wrong with those accounts. As a precaution, ways are explored of modifying Fitch's argument to avoid relying on the distribution principle.

Limits on knowledge have counterparts in limits on action. On at least one interpretation, the relation between an agent and a proposition of *making true* shares the formal features of knowing needed for Fitch's argument: it distributes over conjunction (one makes a conjunction true only if one makes its conjuncts true) and is factive (if one makes something true then it is true). Thus no one can ever make this conjunction true: *p* and no one ever makes *p* true. For if one makes the conjunction true, one makes the first conjunct true, so the second conjunct is false, so the conjunction is false, so one did not make it true after all. If some truths were not made true by anyone, then some truths could not have been made true by anyone. Much of the world is outside both our control and our ken. We should find the limits on our knowledge scarcely more surprising than the limits on our action. Although knowledge and action are central to mind, mind is not central to world.