9
Evidence

9.1 KNOWLEDGE AS JUSTIFYING BELIEF

Tradition has it that the main problems of philosophy include the nature of knowledge. But, in recent decades, questions of knowledge seem to have been marginalized by questions of justification. Thus, according to Crispin Wright,

> knowledge is not really the proper central concern of epistemologico-sceptical enquiry. . . . We can live with the concession that we do not, strictly, *know* some of the things we believed ourselves to know, provided we can retain the thought that we are fully justified in accepting them. (1991: 88; Wright's italics)

Similarly, John Earman argues that accounts of knowledge are irrelevant to the philosophy of science, because in it 'the main concern is rarely whether or not a scientist 'knows' that some theory is true but rather whether or not she is justified in believing it' (1993: 37).[1] Once Gettier showed in 1963 that justified true belief is insufficient for knowledge, and therefore that knowledge is unnecessary for justified true belief, it became natural to ask: if you can have justified true beliefs, why bother with knowledge?[2]

The argument of Chapter 3 indicated that if one is disposed to respond rationally to future evidence, then one's future prospects are better if one now has knowledge than if one now has mere justified true belief. But even if we restrict the comparison between knowledge and justified true belief to what they do for one in the present, there is still a lacuna in the case for the unimportance of knowledge. Grant, for the sake of argument, that knowledge is important now only if it is some-

[1] Earman is discussing externalist accounts of knowledge, but the quoted comment would clearly apply to internalist accounts too. Earman's further point, that 'because science is a community enterprise the only forms of justification that are scientifically relevant are those which are stateable and open to public scrutiny', may be most relevant to externalist accounts. See also Craig 1990b: 272. For the contrary view that scepticism about knowledge entails scepticism about rationality and justification, see Unger 1975: 197–249.

[2] Kaplan 1985 argues along similar lines.

how essential to the present justification of belief.[3] Although it has been shown that *what is justified* need not be knowledge, even when it is true, it has not been shown that *what justifies* need not be knowledge. Only one end of the justification relation has been separated from knowledge. Suppose that knowledge, and only knowledge, justifies belief. That is, in any possible situation in which one believes a proposition p, that belief is justified, if at all, by propositions $q_1, \ldots, q_n$ (usually other than p) which one knows. On that supposition, if justified belief is central to epistemologico-sceptical inquiry and the philosophy of science, then so too is knowledge. Now assume further that what justifies belief is *evidence* (this assumption is briefly discussed in section 9.8). Then the supposition just made is equivalent to the principle that knowledge, and only knowledge, constitutes evidence. This chapter defends that principle; it equates S's evidence with S's knowledge, for every individual or community S in any possible situation. Call this equation $E = K$.[4]

As usual, 'knowledge' is understood as propositional knowledge. The communal case is needed: science depends on public evidence, which is neither the union nor the intersection of the evidence of each scientist. We can ascribe such knowledge by saying that p is known in community S, or that we know p, which is not equivalent to saying that some, many, most, or all of us know p.

The proposed account uses the concept of knowledge in partial elucidation of the concepts of evidence and justification. To some people it will therefore seem to get things back to front. For although knowledge is more than justified true belief, many philosophers still expect to use concepts such as *evidence* and *justification* in a more complex explanation of the concept *knows*; it would then be circular to use the latter to explain the former. Others prefer to use concepts of a different kind, such as *causation* or *reliability*, to explain the concept *knows*; but even they are likely to regard the concept *knows* as so much in need of explanation itself that its pre-theoretic use would lack explanatory value.

That order of explanation has been reversed in this book. The concept *knows* is fundamental, the primary implement of epistemological inquiry. Chapter 1 rejected the programme of understanding knowledge in terms of the justification of belief. That frees us to try the experiment

[3] Wright speaks of justified acceptance rather than belief; although Earman speaks of belief, some philosophers of science contrast it with acceptance and regard the latter as a more appropriate attitude towards scientific theories. The distinction is not important here.

[4] The principle is stated, and applied to the problem of vagueness, in Williamson 1994b: 245–7.

of understanding the justification of belief in terms of knowledge. Of course the concept *knows* is vague; so is the concept *justified*.

For those who remain sympathetic to the orthodox order of explanation, some more irenic points can be made. The equation E = K could be true without being knowable a priori. As a universal generalization over all metaphysically possible situations, it is necessarily true, if true at all (by the S4 principle that a necessary truth is necessarily necessary); but we cannot presume that a necessary truth is knowable a priori. E = K equates the extensions of the concepts *knowledge* and *evidence* in any possible situation; that is enough to make it an informative thesis. By itself, E = K does not equate the concepts themselves; nor is it to be read as offering an analysis of either the concept *evidence* or the concept *knowledge*, or as making one concept prior to the other in any sense. Of course, in offering *arguments* of a broadly a priori kind for E = K, like those below, one commits oneself at least to its a priori plausibility; in the best case for those arguments, they would provide a priori knowledge of E = K. But even if the concepts are equivalent a priori, it does not follow that one is prior to the other.

More positively, we may speculate that standard accounts of justification have failed to deal convincingly with the traditional problem of the regress of justifications—what justifies the justifiers?—*because* they have forbidden themselves to use the concept *knowledge*. E = K suggests a very modest kind of foundationalism, on which all one's knowledge serves as the foundation for all one's justified beliefs. Perhaps we can understand how something could found belief only by thinking of it as knowledge.

9.2 BODIES OF EVIDENCE

When is *e* evidence for the hypothesis *h*, for a subject S? Two conditions seem to be required. First, *e* should speak in favour of *h*. Second, *e* should have some kind of creditable standing. At least as a first approximation, we can model the first condition in probabilistic terms: *e* should raise the probability of *h*. That is, the probability of *h* conditional on *e* should be higher than the unconditional probability of *h*; in symbols, $P(h|e) > P(h)$. The conditional probability $P(h|e)$ is defined as the ratio $P(h \wedge e)/P(e)$ when $P(e) \neq 0$, and is otherwise undefined. Thus the condition that $P(h|e) > P(h)$ obtains if and only $P(h \wedge e) > P(h)P(e)$. What kind of probability is P? It is not a priori, for whether *e* raises the probability of *h* may depend on background information. For example,

the proposition that John belongs to a certain club might raise the probability that he is single relative to the background information that it is a club for singles, but lower it relative to the background information that it is a club for spouses. However, *e* itself should not be built into the background information, for that would give $P(e)$ the value 1, in which case $P(h|e)$ and $P(h)$ would be equal and *e* would not be evidence for anything. Let us leave the nature of P underspecified until the next chapter. Now, *e* may raise the probability of *h* in the sense that $P(h|e) > P(h)$ even if S knows that *e* is false or has no idea whether *e* is true; but then, for S, *e* would not be evidence for *h*. That is why we need the second condition, that *e* should have a creditable standing. A natural idea is that S has a *body of evidence*, for use in the assessment of hypotheses; that evidence should include *e*. The probability distribution P is informed by some but not all of S's evidence. We can therefore formulate a simple schematic proposal:

EV *e* is evidence for *h* for S if and only if S's evidence includes *e* and $P(h|e) > P(h)$.

One consequence of EV is that *e* is evidence for *h* only if *e* is evidence for itself. For if $P(h|e) > P(h)$, then $P(e)$ is neither 0 (otherwise $P(h|e)$ is ill defined) nor 1 (otherwise $P(h|e) = P(h)$). Hence $P(e|e)$ is well defined with the value 1, which is greater than $P(e)$, so *e* is evidence for *e*, by EV with '*e*' substituted for '*h*'. But is it not circular for anything to be evidence for itself? A critic might therefore argue that one's evidence does not consist of a fixed body of propositions; either it depends on the hypothesis under assessment, where no proposition belongs to the evidence relative to its own assessment, or it does not consist of propositions.

The critic is not entitled to assume without argument that classifying *e* as evidence for itself involves circularity of any vicious kind. Certainly EV does not make it trivially easy to have evidence for *e*, for *e* is evidence for itself for S only if S's evidence includes *e*. By E = K, that requires S to know *e*, which may not be easy. The result that *e* is evidence for itself may be as harmless as the consequence of a standard definition of provability in a formal system that every axiom has a one-line proof, consisting of the axiom itself. Of course, if someone asks 'What is the evidence for *h*?', one is not expected to cite *h* itself, but the reason might be that it would be conversationally inappropriate rather than false to do so. In answer to the question 'Who lives in the same house as Mary?' it would be conversationally inappropriate to cite Mary herself; nevertheless, it is true that Mary lives in the same house as Mary (Grice 1989). The question 'What is the evidence for *h*?' is often a challenge to the epistemic standing of *h* and related propositions. In some contexts the challenge is local, restricted to propositions derived in some way from *h*. In other contexts the challenge is global, extending to all propositions with the same kind of pedigree as *h*. In answering the question, one is expected not to cite propositions under challenge, since their status as evidence has been challenged. Thus when the question 'What is the evidence for *e*?' is meant as a challenge to the epistemic standing of *e*, one is expected not to cite *e* in response.

Could we treat the claim that *e* is evidence for itself as false rather than conversationally inappropriate by treating 'evidence' as a context-sensitive term? The idea would be that the question 'What is the evidence for *e*?', meant as a challenge, creates a context in which *e* falls outside the extension of 'S's evidence'. But that seems too drastic. For example, suppose that a doctor asks you, 'Do you feel a tingling sensation?' and you answer, 'No.' If you were asked 'What is your evidence for the proposition that you do not feel a tingling sensation?', you might be at a loss to answer, for the question seems to expect some *further* evidence for the proposition, and you might look in vain for such further evidence. Nevertheless, when we assess the status of your claim that you did not feel a tingling sensation on your evidence, we do not exclude that proposition from your evidence. Its presence justified your claim. This is not to deny that the extension of 'evidence' may vary slightly with context, perhaps corresponding to slight contextual variation in the extension of 'knowledge' (and therefore, presumably, in the extension of 'mental state' too, by section 1.4). The point is just that challenging *e* by itself is not enough to exclude *e* from the extension of 'evidence'.

One sceptical strategy is to exploit the dialectical effects of challenging propositions. If one is never entitled to rely on something under challenge, one will very soon be left with very little. For example, the sceptic can challenge the belief that there are good reasons, and then charge any attempt to provide a good reason for it with begging the question. We should be sceptical of such a sceptic's reliance on the power of challenge. The sceptic relies uncritically on rules of dialectical engagement which evolved to serve more practical purposes, without questioning their appropriateness to the radical questions which scepticism raises. If challenging something thereby makes it dialectically unusable, then the power of challenge might hinder rather than help the pursuit of truth if it is not used with restraint. By refusing to associate questions of evidence too closely with questions of dialectical propriety, we can preserve EV.

EV concerns the evidence-for relation, as do most discussions of evidence.[5] The focus of this chapter is elsewhere. It concerns the nature of the first relatum *e* of the evidence-for relation rather than its relation to the second relatum *h*. Whether EV needs revision will be left open; the present aim is to investigate its constituent 'S's evidence includes *e*'. Chapter 10 develops a theory of evidential probability to address the relation between evidence and what it supports.

Why does it matter what counts as evidence? Consider the idea that one should proportion one's belief in a proposition to one's evidence for it. How much evidence one has for the proposition depends on what one's evidence is. More precisely, a theory of evidence is needed to give bite to what Carnap calls *the requirement of total evidence*:

> [I]n the application of inductive logic to a given knowledge situation, the total evidence available must be taken as a basis for determining the degree of confirmation. (1950: 211; compare Hempel 1965: 63–7)

If too much or too little is counted as evidence, inductive principles will be misapplied. Given the requirement of total evidence, disputes between different theories of evidence are not merely verbal; they involve disagreements as to which inductive conclusions are warranted. Formulations of the total evidence requirement in terms of knowledge encourage E = K, which identifies the total evidence available with the total knowledge available. For example, Peirce writes:

> I cannot make a valid probable inference without taking into account whatever knowledge I have (or, at least, whatever occurs to my mind) that bears on the question. (1932: 461)

Carnap himself describes the evidence as (observational) knowledge. Given E = K, the original idea becomes something like this: one should proportion one's belief in a proposition to the support which it receives from one's knowledge.

The total evidence available must not be built into the probability distribution P in EV, otherwise no part of that evidence could confirm any hypothesis. In general, the total evidence must not be taken as a basis for determining the degree to which an individual piece *e* of that evidence increases the confirmation of a hypothesis *h*, for if *e* is part of the total evidence available then the confirmation of *h* prior to the acquisition of the total evidence presently available is also relevant. This is a form of the problem of old evidence (Glymour 1980: 85–93, Earman 1992: 119–35, Howson and Urbach 1993: 403–8, Maher 1996), which is

[5] The papers in the representative collection Achinstein 1983, for example, are largely concerned with questions about the evidential relation at the expense of questions about its first relatum.

discussed in the next chapter. But that does not undermine the point that the total evidence now available must be taken as the basis for determining the degree to which *h* is now confirmed in the non-comparative sense.

Theories of evidence also play a role when theses of the underdetermination of theory by data are assessed, for the data in question are the actual or potential evidence. If too much or too little is counted as evidence, then the standard for underdetermination is set uninterestingly high or uninterestingly low. E = K implies that undetermination theses of the relevant kind must count all knowable facts as data. Although this condition does not automatically make any argument for underdetermination circular, it is not easily met. Consider, for example, the underdetermination thesis that the theoretical facts do not supervene on the evidential facts: two possible worlds can differ in the former without differing in the latter.[6] One cannot establish this claim just by showing that the theoretical facts do not supervene on the facts which are in some sense observable; one must also show that they do not supervene on all the knowable facts. The gap would be filled by an argument that the knowable facts supervene on the observable facts, for then whatever failed to supervene on the observable facts would fail to supervene on the knowable facts too, by the transitivity of supervenience. But any such argument risks begging the question against the view that at least some theoretical facts are knowable.

9.3 ACCESS TO EVIDENCE

Chapter 8 argued that we are not always in a position to know what our evidence is. Consequently, a theory of evidence cannot be expected to provide a decision procedure which will always enable us to determine in practice whether our evidence includes a given item. In general, a philosophical theory of a concept is not required to provide a decision procedure which will always enable us to determine in practice whether it applies to a given item. The concept of evidence might have been expected to be special in this respect, for if it were problematic whether one's evidence included something, one would need evidence to decide whether one's evidence included it, and an infinite regress looms. It is therefore tempting to suppose both that it must be unproblematic

[6] Compare EI_3 in the useful classification of kinds of empirical indistinguishability in Earman 1993: 21.

whether one's evidence includes any given item, and that an adequate theory of evidence must explain how it manages to be so unproblematic. By the argument of Chapter 8, however, no correct theory of evidence can have that upshot. Certainly the equation E = K does not, but since that does not distinguish it from other theories of evidence, it constitutes no objection to E = K. In obvious symbolism, E = K equates E*p* and K*p*. The transparency of evidence would make E*p* equivalent to KE*p*. Given E = K, that is tantamount to making K*p* equivalent to KK*p*. But we saw in section 5.1 that K*p* does not entail KK*p*. This section explores our limited access to our evidence in the light of the equation E = K.

There is no infallible recipe for deciding in practice whether we know a proposition *p*. Sometimes we reasonably believe ourselves to know *p*, when in fact we do not know *p*, because *p* is false. Reputable authorities assert that Henry V died in 1422; I have no grounds for doubting them, and reasonably believe myself to know by testimony that Henry V died in 1422. But it is not inconceivable that he died in 1423, some elaborate conspiracy being responsible for present evidence to the contrary. According to E = K, if I do know that Henry V died in 1422, then my total evidence includes the proposition that Henry V died in 1422; but if Henry V died in 1423, then my belief that my total evidence includes that proposition is mistaken—my total evidence includes only the proposition that reputable authorities assert that Henry V died in 1422. E = K is an externalist theory of evidence, in at least the sense that it implies that one's evidence does not supervene on one's internal physical states. But if knowing is a mental state, as argued in Chapter 1, then one's evidence does supervene on one's *mental* states.

How does E = K avoid the threatened regress of evidence? The regress comes if evidence-based belief in a proposition *p* must always be preceded by evidence-based belief in a proposition about the evidence for *p*. We can distinguish two senses of 'evidence-based'. Call one's belief in *p* *explicitly* evidence-based if it is influenced by prior beliefs about the evidence for *p*. Explicitly evidence-based beliefs may be more common in science than in everyday life. Call one's belief in *p* *implicitly* evidence-based if it is appropriately causally sensitive to the evidence for *p*. A belief can be both explicitly and implicitly evidence-based. Now, explicitly evidence-based belief in *p* is not always preceded by *explicitly* evidence-based belief in a proposition about the evidence for *p*; this is consistent with E = K and most other theories of evidence. An explicitly evidence-based belief is influenced by a prior state of belief in a proposition about the evidence for *p*, and something has gone wrong if the latter belief is not at least implicitly evidence-based; but it need not be explicitly evidence-based. Thus there is no regress of explicitly evidence-based belief. There would be a different regress if implicitly evidence-based belief in p were always preceded by implicitly evidence-based belief in a proposition about the evidence for p. But the causal sensitivity of the belief in p to the evidence for p need not be mediated by further *beliefs* about the evidence for p. There need be no such beliefs.

How can a belief in p be implicitly evidence-based, if we are liable to misidentify the evidence for p? If the real evidence differs from the apparent evidence, will not the belief be causally sensitive to the latter rather than the former? But, as noted in section 8.7, we are liable to misidentify the apparent evidence, too. Causal sensitivity need not be perfect to be genuine. There can be a non-accidental rough proportionality between the strength of the belief and the strength of the evidence, even if distortions sometimes occur.

Similar questions arise about explicitly evidence-based belief. How can one follow the rule 'Proportion your belief in p to your evidence for p' when one doesn't know exactly what one's evidence is? Given E = K, the rule becomes 'Proportion your belief in p to the support that p receives from your knowledge': but is one not at best following the rule 'Proportion your belief in p to the support that p receives from *what you believe to be* your knowledge'? Consider an analogy. We can follow the rule 'Proportion your voice to the size of the room'. This is not because we are infallible about the size of the room. We sometimes make mistakes; but it does not follow that we are really following the rule 'Proportion your voice to *what you believe to be* the size of the room'. After all, it is often quite hard to know what beliefs one has about the size of a room; we are fallible in our beliefs about such beliefs. That one believes p is not a luminous condition. In general, if the fallibility of our beliefs about X posed a problem, it would not be solved by the move to our beliefs about our beliefs about X, because they are fallible too. But fallibility does not pose a problem here. To make a mistake in following a rule is not to follow a different rule. The rule is a standard of correctness for action, not a description of action. To have applied the rule 'Proportion your voice to the size of the room', one needs beliefs about the size of the room, but they need not have been true—although if they were false, one's application was faulty. Similarly, to have applied the rule 'Proportion your belief in p to the support that p receives from your knowledge', one must have had beliefs about how much support p received from one's knowledge, and therefore about one's knowledge, but those beliefs need not have been true—although if they were false, one's application was faulty.

None of this would be much consolation if our beliefs about our

knowledge were hopelessly unreliable. Sceptics say that those beliefs have no rational basis, but they say the same about most of our other beliefs, too. We have found their reasons for saying so to be inadequate. Although we have no infallible procedure for determining whether we know p, in practice we are often in a position to know whether we know p.

The ways in which we decide whether we know p are not simply the ways in which we decide whether we believe that we know p. If I want to check whether I now really know that Henry V died in 1422, it would be relevant to return to my sources; it would be irrelevant to do that if I merely wanted to check whether I now really believe that I know that he died in 1422.

As Chapter 8 noted, alternative theories of evidence distort the concept in the attempt to make evidence something that we can infallibly identify. Characteristically, they interiorize evidence: it becomes one's present experience, one's present degrees of belief, or the like. Those attempts are quaint relics of Cartesian epistemology. Knowledge of the present contents of one's own mind is neither unproblematic nor prior to knowledge of other things. It is not obvious to me how many shades of blue I am presently experiencing, or to what degree I believe that there was once life on Mars. If one's evidence were restricted to the contents of one's own mind, it could not play the role that it actually does in science. The evidence for the proposition that the sun is larger than the earth is not just my present experiences or degrees of belief. If the evidence is widened to include other people's experiences or degrees of belief, or my past ones, then my identification of it becomes even more obviously fallible. In any case, that does not seem to be the right widening; it is more plausible that the evidence for a scientific theory is the sort of thing which is made public in scientific journals. If evidence is like that, our identification of it is obviously fallible.

9.4 AN ARGUMENT

Here is a schematic argument for E = K:

> All evidence is propositional.
> All propositional evidence is knowledge.
> All knowledge is evidence.
> ______________________________
> All and only knowledge is evidence.

The argument is obviously valid, but its premises are contentious. Its aim is simply to divide the contentiousness of the conclusion into manageable portions; sections 9.5, 9.6, and 9.7 respectively defend the three premises. Since 'knowledge' here means propositional knowledge, each premise follows from the conclusion; thus the conclusion is equivalent to the conjunction of the premises.

One's evidence is propositional if and only if it is a set of propositions. Propositions are the objects of propositional attitudes, such as knowledge and belief; they can be true or false; they can be expressed relative to contexts by 'that' clauses. For present purposes, we do not need a developed theory of propositions. If evidence is propositional, we can refer to evidence by using 'that' clauses: my evidence for the conclusion that the house was empty is *that* it was silent, *that* no lights were on in the evening, *that* the telephone went unanswered, ...

9.5 EVIDENCE AS PROPOSITIONAL

Why should all evidence be propositional? It would not be on a broad interpretation of 'evidence'. In the courts, a bloodied knife is evidence. It is natural to say that my evidence that I am getting a cold includes various sensations. Some philosophers apply the term 'evidence' to non-propositional perceptual states; Quine restricts it to the stimulation of sensory receptors (1969: 75). How can 'All evidence is propositional' do more than stipulate a technical use for the word 'evidence'?

Indiscriminate description of the ordinary use of a term and arbitrary stipulation of a new use are not the only options. We can single out theoretical functions central to the ordinary concept *evidence*, and ask what serves them. That strategy is pursued here. The argument below substantiates the familiar claim that only propositions can be reasons for belief (for example, Unger 1975: 204–6 and Davidson 1986; for opposing views, Moser 1989: 47–125 and Millar 1991). It also suggests a further conclusion: one grasps the propositions that are one's evidence; one can think them.

Consider inference to the best explanation (Harman 1965, Lipton 1991). We often choose between hypotheses by asking which of them best explains our evidence—which of them, if true, would explain the evidence better than any other one would, if true. Fossil evidence enables us to answer questions about terrestrial life in this way. Even if inference to the best explanation is not legitimate in all theoretical contexts, what matters for present purposes is that, where evidence does enable us to answer a question, a central way for it to do so is by inference to its best explanation. Thus evidence is the kind of thing which hypotheses explain. But the kind of thing which hypotheses explain is propositional. Therefore evidence is propositional.

The kind of thing which hypotheses explain is propositional. Inference to the best explanation concerns why-explanations, which can be put in the form '—— because . . .', which is ungrammatical unless declarative sentences, complements for 'that', fill both blanks. We cannot simply *explain Albania*, for 'Albania because . . .' is ill-formed. We can sometimes make sense of the injunction 'Explain Albania!', but only when the context allows us to interpret it as an injunction to explain why Albania exists, or has some distinctive feature. What follows 'why' is a declarative sentence, expressing the proposition to be explained—*that* Albania exists, or *that* it has the distinctive feature. It makes no significant difference if what is to be explained is one thing as contrasted with another (Lipton 1991: 75–98). For example, we may seek to explain why Kosovo rather than Bosnia was peaceful in 1995. The evidence in question would be the propositions that Kosovo was peaceful in 1995 and that Bosnia was not. The same goes for events: 'Explain World War I!' enjoins one to explain why it occurred, or had some distinctive feature. Again, the sensation in my throat is evidence for the conclusion that I am getting a cold in the sense that the hypothesis that I am getting a cold would best explain why I have that sensation in my throat. The evidence to be explained is *that* I have that sensation in my throat—not just that I have *a* sensation in my throat. Even in the courts, the bloodied knife provides evidence because the prosecution and defence offer competing hypotheses as to why it was bloodied or how it came into the accused's possession; the evidential proposition is *that* it was bloodied or *that* it came into the accused's possession. The knife is a source of indefinitely many such propositions.

One can use an hypothesis to explain why A only if one grasps the proposition that A. Thus only propositions which one grasps can function as evidence in one's inferences to the best explanation. By this standard, only propositions which one grasps count as part of one's evidence.

Similar points apply to explicitly probabilistic reasoning. If such reasoning can be assimilated to inference to the best explanation, or vice versa, so much the better. The best way of comparing the conditional probabilities of two hypotheses h and h^* on evidence e, $P(h|e)$ and $P(h^*|e)$, is often by calculating the inverse probabilities of e on h and h^*, $P(e|h)$, and $P(e|h^*)$. For example, a bag contains ten red or black balls; we wish to estimate how many of them are red; we are allowed to gain evidence only by sampling one ball at a time, noting its colour and replacing it. A good way to compare the probabilities of hypotheses about the number of red balls is by calculating the probabilities of the actual outcome *e* of the sampling (say, red fifteen times and black five times) on those hypotheses. One way of using those probabilities is to regard *h* as more probable than h^* given *e* ($P(h|e) > P(h^*|e)$) if and only if *h* makes *e* more probable than h^* does ($P(e|h) > P(e|h^*)$). Bayesians take this method to involve assigning the same prior probability to *h* and h^* ($P(h) = P(h^*)$); they treat as equally legitimate assignments of unequal prior probabilities to the hypotheses—perhaps reflecting differences in explanatory virtues such as simplicity and elegance. To allow for such cases, their general rule weights the probability of *e* on *h* by the prior probability of *h*; thus $P(h|e) > P(h^*|e)$ if and only if $P(h)P(e|h) > P(h^*)P(e|h^*)$, where $P(e)$, $P(h)$, and $P(h^*)$ are all non-zero. For present purposes, it does not matter whether Bayesians are right to introduce prior probabilities here. The point is that such probabilistic comparisons of hypotheses on the evidence depend on the probabilities of the evidence on the hypotheses. But what has a probability is a proposition; the probability is the probability *that* . . .. At least, that is so when 'probability' has to do with the evidential status of beliefs, as now; if we speak in this connection of the probability of an event, we mean the probability *that it occurred.*[7] We might initially suppose that, in $P(x|y)$, only *x* need be a proposition, but the relation between $P(x|y)$ and $P(y|x)$ means that *y* must be a proposition too; what gives probability must also receive it. Moreover, these probabilities, as measures of degrees of belief warranted by evidence, are idle unless the subject grasps *x* and *y*.

More straightforward uses of evidence also require it to be propositional. In particular, our evidence sometimes rules out some hypotheses by being *inconsistent* with them. For example, the hypothesis that only males have varicose veins is inconsistent with much medical evidence. But only propositions can be inconsistent in the relevant sense. If evidence *e* is inconsistent with an hypothesis *h* in that sense, it must be possible to *deduce* ~*h* from *e*; the premises of a deduction are propositions. Moreover, the subject who deduces ~*h* from *e* must grasp *e*.

Only propositions which we grasp serve the central evidential functions of inference to the best explanation, probabilistic confirmation, and the ruling out of hypotheses. Could non-propositional items count as evidence by serving other central functions of evidence? For example, they might serve as the inputs to a non-inferential process whose outputs were beliefs. But suppose that we are choosing between hypotheses according to which best explains our evidence, or is most probable on our evidence, or is not ruled out by our evidence. The argument so far shows that only propositional evidence would be directly relevant to our choice. Moreover, in choosing between hypotheses in those ways, we can use only propositions which we grasp. In those respects, any evidence other than propositions which we grasp would be impotent. Although evidence may well have central functions additional to those considered above, genuine evidence would make a difference to the serving of the functions considered above, whatever else it made a difference to. Certainly, defences of non-propositional evidence have not been based on an appreciation of its impotence in those respects. Since only propositions we grasp make a difference of the requisite kind, only propositions which we grasp are our evidence.

A positive case for that conclusion has now been given. Nevertheless, perceptual experience is often regarded as a kind of non-propositional evidence. Do the considerations above somehow fail to do it justice? The remainder of this section will rebut objections to the view that our perceptual evidence consists of propositions which we grasp.

Experiences provide evidence; they do not consist of propositions. So much is obvious. But to provide something is not to consist of it. The question is whether experiences provide evidence just by conferring the status of evidence on propositions. On that view, consistent with E = K, the evidence for an hypothesis *h* consists of propositions $e_1, \ldots, e_n$, which count as evidence for one only because one is undergoing a perceptual experience ε. As a limiting case, *h* might be e_i. The threatening alternative is that ε can itself be evidence for *h*, without the mediation of any such $e_1, \ldots, e_n$. Both views permit ε to have a non-propositional, non-conceptual content, but only the latter permits that content to function directly as evidence.

If perceptual evidence consists of propositions, which propositions are they? Consider an example. I am trying to identify a mountain by its shape. I can see that it is pointed; that it is pointed may be part of my evidence for believing that it is not Ben Nevis. However, the proposition that it is pointed does not begin to exhaust my present perceptual evidence. No description of the mountain in words seems to capture the richness of my visual experience of its irregular shape. But it does not follow that my evidence is non-propositional. If I want to convey my

[7] Objective probabilities, in the sense of chances determined in the natural world independently of our beliefs, are irrelevant here. We sometimes speak, too, of the probability of one property, concept, or predicate conditional on another—for example, the probability of lung cancer conditional on smoking—but the probabilities relevant to the argument are the probabilities of hypotheses, which, unlike properties, concepts, and predicates, have truth-values. That someone has lung cancer is evidence that he smoked; the unattributed property of lung cancer is not by itself evidence of anything.

evidence, I might point and say 'It is that shape'.[8] Of course, the mere linguistic meaning of the sentence type 'It is that shape' does not convey my evidence, for it is independent of the reference of 'that shape' in a particular context of utterance. Only by using the sentence in an appropriate context do I express the proposition at issue. My token of 'that shape' still expresses a constituent of that proposition, even if you cannot grasp that constituent without having a complex visual experience with a structure quite different from the constituent structure of the proposition. The proposition that the mountain is that shape is contingent; it could have been another shape. The proposition is also known a posteriori; I do not know a priori that I am not including the tip of another mountain behind in the profile. But in ordinary circumstances I can know that the mountain is that shape, and a fortiori grasp the proposition that it is, when 'that shape' does not refer to an absolutely specific shape. Of course, I cannot see *exactly* what shape the mountain is; I can only see roughly what profile it presents to me, and cannot see round the back. That shape must be unspecific enough to give my knowledge that the mountain is that shape an adequate margin for error in the sense of Chapter 5.[9] The knowledge that the mountain is that shape is obtainable in other contexts; you can have it too, and we can retain it in memory. Properties other than shape are similar in those respects.

In unfavourable circumstances, one fails to gain perceptual knowledge, perhaps because things are not the way they appear to be. One does not know that things are that way, and E = K excludes the proposition that they are as evidence. Nevertheless, one still has perceptual evidence, even if the propositions it supports are false. True propositions can make a false proposition probable, as when someone is skilfully framed for a crime of which she is innocent. If perceptual evidence in the case of illusions consists of true propositions, what are they? The obvious answer is: the proposition that things appear to be that way. The mountain appears to be that shape. Of course, unless one has reason to suspect that circumstances are unfavourable, one may not consider the cautious proposition that things appear to be that way; one

[8] See McDowell 1994: 56–9 for such a proposal. It is not being used here to deny that perceptual experience has non-conceptual content (see Peacocke 1992: 84). Christensen 1992: 545 discusses such beliefs in a Bayesian context, asking whether they can 'connect with other beliefs in the way that would be necessary for them to fulfill their intended evidential role'. The connections will not be purely syntactic, but fifty years of confirmation theory have shown that confirmation is not a purely syntactic matter.

[9] The unspecificity makes the present proposal closer to that of McDowell 1994: 170–1 than to that of Peacocke 1992: 83–4.

may consider only the unqualified proposition that they really are that way. But it does not follow that one does not know that things appear to be that way, for one knows many propositions without considering them. When one is walking, one normally knows that one is walking, without considering the proposition. Knowing is a state, not an activity. In that sense, one can know without consideration that things appear to be some way. When I believe falsely that circumstances are favourable, I believe falsely that I am gaining perceptual knowledge about the environment, and therefore that my evidence includes those propositions believed to be known. But our fallibility in identifying our evidence is nothing new, and my actual evidence may justify my false beliefs about my evidence.

In order to grasp the proposition that things appear to be some way, one must grasp the property of appearing, on the assumption that the semantically significant constituents of a sentence express constituents of the proposition expressed by the whole sentence. Although one's grasp of the property of appearing may be inarticulate, one must have some inkling of the distinction between appearance and reality. For instance, one should be willing in appropriate circumstances to give up the belief that things were that way while retaining the belief that they appeared to be that way. In the absence of such dispositions, it is implausible to attribute the qualified belief that things appear to be that way rather than the unqualified belief that they are that way. Perhaps some young children and animals have beliefs and perceptual experiences without even implicitly grasping the property of appearance. Suppose that such a simple creature is given a drug which causes the hallucinatory appearance that there is food ahead; as a result, it comes to believe falsely that there is food ahead. Does it have any evidence for that belief? According to E = K, its evidence cannot be that things appear some way, for it cannot grasp that proposition. Perhaps it knows that the situation is *like* one in which there is food ahead, where the property of likeness covers both likeness in appearance and other kinds of likeness indifferently, so that grasp of the property of likeness does not require grasp of the property of appearing. If the creature does not even know that the situation is like one in which there is food ahead, then we can plausibly deny that it has perceptual evidence that there is food ahead. It does not recognize the features of its perceptual experience which, if recognized, would provide it with evidence. *We* can use the proposition that there appears to be food ahead as evidence, but the simple creature cannot. Although the hallucinatory appearance causes a belief, that causal relation is not an evidential one.

Very simple creatures grasp no properties or propositions and have no beliefs or knowledge. It is sometimes even argued—not very plausibly—that any creature which lacks the distinction between appearance and reality is in this predicament. Simple creatures have no evidence, for they have no degrees of belief, and degrees of belief are what evidence justifies.

S can use as evidence only propositions which S grasps. Since S can use S's evidence as evidence, only propositions which S grasps are S's evidence. What has not yet been argued is that those propositions count as evidence by being known.

9.6 PROPOSITIONAL EVIDENCE AS KNOWLEDGE

Why should all propositional evidence be knowledge? The thesis is that if S's evidence includes a proposition *e*, then S knows *e*. If I do not know that the mountain is that shape, then that it is that shape is not part of my evidence. As in the previous section, the argument is from the function of evidence.[10] Indeed, the thesis draws support from the role of evidence cited there, in inference to the best explanation, probabilistic reasoning, and the exclusion of hypotheses. When we prefer an hypothesis *h* to an hypothesis *h** because *h* explains our evidence *e* better than *h** does, we are standardly assuming *e* to be known; if we do not know *e*, why should *h*'s capacity to explain *e* confirm *h* for us? It is likewise hard to see why the probability of *h* on *e* should regulate our degree of belief in *h* unless we know *e*. Again, an incompatibility between *h* and *e* does not rule out *h* unless *e* is known. But it is prudent to consider the matter more carefully.

Suppose that balls are drawn from a bag, with replacement. In order to avoid issues about the present truth-values of statements about the future, assume that someone else has already made the draws; I watch them on film. For a suitable number *n*, the following situation can arise. I have seen draws 1 to *n*; each was red (produced a red ball). I have not yet seen draw *n*+1. I reason probabilistically, and form a justified belief that draw *n*+1 was red too. My belief is in fact true. But I do not know that draw *n*+1 was red. Consider two false hypotheses:

h: Draws 1 to *n* were red; draw *n*+1 was black.

*h**: Draw 1 was black; draws 2 to *n*+1 were red.

[10] For linguistic arguments that if S's reason or justification is *e* then S knows *e*, see Unger 1975: 206–14. See also Hyman 1999.

It is natural to say that *h* is consistent with my evidence and that *h** is not. In particular, it is consistent with my evidence that draw *n*+1 was black; it is not consistent with my evidence that draw 1 was black. Thus my evidence does not include the proposition that draw *n*+1 was red. Why not? After all, by hypothesis I have a justified true belief that it was red. The obvious answer is that I do not *know* that draw *n*+1 was red; the unsatisfied necessary condition for evidence is knowledge. An alternative answer is that I have not *observed* that draw *n*+1 was red. That is equally good for the purposes of this section (although not for those of the next), for observing the truth of *e* includes *e* in my evidence only by letting me know *e*. If I observe the truth of *e* and then forget all about it, my evidence no longer includes *e*. It is hard to see how evidence could discriminate between hypotheses in the way we want it to if it did not have to be known.

If evidence required only justified true belief, or some other good cognitive status short of knowledge, then a critical mass of evidence could set off a kind of chain reaction. Our known evidence justifies belief in various true hypotheses; they would count as evidence too, so this larger evidence set would justify belief in still more true hypotheses, which would in turn count as further evidence . . .. The result would be very different from our present conception of evidence.

That propositional evidence is knowledge entails that propositional evidence is true. That is intuitively plausible; if one's evidence included falsehoods, it would rule out some truths, by being inconsistent with them. One's evidence may make some truths improbable, but it should not exclude any outright. Although we may treat false propositions as evidence, it does not follow that they are evidence. No true proposition is inconsistent with my evidence, although I may think that it is. If *e* is evidence for *h*, then *e* is true. There is no suggestion, of course, that if *e* is evidence for *h* then *h* is true. For example, that the ground is wet is evidence that it rained last night only if the ground is wet—even if it did not rain last night. If *e* is not true, then at most a counterfactual holds: if *e* had been true, *e* would have been evidence for *h*.[11] If the convincing but lying witness says that the accused was asleep at the time of the murder, then it is part of the evidence for the innocence of the accused that the witness *said* that he was asleep then. It is not part of the

[11] Stampe 1987: 337 takes a similar view of reasons. Millar 1991: 65 says that we ordinarily think of evidence as consisting in facts; that is also how we ordinarily think of the objects of knowledge. If facts are distinguished from true propositions, then the arguments of this chapter can be adjusted accordingly, but the individuation of facts must be reconciled with the individuation of knowledge and evidence. Presumably, the fact that Hesperus is bright is the fact that Phosphorus is bright. See also section 1.5.

evidence for his innocence that he was asleep, for it is consistent with the evidence that he was not. The rival view, that a false proposition can become evidence through a sufficient appearance of truth, gains most of its appeal from the assumption, disposed of in Chapter 8 and section 9.3, that we must have an infallible way of identifying our evidence.

Once it is granted that all propositional evidence is true—and therefore, by the previous section, that all evidence consists of true propositions—adjusting our beliefs to the evidence has an obvious point. It is a way of adjusting them to the truth. Although true evidence can still support false conclusions, it will tend to support truths. The maxim 'Proportion your belief to your evidence' requires more than the mere internal coherence of one's belief system; it does so because evidence must be true. Even if an internally coherent belief system cannot be wholly false, a given belief system with a given degree of internal coherence can be better or worse proportioned to the evidence, depending on what the evidence is. But, equally, the evidence is not a wholly external standard, if it is known.

Another consequence of the claim that propositional evidence is knowledge is that propositional evidence is believed—at least, if knowledge entails belief, which is granted here (see section 1.5). The case of perception may seem to suggest that propositional evidence is not always believed. In conformity with the previous section, a piece of perceptual evidence is, for example, a proposition *e* that things are *that* way. According to E = K, my evidence includes *e* because I know that things are that way. But, a critic may suggest, that does not go back far enough; my evidence includes *e* because it is perceptually apparent to me that things are that way, whether or not I believe that they are that way. Even if I do believe *e*, my evidence included *e* before I came to believe it; according to the critic, I came to believe it because it was perceptually apparent. If 'It is perceptually apparent that A' entails 'A', then the critic's view allows that evidential propositions are always true; what it denies is that they are always believed, and therefore that they are always known.

If my evidence includes a proposition *e*, then I grasp *e*, by section 9.5. Thus, if I fail to believe *e*, my problem is not conceptual incapacity. Perhaps I have simply not had time to form the belief; perhaps I suspect, for good or bad reasons, that I am the victim of an illusion. We can ask the critic whether, for my evidence to include *e*, I must at least be *in a position* to know *e*? If so, then the critic's view does not differ radically from E = K. Given E = K, the evidence in my actual possession consists of the propositions which I know, but there is also the evidence in my potential possession, consisting of the propositions which I am in a

position to know. The critic takes my evidence to be the evidence in my potential possession, not just the evidence in my actual possession. To bring out the difference between that view and E = K, suppose that I am in a position to know any one of the propositions $p_1, \ldots, p_n$ without being in a position to know all of them; there is a limit to how many things I can attend to at once. Suppose that in fact I know p_1 and do not know $p_2, \ldots, p_n$. According to E = K, my evidence includes only p_1; according to the critic, it includes $p_1, \ldots, p_n$. Let *q* be a proposition which is highly probable given $p_1, \ldots, p_n$ together, but highly improbable given any proper subset of them; the rest of my evidence is irrelevant to *q*. According to E = K, *q* is highly improbable on my evidence. According to the critic, *q* is highly probable on my evidence. E = K gives the more plausible verdict, because the high probability of *q* depends on an evidence set to which as a whole I have no access.

The contrast with E = K is more radical if the critic allows my evidence to include *e* even when I am not in a position to know *e*. For example, it is perceptually apparent to me that it is snowing; I am not hallucinating; but since I know that I have taken a drug which has a 50 per cent chance of causing me to hallucinate, I am not in a position to know that it is snowing. According to the radical critic, my evidence nevertheless includes the proposition that it is snowing, because it is perceptually apparent to me that it is snowing; thus my evidence is inconsistent with the hypothesis that I am hallucinating and it is not snowing, even though, for all I am in a position to know, that hypothesis is true. According to E = K, my evidence includes at best the proposition that it appear to be snowing. Surely, if I proportion my belief to my evidence, I shall not dismiss the hypothesis that I am hallucinating and it is not snowing. E = K gives the better verdict. Perceptual cases do not show that we sometimes fail to believe our evidence.

A truth does not become evidence merely by being believed, or even by being justifiably believed, as the example of the proposition that draw $n+1$ was red showed above. Nothing short of knowledge will do. But is even knowledge enough?

9.7 KNOWLEDGE AS EVIDENCE

Any restriction on what counts as evidence should be well-motivated by the function of evidence. By sections 9.5 and 9.6, one's evidence includes only propositions which one knows. If, when assessing an hypothesis, one knows something *e* which bears on its truth, should not *e* be part of one's evidence? Would it not violate the total evidence condition to do otherwise? This section examines attempts to justify some further restriction on evidence, and finds them wanting.

One's knowledge is held together by a tangle of evidential interconnections. For example, my knowledge that Henry V died in 1422 is evidentially related to my knowledge that various books say that he died in 1422. Much of one's knowledge is redundant, in the sense that the proposition known is a logical consequence of other known propositions. Perhaps each proposition which I know is redundant in that sense. If all knowledge is evidence, the evidential interconnectedness and redundancy is internal to one's evidence. The redundancy itself is harmless; it does not make the evidence support the wrong hypotheses. The concern is rather that if all one's knowledge is treated as a single body of evidence, its internal evidential interconnections will be obliterated, and therefore that such an account would falsify the nature of our knowledge.

The alternative, presumably, is for evidence to be self-evident, consisting of epistemically self-sufficient nuggets of information. That is an implausibly atomistic picture of evidence, but it constitutes a challenge to explain how there can be evidential interconnections within a single body of evidence. Section 9.2 provides the basis for such an explanation. According to EV, when e is evidence for an hypothesis h for one, one's evidence includes e, and e raises the probability of h, which requires the probability of e on the relevant distribution to be less than 1. Thus EV already permits one proposition in one's evidence to be evidence for another in a non-trivial way. The internal evidential interconnections are not obliterated.

If all knowledge is evidence, then EV in section 9.2 does have the effect of making evidential interconnections within one's knowledge symmetric. For $P(p \mid q) > P(p)$ if and only if $P(p \wedge q) > P(p)P(q)$; since the latter condition is symmetric in p and q, $P(p \mid q) > P(p)$ if and only if $P(q \mid p) > P(q)$. Thus, given that S's evidence includes both p and q, p is evidence for q for S if and only if q is evidence for p for S by EV. Consequently, given that one knows p and q and that all knowledge is evidence, EV implies that if p is evidence for q for one then q is evidence for p for one. We could avoid this result by modifying EV. For example, we could stipulate that e is evidence for h for S only if S's belief in e does not essentially depend on inference from h. But it might be neater to retain EV unmodified and say that e is *independent* evidence for h for S only if S's belief in e does not essentially depend on inference from h. Since the focus of this discussion is not on the evidence-for relation, we shall not pursue these options further.

The claim that all knowledge is evidence faces another sort of objection. Very little (if any) of what we know is indubitable. Therefore, if all knowledge is evidence, much of our evidence is dubitable. We are uneasy with the idea of uncertain evidence. Is this just the old Cartesian prejudice that only unshakable foundations will do? The worry cannot be so easily dismissed. It takes a particularly sharp form in a Bayesian context. The standard way of accommodating new evidence e is by conditionalizing on it. The new unconditional probability of a proposition is its old probability conditional on e (where the old probability of e was non-zero); $P_{new}(h) = P_{old}(h \mid e)$. In particular, $P_{new}(e) = P_{old}(e \mid e) = 1$. These probability distributions should be distinguished from P in EV above, for both P_{old} and P_{new} are supposed to incorporate all of one's evidence at the relevant times; whereas it was observed that P must incorporate only a proper part of one's evidence. Now if the old probability of h was 1, so is its new probability; for if $P_{old}(h) = 1$ then $P_{old}(h \wedge e) = P_{old}(e)$. Since the new probability of e is 1, it will remain 1 under any series of conditionalizations on further propositions. Thus once a proposition is conditionalized on as evidence, it acquires probability 1, and retains it no matter what further evidence is conditionalized on. But most of our knowledge has no such status. Further evidence could undermine it.[12]

Here is an example. I put exactly one red ball and one black ball into an empty bag, and will make draws with replacement. Let h be the proposition that I put a black ball into the bag, and e the proposition that the first ten thousand draws are all red. I know h by a standard combination of perception and memory, because I saw that the ball was black as I put it into the bag a moment ago. Nevertheless, if after ten thousand draws I learn e, I shall have ceased to know h, because the evidence which I shall then have will make it too likely that I was somehow confused about the colours of the balls. Of course, what I know now is true, and so will never be discovered to be false, but it does not follow that there will never be misleading future evidence against it. My present knowledge is consistent with e; on simple assumptions, e has a probability of $1/2^{10,000}$ on my present evidence. If I subsequently learn e, the probability of h on that future evidence will be less than 1. But if conditionalization on subsequent evidence will give h a probability less than 1, then the present probability of h is less than 1, so h is not part of my present evidence. The problem is general: if misleading future evidence of positive probability can undermine my knowledge that I put a

[12] The use of Jeffrey conditionalization to avoid the problem is criticized in section 10.2.

black ball into the bag, it can undermine most of my present knowledge. It looks as though *h* should count as part of my present evidence, and therefore receive probability 1, only if *h* is bound to be a rational belief for me in the future, come what may. Few propositions will pass that test. Indeed, not even *e* passes the test, for later evidence may make it rational for me to believe that I had misremembered the outcome of the first ten thousand draws; several eyewitnesses may insist that I was misremembering; but since uncertainty about *e* does not make *h* certain, it does not rehabilitate *h* as evidence. By this line of argument, either we know very little, or very little of our knowledge is evidence.

What empirical propositions qualify as evidence by the proposed test that their probability should never subsequently slip below 1? One might suppose that the best candidates would be propositions about the present—traditionally, propositions about the subject's present mental states. Since the test requires evidence to remain certain as time passes, in order for a proposition about the present to be evidence, it must remain certain long after the time it is about has passed. But even if it is absolutely certain for me today that I seem to see a blue patch, it will not be absolutely certain for me tomorrow that I seemed to see a blue patch today; my memories will not be beyond question. It would only exacerbate the problem to individuate propositions so that the present-tensed sentence 'I seem to see a blue patch' expressed the same proposition at different times, for even if such a proposition is certain and so true now, it will be false and so uncertain in the future. It is hard to see what empirical propositions would qualify as evidence by the proposed test. Thus the very possibility of learning from experience is threatened.

The model assumes that probabilities change only by conditionalization on new evidence. This is to assume that evidence can be added but not subtracted over time. The assumption is obviously false in practice, because we sometimes forget. But even if the model is applied to elephants, idealized subjects who never forget, the assumption that evidence cannot be lost is implausible. On any reasonable theory of evidence, an empirical proposition which now counts as evidence can subsequently lose its status as evidence without any forgetting, if future evidence casts sufficient doubt on it. Given E = K, this process is the undermining of knowledge. The next chapter develops a more liberal model within a broadly Bayesian framework in which evidence can be lost as well as gained. If today's evidence is not evidence tomorrow, its probability tomorrow can be less than 1. The requirement that the probability of present evidence should never slip below 1 in the future was just an artefact of an overly restrictive model of updating.

One could have a model of the same structure on which only knowledge of the present, or observational knowledge, counts as evidence.[13] But once it is recognized that evidence is not obliged to meet unusual standards of certainty, such restrictions on evidence look ad hoc. Although knowledge of the present or observational knowledge may be easier to obtain than some other kinds of knowledge, that is no reason against counting other kinds of knowledge as evidence, when we obtain them. For example, our evidence for a mathematical conjecture may consist of mathematical knowledge. If we believe that we know *p*, we shall be disposed to use *p* in the ways in which we use evidence. If our belief is true, we are right to use *p* in those ways. It does not matter what kind of proposition *p* is; as Austin said, '*any* kind of statement could state evidence for *any* other kind, if the circumstances were appropriate' (1962: 116). All knowledge is evidence.

9.8 NON-PRAGMATIC JUSTIFICATION

The present case for E = K is now complete. If evidence is what justifies belief, then knowledge is what justifies belief. But is all justified belief justified by evidence? Why cannot experience itself, or practical utility, justify belief? Why cannot belief sometimes *be* justified without being justified *by* anything at all?

The *pragmatic* justification of belief need not be by evidence. Without any evidence at all, someone believes that her child somehow survived an air crash, and will one day return to her. The belief is the only thing which keeps her going; without it, she would kill herself. Perhaps it is on balance a good thing that she has the belief, and in that sense the belief is justified. But this is not the sense of 'justified' in which justified belief appeared to have marginalized knowledge within epistemology. Could belief be *epistemically* justified except by evidence? Epistemic justification aims at truth in a sense—admittedly hard to define—in which pragmatic justification does not. It is far from obvious that any belief is justified in the truth-directed sense without being justified by evidence. It appears otherwise when evidence is conceived too

[13] For Maher, '*E* is evidence iff *E* is known directly by experience' (1996: 158; he relativizes 'directly' to a set of propositions, 160–1). On his view, if I know *e* (e.g. that a substance s dissolved when placed in water), and deduce *h* (e.g. that s is soluble), thereby coming to know *h*, then *e* but not *h* is evidence for me (1996: 158). On the present view, both *e* and *h* are evidence for me, but *h* is not independent evidence for *e*, since in the circumstances I believe *e* by inference from *h*. If someone now tells me that s is salt, *h* may become independent evidence for *e*.

narrowly, for then the evidence looks too scanty to justify all the beliefs which are in fact justified. But if anything we know can be evidence to anchor a chain of justification, as E = K implies, then evidence plausibly suffices for all truth-directed justification. An epistemically justified belief which falls short of knowledge must be epistemically justified by something; whatever justifies it is evidence. An epistemically justified belief which does not fall short of knowledge is itself evidence, by E = K. If we are aiming at the truth, we should proportion our belief to the evidence.

E = K supports the plausible equation of truth-directed justification with justification by evidence, and therefore with justification by knowledge. On this view, if truth-directed justification is central to epistemology, so too is knowledge.

We can suggest something more radical. Belief does not aim merely at truth; it aims at knowledge. The more it is justified by knowledge, the closer it comes to knowledge itself. If evidence and knowledge are one, then the more a belief is justified by evidence, the closer it comes to its aim. The next two chapters will help to make those suggestions good.

10
Evidential Probability

10.1 VAGUE PROBABILITY

When we give evidence for our theories, the propositions which we cite as evidence are themselves uncertain. Probabilistic theories of evidence have notorious difficulty in accommodating that obvious fact, as section 9.7 noted. This chapter embeds the fact in a probabilistic theory of evidence. The analysis of uncertainty leads naturally to a simple theory of higher-order probabilities. The first step is to focus on the relevant notion of probability.

Given a scientific hypothesis *h*, we can intelligibly ask: how probable is *h* on present evidence? We are asking how much the evidence tells for or against the hypothesis. We are not asking what objective physical chance or frequency of truth *h* has. A proposed law of nature may be quite improbable on present evidence even though its objective chance of truth is 1. That is quite consistent with the obvious point that the evidence bearing on *h* may include evidence about objective chances or frequencies. Equally, in asking how probable *h* is on present evidence, we are not asking about anyone's actual degree of belief in *h*. Present evidence may tell strongly against *h*, even though everyone is irrationally certain of *h*. We will refer to degrees of belief as *credences*; for example, one's prior credence in the proposition that the fair coin will come up heads is normally 1/2; thus credences are *not* the degrees of outright belief discussed in section 4.4.

Is the probability of *h* on our evidence the credence which a perfectly rational being with our evidence would give to *h*? That suggestion comes closer to what is intended, but not close enough. It fails in the way in which counterfactual analyses usually fail, by ignoring side-effects of the conditional's antecedent on the truth-value of the analysandum (Shope 1978). For example, to say that the hypothesis that there are no perfectly rational beings is very probable on our evidence is not to say that a perfectly rational being with our evidence would be very confident that there were no perfectly rational beings. To make the point more carefully, let *p* be a logical truth (a proposition expressed by a logically true sentence) such that in this imperfect world it is very