The Analysis of Knowledge

First published Tue Feb 6, 2001; substantive revision Mon Jan 16, 2006

The objective of the analysis of knowledge is to state the conditions that are individually necessary and jointly sufficient for propositional knowledge: knowledge that such-and-such is the case. Propositional knowledge must be distinguished from two other kinds of knowledge that fall outside the scope of the analysis: knowing a place or a person, and knowing how to do something. The concept to be analyzed -- the analysandum -- is commonly expressed using the schema "*S* knows that *p*", where "*S*" refers to the knowing subject, and "*p*" to the proposition that is known. A proposed analysis consists of a statement of the following form: *S* knows that *p* if and only if -- . The blank is to be replaced by the analysans: a list of conditions that are individually necessary and jointly sufficient. To test whether a proposed analysis is correct, we must ask (a) whether every possible case in which the conditions listed in the analysans are met is a case in which *S* knows that *p*, and (b) whether every possible case in which *S* knows that *p* is a case in which each of these conditions is met. When we ask (a), we wish to find out whether the proposed analysans is sufficient for *S*'s knowing that *p*; when we ask (b), we wish to determine whether each of the conditions listed in the analysans is necessary.

---

1. Knowledge as Justified True Belief

According to the following analysis, which is usually referred to as the "JTB" account, knowledge is justified true belief.

The JTB Analysis of Knowledge:
S knows that *p* iff

   i.   *p* is true;
  ii.   *S* believes that *p*;
 iii.   *S* is justified in believing that *p*.

Condition (i), the truth condition, has not generated any significant degree of discussion. It is overwhelmingly clear that what is false cannot be known. For example, it is false that G. E. Moore is the author of *Sense and Sensibilia*. Since it is false, it is not the sort of thing anybody can know.

Although the truth-condition enjoys nearly universal consent, let us nevertheless consider at least one objection to it. According to this objection, Newtonian Physics is part of our overall scientific knowledge. But Newtonian Physics is false. So it's possible to know something false after all.[1]

In response, let us say that Newtonian physics involves a set of laws of nature $\{L_1, L_2,\ldots, L_n\}$. When we say we know Newtonian physics, this could be interpreted as saying we know that, according to Newtonian physics, $L_1, L_2,\ldots, L_n$ are all true. And that claim is of course true.

Additionally, we can distinguish between two theories, T and T^*, where T is Newtonian physics and T^* updated theoretical physics at the cutting edge. T^* does not literally include T as a part, but absorbs T by virtue of explaining in which way T is useful for understanding the world, what assumptions T is based on, where T fails, and how T must be corrected to describe the world accurately. So we could say that, since we know T^*, we know Newtonian physics in the sense that we know how Newtonian physics helps us understand the world and where and how Newtonian physics fails.

1.1 The Belief Condition

Unlike the truth condition, condition (ii), the belief condition, has generated at least some discussion. Although initially it might seems obvious that knowing that p requires believing that p, some philosophers have argued that knowledge without belief is indeed possible. Suppose Walter comes home after work to find out that his house has burned down. He utters the words "I don't believe it." Critics of the belief condition might argue that Walter knows that his house has burned down (he sees that it has), but, as his words indicate, he does not believe it. Therefore, there is knowledge without belief. To this objection, there is an effective reply. What Walter wishes to convey by saying "I don't believe it" is not that he really does not believe what he sees with his own eyes, but rather that he finds it hard to come to terms with what he sees.

A more serious counterexample has been suggested by Colin Radford (1966). Suppose Albert is quizzed on English history. One of the questions is: "When did Queen Elizabeth die?" Albert doesn't think he knows, but answers the question correctly. Moreover, he gives correct answers to many other questions to which he didn't think he knew the answer. Let us focus on Albert's answer to the question about Elizabeth:

(E) Elizabeth died in 1603.

Radford makes the following two claims about this example:

a. Albert does not believe (E). Reason: He thinks he doesn't know the answer to the question. He doesn't trust his answer because he takes it to be a mere guess.
b. Albert knows (E). Reason: His answer is not at all just a lucky guess. The fact that he answers most of the questions correctly indicates that he has actually learned, and never forgotten, the basic facts of English history.

Since he takes (a) and (b) to be true, Radford would argue that knowledge without belief is indeed possible. But Radford's example is not compelling. Those who think that belief is necessary for knowledge

could reply that the example does not qualify as a case of knowledge without belief because it isn't a case of knowledge to begin with. Albert doesn't know (E) because he has no justification for believing (E). If he were to believe (E), his belief would be unjustified. This reply anticipates what we have not yet discussed: the necessity of the justification condition. Let us first discuss why friends of JTB hold that knowledge requires justification, and then discuss in greater detail why they would not accept Radford's alleged counterexample.

1.2 The Justification Condition

Why is condition (iii) necessary? Why not say that knowledge is true belief? The standard answer is that to identify knowledge with true belief would be implausible because a belief that is true just because of luck does not qualify as knowledge. Beliefs that are lacking justification are false more often than not. However, on occasion, such beliefs happen to be true. Suppose William takes a medication that has the following side effect: it causes him to be overcome with irrational fears. One of his fears is that he has cancer. This fear is so powerful that he starts believing it. Suppose further that, by sheer coincidence, he does have cancer. So his belief is true. Clearly, though, his belief does not amount to knowledge. But why not? Most epistemologists would agree that William does not know because his belief's truth is due to luck (bad luck, in this case). Let us refer to a belief's turning out to be true because of mere luck as *epistemic luck.* It is uncontroversial that knowledge is incompatible with epistemic luck. What, though, is needed to rule out epistemic luck? Advocates of the JTB account would say that what is needed is justification. A true belief, if an instance of knowledge and thus not true because of epistemic luck, must be justified. But what is it for a belief to be justified?[2]

Among the philosophers who favor the JTB approach, we find bewildering disagreement on how this question is to be answered. According to one prominent view, typically referred to as "evidentialism", a belief is justified if, and only if, it fits the subject's evidence.[3] Evidentialists, then, would say that the reason why knowledge is not the same as true belief is that knowledge requires evidence. Opponents of evidentialism would say that evidentialist justification (i.e., having adequate evidence) is not needed to rule out epistemic luck. They would argue that what is needed instead is a suitable relation between the belief and the mental process that brought it about. What we are looking at here is an important disagreement about the nature of knowledge, which will be our main focus further below. In the meantime, we will continue our examination of the JTB analysis.

Let us return to Radford's objection to the belief condition, which we considered above. We are now in a position to discuss further how that objection can be rebutted. Recall that Albert does not take himself to know the answer to the question about the date of Elizabeth's death. He does not because he does not remember having learned the basic facts of British history. Now, it is of course true that he did learn these facts, and is indeed able to recall them. But is this by itself sufficient for knowing them? Philosophers who think that knowledge requires evidence would say that it is not. Albert needs to have evidence for believing that he learned those facts. Until he is quizzed, he has no such evidence. *After* the quiz, when he is told that most of his answers are correct, he does have the requisite evidence. For once he comes to know that he is

able to produce consistently correct answers to the questions he is asked, he has acquired evidence for believing that he must have learned this subject matter at school. This evidence is also evidence for the answers he has given. So at that point, the justification condition is met, and thus (since the other conditions of knowledge are also met) he knows (again) that Elizabeth died in 1603. However, he did not know this before finding out that he must have learned those facts, for at that point his answer to the question lacked justification and thus did not add up to knowledge. Evidentialists would deny, therefore, that Radford has supplied us with a counterexample to the belief condition.[4]

2. The Gettier Problem

In his short 1963 paper, "Is Justified True Belief Knowledge?", Edmund Gettier presented two effective counterexamples to the JTB analysis (Gettier 1963). The second of these goes as follows. Suppose Smith has good evidence for the false proposition

1. Jones owns a Ford.[5]

Suppose further Smith infers from (1) the following three disjunctions:

2. Either Jones owns a Ford or Brown is in Boston.
3. Either Jones owns a Ford or Brown is in Barcelona.
4. Either Jones owns a Ford or Brown is in Brest-Litovsk.

Since (1) entails each of the propositions (2) through (4), and since Smith recognizes these entailments, he is justified in believing each of propositions (2)-(4). Now suppose that, by sheer coincidence, Brown is indeed in Barcelona. Given these assumptions, we may say that Smith, when he believes (3), holds a justified true belief. However, is Smith's belief an instance of knowledge? Since Smith has no evidence whatever as to Brown's whereabouts, and so believes what is true only because of luck, the answer would have to be 'no'. Consequently, the three conditions of the JTB account — truth, belief, and justification — are not sufficient for knowledge.[6] How must the analysis of knowledge be modified to make it immune to cases like the one we just considered? This is what is commonly referred to as the "Gettier problem".

Epistemologists who think that the JTB approach is basically on the right track must choose between two different strategies for solving the Gettier problem. The first is to strengthen the justification condition. This was attempted by Roderick Chisholm.[7] The second strategy is to search for a suitable further condition, a condition that would, so to speak, "degettierize" justified true belief. Let us focus on this second strategy. According to one suggestion, the following fourth condition would do the trick:

(iv) *S*'s belief that *p* is not inferred from any falsehood.[8]

Unfortunately, this proposal is unsuccessful. Since Gettier cases need not involve any inference, there are possible cases of justified true belief in which the subject fails to have knowledge although condition (iv) is met. Suppose, for example, that James, who is relaxing on a bench in a park, observes a dog that, about 8 yards away from him, is chewing on a bone. So he believes

5. There is a dog over there.

Suppose further that what he takes to be a dog is actually a robot dog so perfect that, by vision alone, it could not be distinguished from an actual dog. James does not know that such robot dogs exist. But in fact a Japanese toy manufacturer has recently developed them, and what James sees is a prototype that is used for testing the public's response. Given these assumptions, (5) is of course false. But suppose further that just a few feet away from the robot dog, there is a real dog. Sitting behind a bush, he is concealed from James's view. Given this further assumption, James's belief is true. So once again, what we have before us is a justified true belief that doesn't qualify as an instance of knowledge. Arguably, this belief is directly justified by a visual experience; it is not inferred from any falsehood. But if (5) is indeed a non-inferential belief, then the JTB account, even if supplemented with (iv), gives us the wrong result that James knows (5).

Another case illustrating that clause (iv) won't do the job is the well-known Barn County case (Goldman 1976). Suppose there is a county in the Midwest with the following peculiar feature. The landscape next to the road leading through that county is peppered with barn-facades: structures that from the road look exactly like barns. Observation from any other viewpoint would immediately reveal these structures to be fakes: devices erected for the purpose of fooling unsuspecting motorists into believing in the presence of barns. Suppose Henry is driving along the road that leads through Barn County. Naturally, he will on numerous occasions form a false belief in the presence of a barn. Since Henry has no reason to suspect that he is the victim of organized deception, these beliefs are justified. Now suppose further that, on one of those occasions when he believes there is a barn over there, he happens to be looking at the one and only real barn in the county. This time, his belief is justified and true. But since its truth is the result of luck, it is exceedingly plausible to judge that Henry's belief is not an instance of knowledge. Yet condition (iv) is met in this case. His belief is clearly not the result of any inference from a falsehood. Once again, we see that (iv) does not succeed as a solution to the Gettier problem.

Above, we noted that the role of the justification condition is to ensure that the analysans does not mistakenly identify as knowledge a belief that is true because of epistemic luck. The lesson to be learned from the Gettier problem is that the justification condition by itself cannot ensure this. Even a justified belief, understood as a belief based on good evidence, can be true because of luck. So if a JTB analysis of knowledge is to rule out the full range of cases of epistemic luck, it must be amended with a suitable fourth condition, a condition that succeeds in preventing justified true belief from being "gettiered." Thus amended, the JTB analysis becomes a JTB+ account of knowledge, where the '+' stands for the needed fourth condition.

3. An Alternative Approach: Reliabilism

The analysis of knowledge may be approached by asking the following question: What turns a true belief into knowledge? An uncontroversial answer to this question would be: the sort of thing that effectively prevents a belief from being true as a result of epistemic luck. Controversy begins as soon as this formula is turned into a substantive proposal. According to evidentialism, which endorses the JTB+ conception of

knowledge, the combination of two things accomplishes this goal: evidentialist justification plus degettierization (a condition that prevents a true and justified belief from being "gettiered"). However, according to an alternative approach that has in the last three decades become increasingly popular, what stands in the way of epistemic luck — what turns a true belief into knowledge — is the reliability of the cognitive process that produced the belief. Consider how we acquire knowledge of our physical environment: we do so through sense experience. Sense experiential processes are, at least under normal conditions, highly reliable. There is nothing accidental about the truth of the beliefs these processes produce. Thus beliefs produced by sense experience, if true, should qualify as instances of knowledge. An analogous point could be made for other reliable cognitive processes, such as introspection, memory, and rational intuition. We might, therefore, say that what turns true belief into knowledge is the reliability of our cognitive processes.

This approach — reliabilism, as it is usually called — can be carried out in two different ways. First, there is reliabilism as a theory of justification (J-reliabilism).[9] The most basic version of this view — let's call it 'simple J-reliabilism' — takes knowledge to be justified true belief but, unlike evidentialism, conceives of justification in terms of reliability:

Simple J-Reliabilism:
Part A: *S* knows that *p* iff *S*'s belief that *p* is (i) true and (ii) justified.
Part B: *S* is justified in believing that *p* iff *S*'s belief that *p* was produced by a reliable cognitive process (in a way that degettierizes *S*'s belief).

Second, there is reliabilism as a theory of knowledge (K-reliabilism).[10] According to this approach, knowledge does not require justification. Rather, what it requires (in addition to truth) is reliable belief formation. Let us define this second version of reliabilism thus:

Simple K-Reliabilism:
S knows that *p* if, and only if, *S*'s belief that *p* (i) is true and (ii) was produced by a reliable cognitive process (in a way that degettierizes *S*'s belief).

The degettierization-clauses in parentheses are needed because the Gettier problem is no less of a problem for reliabilism as it is for the JTB approach. We will set this issue aside for now and return to it at the end of this section.

In the following passage, Fred Dretske articulates how K-reliabilism can be motivated:

Those who think knowledge requires something *other than*, or at least *more than*, reliably produced true belief, something (usually) in the way of justification for the belief that one's reliably produced beliefs *are* being reliably produced, have, it seems to me, an obligation to say what benefits this justification is supposed to confer…. Who needs it, and why? If an animal inherits a perfectly reliable belief-generating mechanism, and it also inherits a disposition, everything being equal, to *act* on the basis of the beliefs so generated, what additional benefits are conferred by a justification that the beliefs *are* being produced in some reliable way? If there are no additional benefits, what good is this justification? Why should we insist that no one can have knowledge without it? (Dretske 1989, p. 95)

Further below we will discuss how advocates of the JTB approach might answer Dretske's question. In the meantime, let us focus a bit more on Dretske's account of knowledge. According to Dretske, reliable

cognitive processes convey information, and thus endow not only humans, but (nonhuman) animals as well, with knowledge. He writes:

I wanted a characterization that would at least allow for the possibility that animals (a frog, rat, ape, or my dog) could know things without my having to suppose them capable of the more sophisticated intellectual operations involved in traditional analyses of knowledge. (Dretske 1985, p. 177)

It does indeed seem odd to think of frogs, rats, or dogs as having justified or unjustified beliefs. Yet attributing knowledge to animals is certainly in accord with our ordinary practice of using the word 'knowledge'. So if, with Dretske, we want an account of knowledge that includes animals among the knowing subjects, we might want to abandon the traditional JTB account in favor of K-reliabilism.

Advocates of J-reliabilism take justification, and thus reliable belief formation, to be a necessary condition of knowledge. Advocates of K-reliabilism also take reliable belief formation to be a necessary condition of knowledge, however without saying anything about justification. We might wonder, therefore, whether there is any *substantive* difference between the two views, a difference that goes beyond the mere terminological difference of using vs. not using the word 'justification'. Why not think that J and K-reliabilism actually amount to the same thing?[11]

Simple J-reliabilism and simple K-reliabilism would appear to be extensionally equivalent: whatever is a case of knowledge according to the former is also a case of knowledge according to the latter, and vice versa. This does not mean, however, that there is no important difference between the two views. Suppose B is a belief that, though produced by a reliable faculty or process, is in fact false. About B, K-reliabilism implies one and only one thing: B is not an instance of knowledge. But J-reliabilism implies two things about B: (i) B is not an instance of knowledge; (ii) B is a justified belief. So although the two theories do not differ with regard to which beliefs qualify as instances of knowledge and which do not, they do differ in the following respect: Whereas J-reliabilism yields implications about justification or the lack of it, K-reliabilism does not. This could be viewed as a consideration favoring J-reliabilism. Beliefs that fail to qualify as knowledge can, after all, still exhibit an epistemically desirable quality, namely that of being justified. We might be interested in having an account of this quality even if we do not want to conceive of justification as resulting from the possession of evidence.

According to Dretske, his version of K-reliabilism avoids Gettier problems. He says:

Gettier difficulties … arise for any account of knowledge that makes knowledge a product of some justificatory relationship (having good evidence, excellent reasons, etc.) that *could* relate one to something false…. This is [a] problem for justificational accounts. The problem is evaded in the information-theoretic model, because one can get into an appropriate justificational relationship to something false, but one cannot get into an appropriate informational relationship to something false. (Dretske 1985, p. 179)

However, consider again the case of the barn facades. Henry sees a real barn, and that's why he believes there is a barn near-by. Since the barn he is looking at is an actual barn, it would appear that the perceptual process that causes Henry to believe this does not relate him to anything false. So if perception, on account

of its reliability, normally conveys information, it should do so in this case as well. Alas, it arguably does not. Since Henry would have believed the same had he been situated in front of one of the many barn-facades in the vicinity, we are reluctant to judge that Henry knows there is a barn nearby. There is reason to doubt, therefore, that Dretske's version of K-reliabilism escapes the Gettier problem.

In general terms, since reliable faculties can be just as misleading as a person's evidence, a bare bones reliability condition does little toward solving the Gettier problem. When Henry travels through Barn County, surely his vision works just as well as it would elsewhere. Hence, unless we are told how to gauge reliability relative to the subject's environment, reliabilism offers us no reason to judge that Henry fails to know that there is a barn near-by. Or consider the example of the Japanese toy-dog. When James believes that there is a toy-dog before him, his failure to know this is not due to a sudden deterioration of his vision. Rather, James fails to know because an otherwise reliable faculty, vision, is misleading on this particular occasion. Hence, if reliabilism is to yield the correct outcome about this case, it needs to be amended with a further clause. We need to be told either a principled reason why James's visual faculty fails to be reliable under the circumstances, or else why James fails to know even though his belief is produced by a reliable faculty. Clearly, then, Gettier cases pose as much of a problem for reliabilism as for an evidentialist JTB account. Neither theory, unless amended with a clever degettierization clause, succeeds in stating sufficient conditions of knowledge.[12]

4. Internalism and Externalism

Evidentialists reject both J-reliabilism and K-reliabilism. We will first focus on J-reliabilism and further below discuss why evidentialists reject K-reliabilism as well. Evidentialists reject J-reliabilism because they take justification to be something that is *internal* to the subject. J-reliabilists, on the other hand, take justification to be something that is *external* to the subject.[13]

How are we to understand the difference between the, so to speak, internality and the externality of justification? Let us turn to Roderick Chisholm, one of the chief advocates of internalism. In the third edition of *Theory of Knowledge*, Chisholm says the following:

If a person *S* is internally justified in believing a certain thing, then this may be something he can know just by reflecting upon his own state of mind. (Chisholm 1989, p. 7)

In the second edition of this book, he characterizes internalism in a somewhat different way:

We presuppose … that the things we know are justified for us in the following sense: we can know what it is, on any occasion, that constitutes our grounds, or reasons, or evidence for thinking that we know. (Chisholm 1977, p. 17)

These passages differ in the following respect: in the first Chisholm is concerned with the property of justification (a belief's being justified); in the second, with justifiers: the things that *make* justified beliefs justified. What is common to both passages is the constraint Chisholm imposes. In the first passage,

Chisholm characterizes justification as something that is recognizable *on reflection* and, in the second, as the sort of thing that can be known *on any occasion.* Arguably, this is just a terminological difference. It would not be implausible to claim that what can be recognized through reflection is something that can be recognized on any occasion, and what can be recognized on any occasion is something that can be recognized through reflection. Although this point deserves further examination, let us here simply assume that recognizability on reflection and recognizability on any occasion amount to the same thing. In what follows, we will refer to it as *direct recognizability.*

As already noted, in the first passage Chisholm imposes the direct recognizability constraint on justification, in the second on justifiers. Does this amount to a substantive difference? If the direct recognizability of justifiers implies the direct recognizability of justification, and vice versa, then the two passages we considered would indeed just be alternative ways of stating the same point. Whether they really are is perhaps debatable, but here we will simply assume that it makes no substantive difference whether the characterization of internalism focuses on justification or justifiers.

Chisholm, then, defines internalism by saying that justification is recognizable on reflection, and thus in terms of the accessibility of justification. This type of internalism may therefore be called *accessibility internalism.* Alternatively, internalism could be defined in terms of limiting justifiers to mental states. According to this second approach, internalism says that justifiers must be internal to the mind, i.e., must be mental events or states. Internalism thus defined could be labeled *mental state internalism.*[14] Whether accessibility internalism and mental state internalism are genuine alternatives depends on whether being directly recognizable is an essential property of mental states. If it is, then what appear to be genuine alternatives might in fact not be.[15] Since here we cannot go into the details of this issue, we will cut this matter short and simply define internalism, as suggested by Chisholm, in terms of direct recognizability, while acknowledging that it might be preferable to define it by restricting justifiers to mental states. We will refer to internalism as defined here as "J-internalism," since it imposes the direct recognizability constraint not on knowledge but justification.

J-Internalism:
Justification is directly recognizable. At any time *t* at which *S* holds a justified belief *B*, *S* is in a position to know at *t* that *B* is justified.[16]

J-internalism is to be contrasted with J-externalism, which is simply its negation.

J-Externalism:
Justification is not directly recognizable. It is not the case that at any time *t* at which *S* holds a justified belief *B*, *S* is in a position to know at *t* that *B* is justified. (There are times at which *S* holds a justified belief *B* but is not in a position to know that *B* is justified.)

Next, we will discuss what consequences we can derive from J-internalism. To begin with, we can derive the result that Simple J-reliabilism is an externalist theory. Suppose *S*'s belief *B* has, at time *t*, the property of being reliably formed. *B*'s being reliably formed at *t*, and *S*'s being able to recognize at *t* that *B* is reliably formed, are clearly two different affairs. It could be the case that *B* is reliably formed without *S*'s being able to tell at *t* that *B* is reliably formed. According to Simple J-reliabilism, however, reliability by itself —

without the subject's having any evidence indicating its presence — is sufficient for justification. Simple J-reliabilism, therefore, allows for cases of the following kind: *S*'s belief *B* is reliably formed and therefore justified, but, since *B*'s reliability is, so to speak, "hidden" from *S*, *S* cannot directly recognize that *B* is justified. J-reliabilism is, therefore, an externalist theory.

To illustrate this point, let us consider a familiar example due to Laurence BonJour.[17] Suppose Norman is a perfectly reliable clairvoyant. At time *t*, his clairvoyance causes Norman to form the belief that the president is presently in New York. However, Norman has no evidence whatever indicating that he is clairvoyant. Nor has he at *t* any way of recognizing that his belief was caused by his clairvoyance. Norman, then, cannot at *t* recognize that his belief is justified. So Simple J-reliabilism implies that Norman's belief is justified at *t* although Norman cannot recognize at *t* that his belief is justified.

Second, J-internalism allows us to derive the consequence — as it should — that evidentialism is an internalist theory. The question of what a person's evidence consists of is of course not uncontroversial. Nor is it uncontroversial what kind of cognitive access a subject has to her evidence. However, it would not be without a good deal of initial plausibility to make the following two assumptions. First, a subject's evidence consists of her perceptual, introspective, memorial, and intuitional states, as well as her beliefs. In short, a subject's evidence consists of her mental states. Items other than mental states are never part of a subject's evidence.[18] Second, a subject's mental states are directly recognizable to her.[19] If we now add the further assumption (mentioned above) that the direct recognizability of justifiers implies the direct recognizability of justification, then we get the result that evidentialism is a form of J-internalism. Let us display the argument in detail:

Why Evidentialism is a Version of J-Internalism:

1. According to evidentialism, justifiers consist of a person's evidence.
2. A person's evidence (consisting of her mental states) is directly recognizable to that person.
3. Therefore:
   According to evidentialism, a person's justifiers are directly recognizable to that person.
4. If the justifiers that make a person's justified beliefs justified are directly recognizable to that person, then the justification of that person's justified beliefs is directly recognizable to that person.
5. Therefore:
   According to evidentialism, the justification of a person's justified beliefs is directly recognizable to that person.

The crucial premises in this argument are (2) and (4). Evidentialists would be reluctant to call 'evidence' something that is not directly recognizable to a subject.[20] So (2) would appear to be a premise that evidentialists are likely to endorse. And (4) expresses no more than one part of what we already assumed: that the direct recognizability of justifiers implies the direct recognizability of justification, and *vice versa*. Of course, both premises might be challenged. What seems safe to say, therefore, is the conditional point that, if (2) and (4) capture what is essential to evidentialism, then evidentialism implies internalism about justification.

As mentioned at the beginning of this section, evidentialists also reject K-reliabilism. They do so because, pace Dretske, they think that internal justification — justification in the form of having adequate evidence — is *necessary* for knowledge. In other words, they deny that a belief's origin in a reliable cognitive process is sufficient for the belief's being an instance of knowledge. Let us refer to this position as internalism about knowledge, or K-internalism, and let us define it using the concept of *internal justification*: the kind of justification that meets the direct recognizability constraint.

K-Internalism:
Internal justification is a necessary condition of knowledge. A belief's origin in a reliable cognitive process is *not* sufficient for its being an instance of knowledge.

K-externalism is the negation of K-internalism:

K-Externalism:
Internal justification is not a necessary condition of knowledge. A belief's origin in a reliable cognitive process *is* sufficient for its being an instance of knowledge. Consequently, there are cases of knowledge without internal justification.

In this section, we have merely concerned ourselves with what internalists and externalists disagree about with regard to both justification and knowledge. In the next two sections, we will examine what reasons internalists and externalists can cite in support of their respective views.

5. Why Internalism?

First, both J- and K-internalism can be motivated by appealing to evidentialism as a premise. As we saw in the previous section, evidentialism is plausibly construed as entailing internalism. Consequently, reasons in support of evidentialism are also reasons in support of J-internalism. Moreover, evidentialists would say that internal justification is a necessary condition of knowledge. Evidentialists would support this claim with examples. Consider again BonJour's clairvoyant Norman. Norman has no evidence for thinking that he is a reliable clairvoyant. Suppose Norman's belief *B* is caused by his clairvoyance. Suppose further Norman has no independent evidence for *B*. Evidentialists would say that, since due to the lack of evidence *B* is unjustified, *B* is not an instance of knowledge. Considerations supporting evidentialism, then, are also considerations in favor of K-internalism.[21]

Second, there is an argument for internalism that starts with what is known as the deontological conception of epistemic justification:

Deontological Justification:
S is justified in believing that *p* iff in believing that *p*, *S* does not violate any of his epistemic duties.

The concept of duty employed here must not be confused with ethical or prudential duty. The type of duty in question is specifically epistemic.[22] What exactly epistemic duties are is a matter of controversy. A fairly uncontroversial starting point is to say that epistemic duties are those that arise in the pursuit of truth.[23] Thus we might express the concept of deontological justification alternatively as follows: *S* is justified in believing that *p* iff in believing that *p*, *S* does not fail to do what he ought to do in the pursuit of

truth. Of course, this way of putting things leads us directly to a further question: In the pursuit of truth, exactly what is it that one ought to do? Evidentialists would say: It is to believe what, and only what, one's evidence supports.[24]

Let's call proponents of the deontological concept of justification *deontologists*. If deontologists conceive of epistemic duty in the way suggested in the previous paragraph, then they can argue as follows: To be justified is to meet the duty of believing what one's evidence supports. Evidential support is directly recognizable. Therefore, deontological justification is directly recognizable. Hence, deontological justification is internal justification.

There is also an argument from deontology to internalism that does not depend on evidentialism as a premise.[25] It derives the direct recognizability of justification from the premise that what *determines* epistemic duty is directly recognizable.

From Deontology to Internalism:

1. Justification is a matter of epistemic duty fulfillment.
2. Therefore:
   What determines justification is identical to what determines epistemic duty.
3. What determines epistemic duty is directly recognizable.
4. Therefore:
   What determines justification is directly recognizable.
5. If what determines justification is directly recognizable, then justification itself is directly recognizable.
6. Therefore:
   Justification is directly recognizable.

(2) follows directly from the deontological conception of justification. (5) is nothing new; we have assumed it above already. The argument's main premise is of course (3).[26] Though certainly not implausible, this premise is open to criticism. Clearly, then, the argument is not uncontroversial. Nevertheless, it seems fair to say that it represents a straightforward and not obviously implausible derivation of internalism from deontology.

Third, internalism (J or K) can be supported by objecting to particular externalist accounts of justification or knowledge. Let us use reliabilism for the purpose of illustration. Internalists will argue that reliable belief formation is neither necessary nor sufficient for justification, nor sufficient for knowledge when added to true belief. To challenge sufficiency, internalists would cite cases like BonJour's Norman, the unwittingly reliable clairvoyant. Evidentialists would say that the beliefs arising from his clairvoyance (unless supported by adequate evidence) are neither justified nor instances of knowledge. To support the claim that reliable belief production is not necessary for justification, internalists will appeal to the possibility of being deceived by Descartes's evil demon. Let's suppose you are a victim of such deception, and let's distinguish between the normal world and the evil demon world. Your memories, experiences, and beliefs in the evil demon world mirror your memories, experiences, and beliefs in the normal world. However, whereas your beliefs in the normal world are by and large true, by far most of your beliefs in the evil demon world are false and thus unreliably produced. Simple J-reliabilism implies, therefore, that your

beliefs in the evil demon world are unjustified. To internalists, this is an intuitively implausible result. Here is why. Your beliefs in the normal world are (as we may assume) by and large supported by adquate evidence and therefore justified. However, as far as your evidence is concerned, there is no difference between the evil demon world and the normal world. Your beliefs in the evil demon world, internalists would therefore say, are also by and large supported by adequate evidence and therefore justified. Hence internalists would reject the claim that being produced by reliable faculties is a necessary condition of epistemic justification.[27]

6. Why Externalism?

One reason for externalism lies in the attraction of philosophical *naturalism*. According to Gilbert Harman, this view, when applied to ethics, "is the doctrine that moral facts are facts of nature. Naturalism as a general view is the sensible thesis that *all* facts are facts of nature" (Harman 1977, p. 17). What naturalists in ethics want, according to Harman,

is to be able to locate value, justice, right, wrong, and so forth in the world in the way that tables, colors, genes, temperatures, and so on can be located in the world. (Harman 1984, p. 33)

According to this conception of naturalism, a naturalist in epistemology wants to be able to locate such things as knowledge, justification, certainty, or probability "in the world in the way that tables, colors, genes, temperatures, and so on can be located in the world." How, though, are naturalists to accomplish this? According to one answer to this question, they can accomplish this by identifying the non-epistemic grounds on which epistemic phenomena supervene. Alvin Goldman describes this desideratum as follows:

The term 'justified', I presume, is an evaluative term, a term of appraisal. Any correct definition or synonym of it would also feature evaluative terms. I assume that such definitions or synonyms might be given, but I am not interested in them. I want a set of *substantive* conditions that specify when a belief is justified … I want a theory of justified belief to specify in non-epistemic terms when a belief is justified. (Goldman 1979, p. 1)

However, internalists need not deny that epistemic phenomena supervene on non-epistemic grounds, and that it is the task of epistemology to reveal these grounds. It is doubtful, therefore, that the goal of locating epistemic value in the natural world establishes a link between philosophical naturalism and externalism.[28]

According to a second approach, the way to locate epistemic value in the natural world is to employ the methods of the natural sciences.[29] Appealing to this methodological constraint, externalists might argue that, because the study of justification and knowledge is an empirical study, justification and knowledge cannot be what internalists take it to be, but rather must be identified with reliable belief production: a phenomenon that can be studied empirically. It is far from clear, however, that the fundamental questions of epistemology can be answered by employing the methods of the natural sciences. For example, can empirical sciences solve the Gettier problem? Can they answer the question of whether knowledge requires evidence? Can they tell us whether the beliefs of evil demon victims are justified, or whether BonJour's Norman can acquire knowledge on account of his clairvoyance even though is he as no reason to suppose

that he is in possession of such a faculty? Indeed, is the question of whether epistemology can be done solely by employing empirical science a question that can be answered by empirical science itself? It is not easy to imagine that these questions should be answered affirmatively. But if the methodological constraint in question cannot be sustained with complete generality, then this constraint offers us no compelling reason to think that justification and knowledge are the sort of thing that can only be studied empirically, and thus cannot be what internalist take them to be.

A second reason for externalism (more specifically, J-externalism) has to do with the connection between justification and truth. Internalists conceive of a justified belief as a belief that, relative to the subject's evidence or reasons, is likely to be true. However, such likelihood of truth is compatible with the belief's actual falsity. Indeed, likelihood of truth as internalists conceive of it can be exemplified in the evil demon world, in which your justified beliefs about the world are mostly false. Hence externalists view the connection between internalist justification and truth as being too thin and therefore demand a stronger kind of likelihood of truth.[30] Reliability is usually taken to fill the bill.[31] William Alston, for example, has argued that, without a reliability constraint, the connection between justification and truth becomes too tenuous.[32] He argues that only reliably formed beliefs can be justified, and defines a reliable belief-producing mechanism as one that "would yield mostly true beliefs in a sufficiently large and varied run of employments in situations of the sorts we typically encounter" (Alston 1993, p. 9). Suppose we endorse this conception of justification. Let's suppose further that most of our beliefs are justified. It then follows that most of the beliefs we form in ordinary circumstances would have to be true most of the time. Such a belief system could still be brought about by an evil demon. However, it would not be a belief-system consisting of mostly false beliefs, and thus the evil demon responsible for it wouldn't be quite as evil as he could be. So what Alston-type justification rules out is this: a belief system of mostly justified beliefs that is generated by an evil demon who sees to it that most of our beliefs are false. This, then, is the benefit we can secure when, as externalists suggest, we make reliability a necessary element of justification.

Internalists would object that a strong link between justification and truth runs afoul of the rather forceful intuition that the beliefs of an evil demon victim are justified even when they are mostly false. In response, externalists might concede that the sort of justification internalists have in mind and attribute to evil demon victims is a legitimate concept, but question the epistemological relevance of that concept. Of what *epistemic* value (of what value to the acquisition of knowledge), they might ask, is internal justification if it is the sort of thing an evil demon victim can enjoy, a person whose belief system is massively marred by falsehood? Internalists would reply that internal justification should not be expected to supply us with a guarantee of truth, and that its value derives (at least in part) from the fact that internal justification is necessary for knowledge.

A third reason for externalism has to do with Dretske's question about justification: "Who needs it, and why?" Dretske would say, of course, that nobody needs it (for the acquisition of knowledge, that is) because reliable belief production is sufficient for turning true belief into knowledge. With this, internalists disagree.[33] As we have seen, they take the existence of examples like BonJour's clairvoyant Norman as a decisive reason to reject this sufficiency claim. Internalists, therefore, would answer Dretske's question

thus: Those who wish to enjoy knowledge need justification, and they need it because one does not know that p unless one has adequate evidence for believing that p.

In reply to this, Dretske might repeat a point — one that amounts to a fourth reason for externalism — from the passage we considered above: he takes animals such as frogs, rats, apes, and dogs to have knowledge. This is surely in line with the way we ordinarily use the concept of knowledge. The owner of a pet who does not attribute knowledge to it would be hard to find. But are animals capable of the sophisticated mental operations required by beings who enjoy the sort of justification internalists have in mind? It would seem not.[34]

7. Two Analyses of Knowledge

K-internalism and K-externalism, then, are supported by conflicting intuitions. On the one hand, there are examples like BonJour's clairvoyant Norman, examples that strongly suggest that internal justification *is* necessary for knowledge. On the other hand, there is Dretske's point that knowledge is enjoyed by not only humans but animals as well. This strongly suggests that internal justification is *not* necessary for knowledge. Both of these thoughts are inherently plausible. Might it be possible to reconcile them? If animals could have the sort of justification internalists have in mind, internalism would be compatible with animal knowledge. Certainly, animals have sensory experiences, just as humans do. Some internalists think that sensory experiences, in and by themselves, constitute evidence. Such internalists might not shy away from attributing internal justification and therefore knowledge to animals. Other internalists, however, think that S's sensory experiences constitute evidence only if S can coherently view them as a reliable guide to truth. That, it would seem, is a condition animals can't meet.

Suppose animals are not the sort of beings that can have internally justified or unjustified beliefs. If so, we get two alternative and irreconcilable analyses of knowledge: one internalist, the other externalist. Let us state a gloss of the respective analyses. In these analyses, the term "internal justification" stands for the kind of concept internalists have in mind, and the term "external justification" for the kind of concept externalists employ.

External Knowledge (EK):
S knows that p iff

    i. p is true;
   ii. S believes that p;
  iii. S is externally justified in believing that p (in a way that degettierizes S's belief).

Internal Knowledge (IK):
S knows that p iff

    i. p is true;
   ii. S believes that p;
  iii. S is internally justified in believing that p;
  iv. S's belief that p is degettiered.

EK and IK agree and differ in the following respects:

a. According to both EK and IK, knowledge requires true belief. The question each of these analyses is intended to answer is: what do we need to add to true belief to get knowledge?

b. According to both, whether or not one knows is an external matter. K-internalists acknowledge the externality of knowledge for two reasons. The first is that knowledge requires truth; the second is that knowledge requires degettierization. Let us consider each of these reasons in turn.

First, consider an evil demon victim's false belief that he has hands. By the victim's own lights, it certainly looks as though he has hands. Surely, the victim would take himself to *know* that he has hands. Since he has no hands, he is mistaken in thinking he knows he has hands. His failure to know, however, is not directly recognizable to him. For unless his evidential situation were to change radically, no amount of reflection will enable him to figure out that he has no hands. So because of the truth condition, it is not always directly recognizable whether or not one knows. Knowledge, therefore, is essentially external.

Second, let us examine why degettierization is an external matter. Call the condition needed to rule out Gettier-cases the 'G-condition'. If the G-condition is met, then *S* is *not* in a Gettier situation. If the G-condition is not met, then *S is* in a Gettier situation. Whether or not the G-condition is met might not be directly recognizable to *S*, just as whether or not *S*'s beliefs are reliably produced might not be directly recognizable to *S*'s. For example, BonJour's Norman has a faculty (his clairvoyance) whose reliability is hidden from him. On reflection, Norman cannot tell that that he is a reliable clairvoyant. (Of course, future experiences might reveal this to him.) Similarly, evil demon victims cannot through reflection figure out that their perceptual faculties are unreliable. Likewise, a subject who is in a Gettier situation cannot directly recognize — find out through reflection alone — that he is. Consider Henry in Barn County: that there is an abundance of barn facades in the area is a feature of his situation that is (at least for the time being) hidden from him. Therefore, it's not directly recognizable to him that he is in a Gettier situation. This point can be generalized. It is an essential aspect of the G-condition that, when it is not met, the subject is not in a position to recognize this directly. Hence degettierization, and thereore knowledge, are essentially external.

c. IK requires internal justification, EK does not. That is the one condition where the two analyses differ. As a result of this difference, EK includes within the scope of knowledge animals, but fails to accommodate the intuition underlying BonJour's case of clairvoyant Norman and other cases like that. IK, on the other hand, does accommodate this intuition, but — counter-intuitively, as K-externalists would say — excludes animals from the range of subjects that can have knowledge.

If the internalism/externalism controversy is seen as essentially a controversy over the nature of *justification*, then the debate over J-internalism vs. J-externalism would appear to be a case of talking past each other. J-internalists and J-externalists simply intend justification to achieve different things. They each operate with a different concept of justification. J-externalists take justification to be the sort of thing that turns true belief into knowledge, and they view the Gettier problem merely as the problem of adding the right sort of bells and whistles to the justification-condition. J-internalists, on the other hand, cannot view degettierization as something that can, in the form of a suitable clause, be tacked on to the justification condition, for degettierization is an external matter. Rather, internalists take justification to be the sort of thing that turns true *and* degettiered belief into knowledge. Since J-internalists and J-externalists assign different roles to justification, what they ultimately disagree about is not the nature of justification, but the sort of thing in relation to which the theoretical role of epistemic justification is fixed: *knowledge*. Internalists assign justification the role of turning true and degettiered belief into knowledge because they think that internal justification is necessary for knowledge. In contrast, externalists (J-externalists, that is) assign a different role to justification — that of turning true belief into knowledge — because they think

16

that internal justification is *not* necessary for knowledge. It is this difference in their respective views on the nature of knowledge that leads to different views on the nature of justification.

Thus we are confronted with a fundamental disagreement about the nature of knowledge. Externalists such as Dretske would say that the desideratum of making knowledge a natural phenomenon that is instantiated equally by humans and animals must trump the demand that knowledge require the possession of justification in the form of adequate evidence. Externalists of that persuasion would have to say, therefore, that Norman, the unwitting clairvoyant, has knowledge just as much as a mouse that knows where to look for the cheese. Internalists would argue the other way around. To them, Norman-type cases establish the necessity of adequate evidence. And so they would say that, just as Norman's reliable clairvoyance (by itself, in the absence of evidence) does not give him knowledge, a mouse's reliable cognitive mechanisms do not give it knowledge of where to look for the cheese. Externalists would say that it merely seems to us that Norman lacks knowledge when in fact he has it. Internalists would say that it merely seems to us that animals know when in fact they do not.

It might be a mistake to expect that there is a decisive argument that settles the dispute between internalists and externalists one way or another. One way to respond to the intracatability of the debate is to acknowledge that there simply is not *one* concept of knowledge for which there is an analysis that has any chance of meeting with broad assent. Rather, we might conclude that, when we use the word "knowledge", we have sometimes one concept and at other times another concept in mind. If we take this approach, we can distinguish beween *animal knowledge* and *reflective knowledge*. The former, we might say, is reliably formed true belief (that meets a suitable Gettier-clause built into the reliability condition), and the latter is internally justified true belief (that meets a suitable, separate Gettier-condition). Whereas the former kind of knowledge can be shared by animals and humans alike, the latter kind is available only to beings who are capable of intellectual reflection.[35]

To sum up, if we attempt to articulate an analysis of knowledge, we must find answers to the following questions:

- How can the analysis of knowledge be made immune to Gettier cases?
- Does knowledge require justification?
- If it does, is the nature of justification internal or external?

As we have seen, how these questions are to be answered is extremely controversial. Most likely, there isn't one single concept of knowledge that permits consensus on what the necessary and sufficient conditions of knowledge are. Rather, it might be that we must distinguish between animal knowledge and reflective knowledge, and that each of these concepts has its own analysis. In addition to the problems we have discussed in this essay, there are further issues that bear, in a broader sense, on the analysis of knowledge. One of these is:

- What is the extent of our knowledge? Do we know about as much as we think we do?

When we discuss this question, we are confronted with a paradox. On the one hand, there is a seemingly sound argument for the conclusion that we don't even know that we have hands, and thus know much less than we are inclined to think. On the other hand, we are convinced that we do know that we have hands. If this conviction is right, the argument can't be sound after all. The following, supplementary chapter discusses the issues that arise when we try to solve this paradox and examines how they bear on our understanding of the concept of knowledge.

Supplement to <u>The Analysis of Knowledge</u>

Knowledge and Skepticism

---

1. The Skeptical Paradox

Skeptics claim that we know radically less than we think we do. For example, skeptics might claim that we have next to no knowledge of the past, the future, or other minds. Here we will consider the skeptical claim that we have next to no knowledge of the external world: the world of physical objects that we at least seem to perceive. One well-known argument in support of this claim appeals to the possibility of being a BIV: a brain in a vat. According to the BIV Hypothesis, you are a mere BIV without a normal body. This of course means, among other things, that you don't have hands. The nerve endings of your brain are stimulated in a manner so sophisticated that the perfect illusion of a normal life is generated. Let's distinguish between the *vat world* and (what you take to be) the *normal world*. According to the BIV Hypothesis, your introspective and perceptual experiences, your memories, thoughts, fears, and desires -- in short, the totality of your mental states -- in the vat world mirror those you have in the normal world. Hence, according to the skeptical argument, there is nothing you can appeal to that would give you a reason to think that the vat world is not the actual world. You have no evidence whatever for thinking that you are not a BIV. So you don't know that you are not a BIV. That's one of the premises on which the skeptical conclusion is based.

As already mentioned, in the vat world, you don't have hands. Since you can't distinguish between the normal world and the vat world, you can't distinguish between having and not having hands, and therefore can't tell whether or not you have hands. Thus we get the other premise of the argument: If you don't know that you are not a BIV, then you don't know that you have hands.

The skeptical argument against knowledge of the external world, then, goes as follows:

The BIV Argument:

   (C)   If I don't know that I am not a BIV, then I don't know that I have hands

(~K~B I don't know that I am not a BIV.

Therefore:

 (~KH) I don't know that I have hands.

    The conclusion of the BIV Argument mentions just one example of what you don't know if you don't know that you are not a BIV. For example, you wouldn't know either what time it is, what your present location is, what your gender is, who your parents are, and so forth. So if the BIV Argument is sound, we must conclude that our knowledge of the external world is rather limited.

    The conclusion, ~KH, strikes us as obviously false. The premises, however, seem highly plausible. That's why the argument presents us with a paradox. Responding to this paradox raises a dilemma. We can either accept the premises but must then accept a conclusion that strikes us as crazy. Alternatively, we can deny the conclusion, but must then explain which of the argument's seemingly plausible premises is false.

    Next, we will consider four different responses to the BIV Argument.

2. The Relevant Alternatives Theory (RAT)

    Advocates of RAT reject the first premise of the BIV Argument.[1] The motivation underlying this premise arises from a general principle (let an *alternative* to *p* be any proposition that is incompatible with *p*):

The Elimination of Alternatives Principle
If I know that *p*, and I know that *q* is an alternative to *p*, then I know (or I am at least in a position to know) that *q* is false.[2]

    The Elimination of Alternatives Principle seems to guarantee the success of the BIV Argument. You know that your being a BIV is an alternative to your having hands: If you have hands, then you are not a BIV. If you are a BIV, then you don't have hands. Now, you cannot, it would appear, discriminate between the normal world and the vat world, and thus cannot know that you are not a BIV. But your having hands is incompatible with your being a BIV. Hence, it would seem, if the Elimination of Alternatives Principle is true, you cannot know that you have hands.

    So it looks like the BIV Argument stands or falls with the Elimination of Alternatives Principle. Advocates of RAT suggest, therefore, to get rid of it. As a replacement, they suggest the following modification of it: to know that *p*, you must merely be in a position to know the falsehood of all *relevant* known alternatives to *p*. So according to RAT, the following principle is true:

The Elimination of Relevant Alternatives Principle
If I know that *p*, and I know that *q* is an alternative to *p*, then I know (or I am at least in a position to know) that *q* is false — provided *q* is a relevant alternative to *p*.

A canary is, just like a goldfinch, a little yellow bird. Thus when it comes to a task such as classifying the bird in your yard as a goldfinch, its being a canary is a relevant alternative.[3] A Ferrari is, just like a Lamborghini, a slick-looking, expensive sports car. Hence, when you take yourself to know that the car you are seeing a Lamborghini, its being a Ferrari is a relevant alternative. Proponents of RAT would argue, however, that in neither of these two cases is your being a BIV a relevant alternative. To know that the bird in your yard is a goldfinch, you must know that it's not a canary, but you need not know that you are not a BIV. Likewise, to know that the car you are seeing is a Lamborghini, you must know that it's not a Ferrari, but you need not know that you are not a BIV.

RAT is faced with three serious problems. Consider the proposition 'I have hands'. Is the BIV alternative relevant with regard to that proposition? If not, why not? If it is relevant, then RAT is not an effective response to the BIV Argument, for then we still end up with the conclusion that you don't know that you have hands. So for RAT to succeed, its advocates would have to answer the following questions: What are the criteria of relevance? How can we decide, when it comes to knowing that I have hands, whether my being a BIV is a relative alternative?[4]

A second problem for RAT arises from the Elimination of Relevant Alternatives Principle. Suppose that the proposition 'I am a BIV' is not a relevant alternative to the proposition 'I have hands'. If so, you can know that you have hands without knowing that you are not a BIV. But then we get the possibility of what has become known as an 'abominable conjunction': I know that I have hands, but I do not know that I am not a (handless) BIV.[5] Opponents of RAT would argue that this conjunction is so counter-intuitive that it may be viewed as a *reductio ad absurdum* of any theory that entails it as one of the theory's consequences.

The third problem arises because the Elimination of Alternatives Principle, which RAT rejects, enjoys a rather high degree of plausibility. Unless the theory is backed up with a principled account stating criteria of relevance, the rejection of the principle in favor of the Elimination of Relevant Alternatives Principle seems *ad hoc*, motivated not by an independent rationale but merely by the desire to rebut the BIV Argument.

3. Moore-Type Anti-Skepticism

G. E. Moore pointed out that skeptical arguments can be turned on their head. Moreover, he famously attempted to prove the existence of the external world by presenting his hands and saying 'Here is one hand, and there is another.'[6] Applying this strategy, we could reject the second premise of the BIV Argument as follows:

Moorean Anti-BIV:
(C)   If I don't know that I am not a BIV, then I don't know that I have hands

(KH) I know that I have hands.

Therefore:

(K~B) I know that I am not a BIV.

It seems crazy to deny that we know we have hands. So why not take our knowledge of our hands as a starting point for arguing that, since we know we have hands, we therefore know we are not BIVs?

The problem with Moorean Anti-BIV is that it seems to be a case of question begging. If indeed I have no evidence in support of my belief that I am not a BIV, it is unclear how I can know that I am not a BIV. If we find the first premise plausible, it becomes therefore unclear how I can know that I have hands — for if I were a BIV, then I would not have any hands. How, then, can it be a satisfactory response to the BIV-Argument to simply assume that I have hands, and deduce from this assumption that I know I'm not a BIV? Consider an analogy. Suppose we mean by 'God' an omnipotent, omniscient, and benevolent being that created the world. A well-known argument against the existence of God goes as follows:

The Argument from Evil (AFE):
(E~G) If Evil exists, then God does not exist

(E)   Evil exists.

Therefore:

(~G) God does not exist.

Suppose the Theist responds as follows:

Anti-AFE:
(E~G) If evil exists, then God does not exist

(G)   God exists.

Therefore:

(~E) Evil does not exist.

It seems clear that Anti-AFE begs the question against AFE, and so is not a good response to AFE. There is overwhelming, antecedent evidence in support of E: Evil exists, that much is hardly contestable. As a result, AFE succeeds in casting serious doubt on the existence of God. That's why, in appealing to the existence of God as a premise, Anti-AFE begs the question against AFE.

If Anti-AFE and Moorean Anti-BIV are analogous, then the charge that Moorean Anti-BIV begs the question sticks. Many epistemologists would say that the two arguments are indeed analogous. To the extent E is plausible, G is implausible. Therefore, denying E on the basis of G is illegitimate. A properly

justified rejection of E must not appeal to the existence of God. Likewise, ~K~B enjoys a good deal of antecedent plausibility, thus calling KH into question. Therefore, denying ~K~B on the basis of KH is not a legitimate move but rather a case of question begging. A properly justified rejection of ~K~B must be based on a ground other than KH.[7]

Suppose, on the other hand, good evidence for believing "I'm not a BIV" is actually available. In the next section, we will assume we know that the know-how and technology needed for successful envatment (for making a mere brain think it's a normal person having an ordinary life) does not exist. Arguably, we know this in pretty much the same way in which we know that the easter bunny, the Loch Ness monster, and the abominable snowman do not exist. Such creatures belong, we know, to the realm of fiction just as much as BIVs. If that is correct, there would then be a crucial difference between the BIV Argument and Moorean Anti-BIV on the one hand, and AFE and Anti-AFE on the other hand. For if we have excellent evidence for believing that BIVs don't exist, then the second premise of the BIV Argument would be, unlike the second premise of AFE, implausible. The BIV Argument would then fail to succeed in calling the existence of my hands in question. As a result, I would not beg the question if I employed Moorean Anti-BIV as a response to the BIV Argument.[8]

4. The Ambiguity Response

Let us distinguish between the concept of knowledge and the word 'knowledge'. Should we take it for granted that, when you and I use the word 'knowledge', we *mean* the same, that is, have the same concept of knowledge in mind? According to the Ambiguity Response, we should not take this for granted. Obviously, there are many different concepts of knowledge. Knowledge might be viewed as:

A. true belief;
B. justified true belief;
C. justified, true, and degettiered belief;
D. reliably formed true belief.

When we use the word knowledge in ordinary life, sometimes we have (A) in mind. At other times, we don't attribute knowledge of *p* to a person unless we think that person has a good reason for *p*. Then we have (B) in mind. Those who have taken an epistemology course might replace (B) with (C). And sometimes we attribute knowledge to our pets. Then it would seem we have something like (A) or (D) in mind.

It would appear, then, that what we mean when we use the word 'know' can vary from one situation to another. Hence, when the Skeptic and the Anti-Skeptic disagree on what we know, we must distinguish between two possibilities. First, it might be that they have exactly the same concept of knowledge in mind. In that case, their disagreement is substantive. It can't be that both of them are right. One of them must be mistaken. Second, it might be that they have different concepts of knowledge in mind. If so, their disagreement will be merely verbal. As a result, they might both be right.

According to the ambiguity response, the use of the word 'know' in the BIV-Argument can be interpreted in two different ways. According to the first interpretation, the argument is sound but yields an innocuous conclusion. According to the second interpretation, the BIV Argument has an interesting and indeed disturbing conclusion, but fails to be sound. Let us look at the details of this response.

The conceptual distinction to which the ambiguity response appeals is that between fallibilism and infallibilism. Infallibilism, defined as a gloss applicable to various approaches to the analysis of knowledge, can be defined as follows:

Infallibilism:
S's knowing that *p* requires *S*'s satisfying an evidential or reliability condition C, such that it is *not* possible for *S* to satisfy C while *p* is false.

Infallibilism is the negation of fallibilism:

Fallibilism:
S's knowing that *p* requires satisfying an evidential or reliability condition C, but C is not such that it is impossible for *S* to satisfy C while *p* is false.

Applied to evidentialism, fallibilism is the view that knowledge-giving reasons *need not* be entailing reasons, whereas infallibilism is the view that knowledge-giving reasons *must* be entailing reasons. Equipped with this distinction, we can distinguish between a fallibilist and an infallibilist sense of the word 'know' and thus discriminate between three different versions of the BIV Argument:

V results from giving the word 'know' an infallibilist interpretation in the premises and a fallibilist interpretation in the conclusion;

V results from giving the word 'know' a fallibilist interpretation in both the premises and the conclusion;

V results from giving the word 'know' an infallibilist interpretation in the premises and the conclusion a well.

V1 is an instance of equivocation and thus obviously invalid. In response to V2, ambiguity theorists will concede that the argument has an upsetting conclusion, but claim that the argument is unsound because its second premise, ~K~B, is false. In response to V3, they will agree that the argument is sound. But they will not perceive this as a worrisome outcome, since they will say that the conclusion — I don't have *infallible* knowledge of my hands — is not at all surprising or remarkable. We can gladly admit that the extent of infallible knowledge is extremely limited. What matters is rather the extent of *fallible* knowledge. And as far as that is concerned, the BIV-Argument does not succeed in establishing a negative conclusion because the fallibilist interpretation of the second premise — I don't have fallible knowledge of my not being a BIV — is false.

If we assume that, when presenting the BIV-Argument, the Skeptic has infallible knowledge in mind, whereas the Anti-Skeptic is thinking of fallibilist knowledge, their disagreement will be merely verbal. They will both be right. But that, according to the ambiguity response, is no reason for the Skeptic to

celebrate. For from the point of view of the ambiguity theory, the conclusion that we don't have infallible knowledge of our hands is completely unremarkable, and therefore nothing to worry about. [9]

The ambiguity response invites, among others, the following two objections. First, some would object that the fallilbilist concept of knowledge is incoherent. David Lewis, for example, rejects fallibilism:

If you are a contented fallibilist, I implore you to be honest, be naive, hear it afresh. 'He knows, yet he has not eliminated all possibilities of error.' Even if you've numbed your ears, doesn't this overt, explicit fallibilism still sound wrong?[10]

Second, it is not quite clear how one can have even fallibilist knowledge of one's not being a BIV. Skeptics who employ the BIV Hypothesis would say that there is no difference between my evidence in the normal world and my evidence in the vat-World. Therefore, I have simply no evidence at all in support of the belief that I am not a BIV. And if that is true, I don't even know fallibly that I am not a BIV. Hence, unless the ambiguity response is supplemented with an explanation of why ~K~B is false, it doesn't have much bite. Let us briefly examine how such a supplementary account might go.

5. Fallibilist Anti-Skepticism

Let us agree, then, that we are using the word 'know' in the fallibilist sense. In that sense of the word 'know', S's belief that p can be an instance of knowing that p even if S's evidence for p does not entail p. Now, it would appear that there are all kinds of things the non-existence of which is supported by non-entailing but nevertheless adequate evidence. About such things, we know that they don't exist. For example, you know that there isn't a million dollars hidden inside of your mattress. You know that there isn't a nuclear bomb in your basement. You know that the easter bunny, the Loch Ness Monster, and the abominable snowman don't exist. Likewise, you arguably know that the following does not exist: the know-how and technology needed for turning people into BIVs and making them suffer the delusion of a normal life. An advocate of the ambiguity response could, then, argue as follows:

Fallibilist Anti-BIV:
C* If I know that BIV know-how and technology do not exist, then I know that I am not a BIV

K~ˊ I know that BIV know-how and technology do not exist.

Therefore:

K~ǀI know that I am not a BIV.

Does this argument beg the question against the BIV-Argument? It does not appeal to knowledge of my hands as a premise. Therefore, it certainly does not beg the question in the way in which, according to many epistemologists, Moorean Anti-BIV does. On the other hand, if I claim I know that BIV know-how and technology do not exist, then I implicitly appeal to knowledge of the external world. But knowledge of the external world is just what the BIV Argument ultimately calls into question. Thus, it might be argued, Fallibilist Anti-BIV begs the question after all.

However, whether Fallibilist Anti-BIV is indeed question-begging is far from clear. Consider a pair of arguments: an original argument, A, and a counter-argument, Anti-A, which denies one of the premises of A. Suppose Anti-A uses a premise that presupposes or somehow implies that A's conclusion is false. Is that by itself sufficient to render Anti-A question-begging? Let us suppose it is not.[11] Whether Anti-A is question-begging or not, let's suppose, depends on the *evidence* in support of the premises of each argument. Suppose there is one body of evidence supporting the premises of A, and another body of evidence supporting the premises of Anti-A. In that case, neither arguments begs the question against the other. Rather, the two arguments are simply a reflection of the fact that the total evidence available pulls into two different directions. The BIV Argument and Fallibilist Anti-BIV might be related in just that way. My mental states in the normal world and in the BIV world are identical. That's a reason for thinking I cannot know that I'm not a BIV. On the other hand, I have excellent reasons for thinking that the know-how and technology for turning people into BIVs does not exist. That's a reason for thinking I know after all that I'm not a BIV. Of course, these reasons might not be equally strong. If you think the latter reason is more decisive than the former, then you could conclude that Fallibilist Anti-BIV is indeed an effective rebuttal of the BIV-Argument.

6. The Contextualist Response

Contextualism is closely related to the ambiguity response. Like advocates of the ambiguity response, contextualists hold that what speakers mean when they use the word 'knowledge' is not always the same. However, whereas the ambiguity response says nothing about how the word 'know' acquires its meaning in specific contexts, contextualism focuses on just that question. Call a sentence of the form '*S* knows that *p*' a *knowledge attribution*, and the subject who utters such a sentence the *attributor*. According to contextualism, the meaning of knowledge-attributions (which concept of knowledge the attributor has in mind) is fixed by the attributor's conversational context. If the conversation in which the attributor is engaged brings skeptical alternatives to the fore, then the attributor is in a high-standard context. As a result, the attributor's utterance, '*S* knows that *p*' is true only if *S*'s belief that *p* meets high standards of knowledge. If, on the other hand, skeptical alternatives are not salient at all in the attributor's conversational context, then the attributor is in a low-standard context. In such a context, an utterance such as '*S* knows that *p*' can be true even if *S* meets merely low standards of knowledge.[12]

How are we to think of the distinction between low-standards and high-standards knowledge? The distinction pretty much corresponds to that between fallibilist and infallibilist knowledge. If I have low-standards knowledge of my hands, my epistemic situation allows for knowledge of my hands even though I *could* be in that very situation while not having hands. On the other hand, if I have high-standards knowledge of my hands, my epistemic situation allows for knowledge of my hands only if I *could not* be in that situation while not having hands.

The following pair of claims, then, is constitutive of contextualism: As long as error-possibilities are ignored, the standards of knowledge remain low, and the concept expressed by the word 'knowledge' is that of low-standard knowledge. But when error possibilities become salient, then the standards of

knowledge rise. Thus speakers who are confronted with error possibilities have high-standard knowledge in mind when they utter knowledge attributions. According to contextualists, it is precisely this thought that affords us a satisfactory resolution of the skeptical paradox.

Contextualists point out that a satisfactory response must go beyond merely denying one of the premises of the BIV-Argument. Rather, a good response must explain why this argument makes us *vacillate*, why we find it both compelling and, at the same time, crazy.[13] Contextualism is supposed to provide us with just such an explanation. Here is how that explanation goes: When we consider the BIV Argument, the alternative of being a BIV becomes salient. As a result, our standards of knowledge rise and we have high-standard knowledge in mind when we use the word 'know'. In a situation like that, we speak truly when we say 'I don't know that I'm not a BIV' and 'I don't know that I have hands'. That explains the appeal of the BIV-Argument. On the other hand, when we are in ordinary situations and don't bother to consider skeptical alternatives, our standards of knowledge are low. We then speak truly when we say 'I know I have hands'.[14] That explains why the conclusion of the BIV-Argument strikes us as crazy. The contextualist response to the skeptical paradox, then, achieves two things: it explains the appeal of the skeptical argument and rescues, at the same time, our conviction that we know we have hands.[15]

The contextualist literature has elicited many objections.[16] Let us focus on just two. First, we may wonder whether the meaning of knowledge attributions is really context-sensitive in just the way contextualists assert. According to contextualists, what speakers mean when they use the word 'know' is determined by the salience or non-salience of error-possibilities. When error possibilities are salient, the standards of knowledge rise. As long as error possibilities are ignored, they stay low. Thus, when speakers use the word 'know' in high-standards contexts, they have a demanding concept of knowledge in mind. In that demanding sense of 'know', we know very little. But as long as speakers remain in low-standard contexts, they have a non-demanding concept of knowledge in mind. In that low-standard sense of 'know', we know a lot. But does what we mean by 'know' really change in precisely that way? Obviously, there are exceptions. It seems safe to say that, when philosophers such as G. E. Moore or Roderick Chisholm considered skeptical alternatives, they meant by 'knowledge' exactly what they meant by that word in other, ordinary situations, simply because, as philosophers, they *intended* to use that word in a way consistent with their general views on the nature of knowledge. In fact, contextualists have acknowledged that it is possible to resist the upward pressure created by salient error possibilities. So they have made a concession: It's possible to retain a low-standard meaning of 'knowledge' in situations in which error-possibilities are salient.[17] What follows from the existence of philosophers (and perhaps some non-philosophers as well) whose use of the word 'know' does not vary from one context to another? What follows, we might conclude, is that the the word 'know' is less context-sensitive, and that the phenomenon of vacillation towards skepticism is less common, than contextualists suggest.

Second, we might wonder what benefits the appeal to contexts is expected to bestow. Isn't it the case that, whatever context we are in, we can always exercise the option of semantic ascent and distinguish between a fallibilist and an infallibilist meaning of 'know'?

Consider an example contextualists are fond of using: what we mean by the word 'tall' is obviously sensitive to context. Suppose John's height is 6'6". In an ordinary context, we might say that John is tall. But a basketball coach might say John isn't tall. Clearly there is no contradicition between these utterances. Since the use of 'tall' in an ordinary context and in a basketball coach's context are governed by different standards, both utterances will be true. But such context sensitivity can easily be eliminated. The utterance 'John is a tall person' will be true in either context, and the utterance 'John is not a tall basketball player' will be true in either context. The device that enables us to ascend, so to speak, from context dependence to a higher semantic plane is that of conceptual disambiguation, in this case, the device of making a conceptual distinction between tall persons and tall basketball players.

Likewise, I can employ the ambiguity theory and make a conceptual distinction between fallibilist and infallibilist knowledge. This enables me to say 'I have fallibilist knowledge of my hands' and 'I fallibly know that I'm not a BIV', and to say 'I don't have infallible knowledge of my hands' and 'I don't infallibly know that I'm not a BIV'. The meaning of such utterances is not at all context-sensitive. Arguably, they will be true in *any* context. Why, then, should we think that a satisfactory response to skepticism requires of us to appeal to context?

Contextualists might reply that, when we compare contextualism with the ambiguity response, contextualism is superior in the following respect: it enables us to explain why our reaction to skepticism is that of vacillation, of finding skeptical arguments both crazy and compelling. Ambiguity theorists, however, would say they can explain the vacillation phenomenon without appealing to the context-sensitivity of the word 'know'. Whether we know, even in the fallibilist sense of 'know', that we are not BIVs is not an easy question to decide. My mental states in the vat world mirror my mental states in the normal world. That's a good reason for thinking I can't have any evidence at all for concluding I'm not a BIV, and therefore can't even have fallibilist knowledge of not being a BIV. On the other hand, I do seem to have good reasons for thinking that BIV technology does not exist. If that's true, and if the appeal to such reasons is not question-begging, then I do have fallibilist knowledge of not being a BIV. So whether or not I can know I'm not a BIV is, even when we consider fallible knowledge, a complicated question. That's why it's possible to experience vacillation when trying to resolve the skeptical paradox, and why this paradox is unlikely to be resolved to everybody's satisfaction.[18]

The Value of Knowledge

First published Tue Aug 21, 2007

Value of knowledge has always been a central topic within epistemology. An important question to address, which can be traced right back to Plato's *Meno*, is: what is it about knowledge (if anything) that makes it more valuable than mere true belief? Interest in this topic has re-emerged in recent years, in response to a rediscovery of the *Meno* problem regarding the value of knowledge (e.g., Kvanvig 2003) and in response to a concern that contemporary accounts of knowledge are unable to explain the (putative) distinctive value of knowledge (e.g., Williamson 2000). Moreover, recent discussions of the value of knowledge have begun to explore the possibility that it is not knowledge which is the distinctively valuable epistemic standing, but rather a different epistemic standing altogether, such as understanding.

---

1. The Meno Problem

Much of the debate regarding epistemic value has focussed on the value of knowledge. This is unsurprising, since the epistemological focus, both in the contemporary literature and historically, has almost exclusively been on this notion. If knowledge is not of special value, however, then this focus is somewhat mysterious. We will call the general question of why knowledge is valuable the *value problem*.

The question why knowledge is distinctively valuable has an important historical precedent in Plato's *Meno* in which Socrates raises the question of why knowledge is more valuable than mere true belief. Initially, we might appeal to the fact that knowledge appears to be of more practical use than true belief in order to mark this difference in value, but, as Socrates notes, this claim is far from obvious on closer inspection. After all, a true belief about *the correct way to Larissa* is surely of just as much practical use as *knowledge of the way to Larissa*—both will get us to our destination. Given that we clearly do value knowledge more than mere true belief, the fact that there is no obvious explanation of why this should be so creates a problem. We will call the issue of why knowledge is more valuable than mere true belief, the *Meno problem*.

Plato's own solution to this problem was to argue that the property distinctively possessed by knowledge is that of being 'tied-down' to the truth, like the mythical tethered statues of Daedalus which were so life-

like that they were tied to the ground to ensure that they did not run away. In contrast, mere true belief, argues Plato, is apt to run away and be lost. Put more prosaically, the point being made here is that knowledge, unlike mere true belief, gives one a confidence that is not easily lost, and it is this property that accounts for the distinctive value of knowledge over mere true belief.

For example, if one knows the way to Larissa, rather than merely truly believes that such-and-such is the correct way to go, then one is less likely to be perturbed by the fact that the road, initially at least, seems to be going in the wrong direction. Mere true belief at this point may be lost, since one might lose all confidence that this is the right way to go. In contrast, if one knows that this is the right way to go, then one will be more sanguine in the light of this development, and thus will in all likelihood press on regardless (and thereby have one's confidence rewarded by getting where one needs to go).

Like most commentators, then, Plato responds to the *Meno* problem by trying to find a way to meet it head-on—i.e., by trying to find a way to show that knowledge is of more value than mere true belief after all. He thus aims for a *non-revisionary* response to the problem, and we shall consider others below. Alternatively, of course, one could argue that the way to deal with this problem is simply to reject the intuition in play and argue that knowledge isn't of more value than mere true belief after all. Ideally, one would supplement such an account with an explanation of why knowledge might seem to be more valuable than mere true belief even though in fact it isn't. This would be a *revisionary* response to the problem. While few have found revisionary responses to the *Meno* problem attractive, analogous revisionary responses to other comparable problems have been relatively common, as we will see below.

2. Two Other Value Problems for Knowledge

While much of the focus of the discussion of the value of knowledge has tended to cluster around the *Meno* problem, there are in fact two further related problems in this regard. The first is what we might call—following Duncan Pritchard (2007*a*: §2)—the *secondary value problem*(with the *Meno* problem as the *primary value problem* for knowledge). Whereas the *Meno* problem concerns the question of why knowledge is more valuable than mere true belief, the secondary value problem concerns the issue of why knowledge is more valuable than any proper subset of its parts. That is, why do we specifically desire knowledge rather than any epistemic standing that falls short of knowing (including, but not restricted to, mere true belief)? The importance of this distinction between the two value problems can be brought out by considering a possible response to the primary value problem which is not thereby a response to the secondary value problem.

Suppose, for example, that knowledge is justified true belief plus some additional component that deals with Gettier-style cases. Suppose further, however, that justification adds value to a mere true belief. If this last point is right, then one might reasonably argue that the fact that knowledge entails justification offers a way of dealing with the primary value problem, since there would now be a property of knowledge which mere true belief lacks and which affords greater value to knowledge over mere true belief. It would not follow, however, that we would thereby have a response to the secondary value problem. This is because

justified true belief is a proper subset of knowledge on our present suppositions, and thus the greater value of knowledge over mere true belief would not translate into a greater value of knowledge over any proper subset of its parts, including justified true belief.

If one's account of the value of knowledge ended at this point, one would thus be offering a non-revisionary response to the *Meno* problem while simultaneously offering a revisionary response to the secondary value problem. Indeed, this is, in effect, the line taken by Mark Kaplan (1985), who agues that the moral of the post-Gettier literature is that what is really of epistemic value is justified true belief, and not knowledge (knowledge being justified true belief plus an additional component to rule-out Gettier-style cases). Kaplan's point is that it is of no practical consequence to us whether we have Gettier-proof justified true belief (i.e., knowledge) rather than just justified true belief, and hence there is no specific reason to value knowledge over justified true belief. Moreover, Kaplan can explain why we might ordinarily have the intuition that knowledge is of special epistemic value by noting that knowledge could very easily be confused with mere justified true belief. (For criticism of Kaplan's view in this regard, see Conee 1988).

If seems then that if one wishes to account for the distinctive value of knowledge, one must resolve both the *Meno* and the secondary value problem. Indeed, there may even be a third value problem for knowledge in play here. After all, one could respond to the secondary value problem by arguing that knowledge is more valuable as a matter of degree than that which falls short of knowledge. It is unclear, however, whether this way of thinking about the value of knowledge can do justice to the idea that knowledge is distinctively valuable. That is, the picture that one is left with is one on which knowledge simply marks a point on a continuum of epistemic value, but on this picture it is far from clear why the focus of epistemological theorizing has been *this* point on the continuum rather than some other point (a point just before the one that knowledge marks perhaps, or one just after). Thus, one might argue that what is required is an account of why knowledge is more valuable than that which falls short of knowledge not merely as a matter of degree but of kind, what Pritchard (2007*a*: §2) calls the *tertiary value problem*. In effect, the challenge posed by the tertiary value problem is to explain what special kind of value enters the picture once one gets to the point on the continuum that knowledge marks.

One further point is in order before we continue. It ought to be clear that what we are seeking when we look for a response to one of these three value problems is not an account of why knowledge is *always* more valuable than the corresponding epistemic commodity (e.g., mere true belief). To take the *Meno* problem as an illustration on this score, no-one would surely want to hold that knowledge is always of more (overall) value than mere true belief, since there are bound to be cases in which it would better for you, all things considered, to merely truly believe *p* than to know *p* (as when knowing *p* would kill you, say, but merely truly believing *p* would win you the lottery instead). However, while it is clear that the requirement laid down on successful resolutions of the various value problems for knowledge is weaker than the demand that knowledge is always more valuable than the corresponding epistemic commodity, it isn't at all clear how best to understand this weaker demand. (Note what goes here for the value problems regarding knowledge applies just as equally when it comes to analogous problems that face other epistemic standings).

3. Reliabilism and the Meno *Problem*

The first contemporary wave of work on the value problem largely concerned whether this problem raised a distinctive difficulty for reliabilist accounts of knowledge—i.e., those views which essentially define knowledge in terms of true belief that arises out of reliable belief-forming processes. In particular, the claim was that reliabilism was unable even to offer an answer to the primary value problem.

A fairly clear statement of what is at issue here is given in a number of places by Linda Zagzebski (e.g., 2003; cf. DePaul 1988; 1993; Zagzebski 1996; Jones 1997; Swinburne 1999; 2000; Riggs 2002; Kvanvig 2003; Sosa 2007: ch. 4). To begin with, Zagzebski argues that the reliability of the process by which something is produced does not automatically add value to that thing, and thus that it cannot be assumed that the reliability of the process by which a true belief is produced will add value to that true belief. In defence of this claim, she offers the analogy of a cup of coffee. She claims that a good cup of coffee which is produced by a reliable coffee machine (i.e., one that regularly produces good cups of coffee) is of no more value than an equally good cup of coffee that is produced by an unreliable coffee machine.

Furthermore, Zagzebski claims that true belief is in the relevant respects like coffee: a true belief formed via a reliable belief-forming process is no more valuable than a true belief formed via an unreliable belief-forming process. In both cases, the value of the reliability of the process accrues in virtue of its tendency to produce a certain valuable effect (good coffee/true belief), but this means that where the effect has been produced—where one has a good cup of coffee or a true belief—then the value of the product is no greater for having been produced in a reliable way.

Elsewhere in the literature, this problem has been called the "swamping problem", on account of how the value of true belief 'swamps' the value of the true belief being produced in a reliable (i.e., truth-conducive) way. So expressed, the moral of the problem seems to be that where reliabilists go awry is by treating the value of the process as being solely captured by the reliability of the process—i.e., its tendency to produce the desired effect. Since the value of the effect swamps the value of the reliability of the process by which the effect was achieved, this means that reliabilism has no resources available to it to explain why knowledge is more valuable than true belief.

It's actually not clear that this is a problem that is specific to reliabilism. That is, it seems that if this is a *bona fide* problem then it will affect any account of the value of knowledge which has the same relevant features as reliabilism—i.e., which regards the greater value of knowledge over true belief as instrumental value, where the instrumental value in question is relative to the valuable good of true belief. Presumably, there could be non-reliabilist views that had these features.

Even granting the main elements of the swamping argument, there are moves that the reliabilist can make in response (see, e.g., Goldman & Olsson *forthcoming*). For example, it is surely open to the reliabilist to argue that the greater instrumental value of reliable true belief over mere true belief does not need to be understood purely in terms of instrumental value relative to the good of true belief. There could, for

instance, be all sorts of *practical* benefits of having a reliable true belief which generate instrumental value. (Indeed, it is worth noting that line of response to the *Meno* problem sketched by Plato which we noted above seems to specifically appeal to the greater practical instrumental value of knowledge over mere true belief).

Moreover, there is reason to think that that this objection will only at best have an impact on the most extreme forms of reliabilism—i.e., those views which *simply* say that knowledge is true belief derived from a reliable process. In contrast, more refined versions of reliabilism, such as the sort of agent reliabilist account offered by John Greco (e.g., 1999; 2000), might be thought to be untouched by this sort of argument. This is because, according to agent reliabilism, it is not any sort of reliable process which is knowledge-conducive, but only those processes that are stable features of what Greco calls the agent's "cognitive character". The main motivation for this restriction on reliable processes is that it excludes certain kinds of reliable processes—what Greco calls "strange and fleeting processes"—which notoriously cause problems for the view (such as processes where the reliability is due to some quirk in the subject's environment, rather than because of any cognitive trait possessed by the agent herself). Plausibly, however, one might argue that the reliable traits that make up an agent's cognitive character have some value independently of the instrumental value they possess in virtue of being reliable (i.e., that they have some final or intrinsic value). If this is right, then this opens up the possibility that agent-reliabilists can evade the value problem that Zagzebski identifies for pure reliabilists. (Note, however, that Greco himself does not respond to the value problem in this way**.** His proposal will be considered in this regard in due course).

Zagzebski's diagnosis of what is motivating this problem for reliabilism seems**,** however, explicitly to exclude such a counter-response. She argues that what gives rise to this difficulty is the fact that the reliabilist has signed up to a "machine-product model of belief" (see especially, Zagzebski 2003), where the product is external to the cause. It is not clear what exactly Zagzebski means by this point, but she thinks it shows that even where the reliable process is independently valuable (i.e., independently of its being reliable), it still doesn't follow that the value of the cause will transfer to add value to the effect. Zagzebski again offers the 'coffee' analogy to illustrate this: even if a reliable coffee machine were independently valuable, it would not thereby confer additional value on a good cup of coffee.

Perhaps the best way to understand what Zagzebski has in mind here is to consider what she thinks *is* required in order to resolve this problem. She argues that what is needed is an 'internal' connection between product and cause, such as the kind of internal connection that exists between an act and its motive which is highlighted by how we explicitly evaluate actions in terms of the motives that led to them. On this picture, then, we are not to understand knowledge as a state consisting of a known belief, but rather as a state which consists of both the true belief *and* the source from which that true belief was acquired. In short, then, the problem with the machine-product model of belief that Zagzebski claims to identify is that it leads us to evaluate the state of the knowledge independently of the means by which the knowledge was acquired. If, in contrast, we have a conception of knowledge that incorporates into the very state of knowledge the way that the knowledge was acquired, then, Zagzebski argues**,** we can avoid this problem.

Zagzebski's contention is that once one effects this transition away from the machine-product model of belief, one can allow that the independent value of the reliable process can ensure that knowledge, by being produced in this way, is more valuable than mere true belief. In particular, if the process by which one gained the true belief is an epistemic virtue—a character trait which Zagzebski thinks is both reliable and intrinsically valuable—then this can ensure that the value of the knowing state in this case is more valuable than any corresponding state which simply consisted of a true belief.

Other commentators in the virtue epistemology camp, broadly conceived, have put forward similar suggestions. For example, Wayne Riggs (2002) and Greco (e.g., 2002) have argued for a 'credit' version of virtue epistemology, according to which the agent, in virtue of bringing about the positively valuable outcome of a true belief, is due credit as a result. Rather than treating the extra value of knowledge over true belief as deriving simply from the agent's attainment of the target true belief, however, Riggs and Greco instead argue that we should regard the agent's knowing as the state the agent is in when she is responsible for her true belief. Only in so doing, they claim, can we answer the value problem.

Interestingly, however, other virtue epistemologists, most notably Ernest Sosa (2003), have also advocated a 'credit' view of this sort, yet seem to stay within the machine-product picture of belief that Zagzebski thinks is so problematic. That is, rather than analyse the state of knowing as consisting of both the true belief and its source, they regard the state of knowing as distinct from the process, yet treat the fact that the process is intrinsically valuable as conferring additional value on any true belief so produced. With Sosa's view in mind, it is interesting to ask just why we need to analyse knowledge in the way that Zagzebski and others suggest in order to get around the value problem.

The most direct way to approach this question is by considering whether it is really true that a valuable cause cannot confer value on its effect where cause and effect are kept separate in the way that Zagzebski claims is problematic in the case of knowledge. One commentator who has objected to Zagzebski's argument by querying this claim on her part is Berit Brogaard (2007; cf. Percival 2003; Pritchard 2007*a*: §2; 2007*b*), who claims that a valuable cause can indeed confer value on its effect in the relevant cases. Brogaard claims that virtue epistemologists like Zagzebski and Riggs endorse this claim because they adhere to what she call a "Moorean" conception of value, on which if two things have the same intrinsic properties, then they are equally valuable. Accordingly, if true belief and knowledge have the same intrinsic properties (which is what would be the case on the view of knowledge that they reject), it follows that they must have the same value. Hence, it is crucial to understand knowledge as having distinct intrinsic properties from true belief before one can hope to resolve the value problem.

If one holds that there is only intrinsic and instrumental value, then this conception of value is compelling, since objects with the same intrinsic properties trivially have the same amount of intrinsic value, and they also plausibly have the same amount of instrumental value as well (at least in the same sort of environment). As Brogaard points out, however, the Moorean conception of value is problematic because—as Wlodek Rabinowicz and Toni Roennow-Rasmussen (1999; 2003) have pointed out—there seem to be objects which we value for their own sake but whose value derives from their being

extrinsically related to something else that we value. That is, such objects are *finally* (i.e., non-instrumentally) valuable without thereby being intrinsically valuable. (For criticism of this account of final value, see Bradley 2002).

The standard example in this regard is Princess Diana's dress. This would be regarded as more valuable than an exact replica simply because it belonged to Diana, which is clearly an extrinsic property of the object. Even though the extra value that accrues to the object is due to its extrinsic properties, however, it is still the case that this dress is (properly) valued for its own sake, and thus valued non-instrumentally.

Given that value of this sort is possible, then it follows that it could well be the case that we value one true belief over another because of its extrinsic features—i.e., that the one true belief, but not the other, was produced by a reliable cognitive trait that is independently valuable. For example, it could be that we value forming a true belief via a reliable cognitive trait more than a mere true belief because the former belief is produced in such a way that it is of credit to us that we believe the truth. There is thus a crucial lacuna in Zagzebski's argument, and hence she needs to do more to motivate the claim that we must reject the machine-product model of belief in order to respond to the value problem.

A different response to the challenge that Zagzebski raises for reliabilism is given by Michael Brady (2006). In defence of reliabilism, Brady appeals to the idea that to be valuable is to be a fitting or appropriate object of positive evaluative attitudes (such as admiration or love) (e.g., Brentano 1969; Chisholm 1986; Wiggins 1987; Gibbard 1990; Scanlon 1998). That one object is more valuable than another is thus to be understood, on this view, in terms of the fact that that object is more worthy of positive evaluation. Thus, the value problem for reliabilism on this conception of value comes down to the question why knowledge is more worthy of positive evaluation on this view than mere true belief. Brady's contention is that, at least within this axiological framework, it *is* possible for the reliabilist to offer a compelling story about why reliable true belief (and thus knowledge) is more valuable than mere true belief.

Central to Brady's argument is his claim that there are many ways one can positively evaluate something, and thus many different ways something can be valuable. Moreover, Brady argues that we can distinguish *active* from *passive* evaluative attributes, where the former class of attitudes involve pursuit of the good in question. For example, one might actively value the truth, where this involves, for instance, a striving to discover the truth. In contrast, one might at other times merely passively value the truth, such as simply respecting or contemplating it.

With this point in mind, Brady's central thesis is that on the reliabilist account knowledge is more valuable than true belief because certain active positive evaluative attitudes are fitting only with regard to the former (i.e., reliable true belief). In particular, given its intrinsic features, reliable true belief is worthy of active love, whereas an active love of unreliable (i.e., accidental) true belief because of its intrinsic features would be entirely inappropriate because there is nothing that we can do to attain unreliable true belief that wouldn't conflict with love of truth.

This is an intriguing proposal, and certainly opens up a possible avenue of defence against Zagzebski's attack on reliabilism since she doesn't even consider the possibility of applying this axiological framework here. One problem that it faces, however—as Pritchard (2007*a*: §3) points out—is that it is unclear whether we can make sense of the distinction Brady draws between active and passive evaluative attitudes, at least in the epistemic sphere. When Brady talks of passive evaluative attitudes towards the truth he gives examples like contemplating, accepting, embracing, affirming, and respecting. Some of these attitudes are not clearly positive evaluative attitudes, however. Moreover, some of them are not obviously passive either. For example, is to contemplate the truth really to evaluate it *positively*, rather than simply to consider it? Furthermore, in accepting, affirming or embracing the truth, isn't one *actively* positively evaluating the truth? Wouldn't such evaluative attitudes manifest themselves in the kind of practical action that Brady thinks is the mark of active evaluative attitudes? More needs to be said about this distinction before it can do the philosophical work that Brady has in mind.

4. Virtue Epistemology and the Value Problem

So far this discussion has taken it as given that, whatever problems reliabilism faces in this regard, there are epistemological theories available—some form of virtue epistemology, for example—that can deal with them. But not everyone in the contemporary debate accepts this. Perhaps the best known sceptic in this respect is Jonathan Kvanvig (2003), who in effect argues that while virtue epistemology (along with a form of epistemic internalism) can resolve the primary value problem, the real challenge that we need to respond to is that set by the secondary value problem; and Kvanvig says that there is no solution available to *that*. That is, Kvanvig argues that there is an epistemic standing—in essence, justified true belief—which falls short of knowledge but which is no less valuable than knowledge. He concludes that the focus of epistemology should not be on knowledge at all, but rather on *understanding*, an epistemic standing that Kvanvig maintains is clearly of more value than knowledge *and* those epistemic standings that fall short of knowledge (such as justified true belief).

What Kvanvig says about understanding will be considered below. First though, let us consider the specific challenge that he poses for virtue epistemology. In essence, Kvanvig's argument rests on the assumption that it is essential to any virtue-theoretic account of knowledge—and any internalist account of knowledge as well, for that matter (i.e., an account that makes a subjective justification condition necessary for knowledge possession)—that it also includes an anti-Gettier condition. If this is right, then it follows that even if virtue epistemology has an answer to the primary value problem (and Kvanvig concedes that it does), it will not thereby have an answer to the secondary value problem since knowledge is not simply virtuous true belief. Moreover, Kvanvig argues that once we recognise what a gerrymandered notion a non-Gettierized account of knowledge is, it becomes apparent that there is nothing valuable about the anti-Gettier condition on knowledge that needs to be imposed. But if that is right, then it follows by even virtue epistemic lights that knowledge—i.e., non-Gettierized virtuous true believing—is no more valuable than one of its proper sub-sets—i.e., mere virtuous true believing.

There are at least two aspects of Kvanvig's argument that are problematic. To begin with, it isn't at all clear why the anti-Gettier condition on knowledge fails to add value, something that seems to be being assumed here. More generally, Kvanvig seems to be implicitly supposing that if an analysis of knowledge is ugly and gerrymandered then that is itself reason to doubt that knowledge is particularly valuable (at least assuming that there are epistemic standings that fall short of knowledge which can be given an elegant analysis). While a similar assumption about the relationship between the elegance (or otherwise) of the analysis of knowledge and the value of the analysandum is commonplace in the contemporary epistemological literature—see, for example, Zagzebski (1999) and Williamson (2000: chapter 1)—this assumption is contentious. (For critical discussion of this assumption, see DePaul *forthcoming*).

In any case, a more serious problem is that Kvanvig seems not to have noticed that many virtue epistemologists—among them Sosa (1988; 1991; 2007), Zagzebski (e.g., 1996; 1999) and Greco (2002; 2007; *forthcoming a*; *forthcoming b*)—think that their view *can* deal with Gettier problems without needing to add an additional anti-Gettier condition on knowledge. The way this is achieved is by making the move noted above of treating knowledge as a state that includes both the truly believing and the virtuous source by which that true belief was acquired. Greco (*forthcoming a*: 14), for example, makes a distinction between (i) a belief's being true and virtuously formed, and (ii) a belief's being true *because* virtuously formed. On the virtue-theoretic account he proposes, knowledge is to be analysed as the latter, and it is only when so analysed, claims Greco, that virtue epistemology can respond to the Gettier problem. Kvanvig, however, resolutely reads virtue epistemologists as uniformly offering an account of knowledge cast along the lines of (i).

Moreover, building on earlier work by Sosa and Zagzebski on this score, Greco (e.g., *forthcoming a*) argues that it is only if the virtue epistemological proposal is read in this way that it is able to answer the value problem. More specifically (though he does not put the point in these terms), the answer to the value problem offered by virtue epistemology on this construal is able to respond to not only the secondary value problem but also the tertiary value problem. This is because knowledge, on this view, is simply the cognitive aspect of a more general notion, that of achievement. That is, Greco argues that achievements are successes that are because of ability, and thus, given that on his view knowledge is cognitive success (i.e., true belief) that is because of cognitive ability, knowledge is cognitive achievement. The import of this claim to our current discussion is that achievements are, plausibly at least, of final value (Greco himself describes the kind of value in play here as intrinsic value, but it is clear from how he describes it that it is specifically final value that he has in mind). If this is right, then cognitive achievements—i.e., knowledge—will also have final value, and thus one is well on one's way to answering the tertiary value problem (and thus the secondary value problem also).

There are thus two key theses to this account of the value of knowledge—that achievements are finally valuable, and that knowledge is a form of achievement—both of which could be called into question. As regards the first thesis, one might object that some successes that are because of ability—i.e., achievements, on this view—are too trivial or easy or wicked to count as finally valuable. As Pritchard (2007*b*) argues, however, this line of objection is far from decisive. After all, it is open to Greco to argue that the claim is

only that all achievements *qua* achievements are finally valuable, not that the overall value of every achievements is particularly high. It is thus consistent with the proposal that some achievements have a very low—perhaps even negative, if that is possible—value in virtue of their other properties (e.g., their triviality). Indeed, a second option in this regard is to allow that not all achievements enjoy final value whilst nevertheless maintaining that it is in the nature of achievements to have such value (e.g., much in the way that one might argue that it is in the nature of pleasure to be a good, even though some pleasures are bad). Since, as noted above, all that is required to meet the (tertiary) value problem is to show that knowledge is generally distinctively valuable, this claim would almost certainly suffice for Greco's purposes.

In any case, it is the second claim that Greco makes—i.e., that knowledge is to be understood as a kind of achievement—that is the most controversial. There are two key problems with this claim. The first is that there sometimes seems to be more to knowledge than a cognitive achievement; the second is that there sometimes seems to be less to knowledge than a cognitive achievement.

As regards the first claim, notice that achievements seem to be compatible with at least one kind of luck. Pritchard (2007*b*) offers the following example to illustrate this point. Suppose that an archer hits a target by employing her relevant archery abilities, but that the success is 'gettierized' by luck intervening between the archer's firing of the arrow and the hitting of the target. For example, suppose that a freak gust of wind blows the arrow off-course, but then a second freak gust of wind happens to blow it back on course again. The archer's success is thus lucky in the sense that it could very easily have been a failure. When it comes to 'intervening' luck of this sort, Greco's account of achievements is able to offer a good explanation of why the success in question does not constitute an achievement. After all, we would not say that the success was because of the archer's ability in this case.

Notice, however, that not all forms of luck are of this intervening sort. For suppose that nothing intervenes between the archer's firing of the arrow and the hitting of the target. However, the success is still lucky in the relevant sense because, unbeknownst to the archer, she just happened to fire at the only target on the range that did not contain a forcefield which would have repelled the arrow. Is the archer's success still an achievement? Intuition would seem to dictate that it is; it certainly seems to be a success that is because of ability, even despite the luckiness of that success. Achievements, then, are, it seems, compatible with luck of this 'environmental' form even though they are not compatible with luck of the standard 'intervening' form.

The significance of this conclusion for our purposes is that knowledge is incompatible with *both* forms of luck. In order to see this, one only needs to note that an epistemological analogue of the archer case just given is the famous barn façade example. In this example, we have an agent who forms a true belief that there is a barn in front of him. Moreover, his belief is not subject to the kind of 'intervening' luck just noted and which is a standard feature of Gettier-style cases (it is not as if, for example, he is looking at what appears to be a barn but which is not in fact a barn, but that his belief is true nonetheless because there is a barn behind the barn shaped object that he is looking at). Nevertheless, his belief is subject to

environmental luck in that he is, unbeknownst to him, in barn façade county in which every other barn-shaped object is a barn façade. Thus, his belief is only luckily true in that he could very easily have been mistaken in this respect. Given that this example is structurally equivalent to the 'archer' case just given, it seems that just as we treat the archer as exhibiting an achievement in that case, so we should treat this agent as exhibiting a cognitive achievement here. The problem, however, is that it is almost universally accepted that the agent in the barn façade case lacks knowledge. Knowledge, is seems, is incompatible with environmental luck in a way that achievements, and thus cognitive achievements, are not.

Greco (2007; *forthcoming a*; *forthcoming b*; cf. Pritchard 2007*b*; 2007*c*; *forthcoming a*) has made a number of salient points regarding this case. For example, he has argued for a conception of what counts as a cognitive ability according to which the agent in the barn façade case would not count as exhibiting the relevant cognitive ability. Moreover, he has argued that, in any case, there are grounds to think that there may be something special about the concept of knowledge which would mean that knowledge might be more resistant to certain kinds of luck than achievements more generally.

Even if these claims can be made to stick, however, there is a second problem on the horizon, which is that it seems that there are some cases of knowledge which are not cases of cognitive achievement. One such case is offered by Jennifer Lackey (2007), albeit to illustrate a slightly different point. Lackey asks us to imagine someone arriving at the train station in Chicago who, wishing to obtain directions to the Sears Tower, approaches the first adult passer-by she sees. Suppose the person she asks is indeed knowledgeable about the area and gives her the directions that she requires. Intuitively, any true belief that the agent forms on this basis would ordinarily be counted as knowledge. Indeed, if one could not gain testimonial knowledge in this way, then it seems that we know an awful lot less than we think we know. What is significant about this case, however, is that we would not intuitively regard the truth of the agent's belief as being because of her cognitive abilities. Indeed, if anything, we would think that her cognitive success was down to her *informant's* cognitive abilities. Thus, it seems that there are cases of knowledge which are not also cases of cognitive achievement.

It is worth being clear about the nature of this objection. Lackey takes cases like this to demonstrate that one can possess knowledge without it being of any credit to one that one's belief is true. As Pritchard (e.g., 2007*b*) points out, however, this conclusion is surely too strong, in that the agent *is* employing her cognitive abilities to some degree, and so surely deserves *some* credit for the truth of the belief formed (she would not have asked just anyone, for example, nor would she have simply accepted just any answer given by her informant). The point is thus rather that whatever credit the agent is due for having a true belief, it is not the kind of credit that reflects a *bona fide* cognitive achievement because of how this cognitive success involves 'piggy-backing' on the cognitive efforts of others.

The obvious way in which someone like Greco would respond to this sort of case is to either claim that, despite first appearances, the agent concerned does not have knowledge or else claim that she does have knowledge but that, on closer inspection, this is a genuine cognitive achievement after all. Neither alternative looks particularly appealing, though no doubt a strong case can be constructed in support of at

least one of these responses. The fundamental problem facing the view, however, as Pritchard (2007*b*) points out, is that once one combines this problem with the one mentioned earlier (i.e., the problem that some cognitive achievements don't seem to be cases of knowledge), the view starts to look far from compelling. Indeed, Pritchard (2007*b*) argues that the moral that should be drawn is that there is no adequate response to the tertiary value problem available after all (the virtue-theoretic account being the most promising line on this score) and that a revisionary response to this problem should therefore be favoured. In particular, the claim is that given the close connection between knowledge and cognitive achievements, and the general final value of achievements, it is not surprising that knowledge is thought to be distinctively (i.e., finally) valuable even though closer reflection reveals that it is actually a distinct epistemic standing, that of cognitive achievement, that is distinctively valuable.

5. Understanding and Epistemic Value

As noted above, the main conclusion that Kvanvig (2003) draws from his reflections on the value problem is that the real focus in epistemology should not be on knowledge at all but on understanding, an epistemic standing that Kvanvig does think is especially valuable but which, he argues, is distinct from knowing (i.e., one can have knowledge without the corresponding understanding, and one can have understanding without the corresponding knowledge). It is perhaps this aspect of Kvanvig's book that has prompted the most critical response, so it is worth briefly dwelling on the debate regarding his claims in this respect in a little more detail here.

To begin with, one needs to get clear what Kvanvig has in mind when he talks of understanding, since many commentators have found the conception of understanding that he targets problematic. The two usages of the term 'understanding' in ordinary language that Kvanvig focuses on—and which he regards as being especially important to epistemology—are "when understanding is claimed for some object, such as some subject matter, and when it involves understanding that something is the case." (Kvanvig 2003: 189) The first kind of understanding he calls "objectual understanding", the second kind "propositional understanding". In both cases, understanding requires that one successfully grasp how one's beliefs in the relevant propositions cohere with other propositions one believes (e.g., Kvanvig 2003: 192, 197-8). This requirement entails that understanding is directly factive in the case of propositional understanding and indirectly factive in the case of objectual understanding (i.e., the agent needs to have at least mostly true beliefs about the target subject matter in order to be truly said to have objectual understanding of that subject matter).

Given that understanding—propositional understanding at any rate—is factive, Kvanvig's argument for why understanding is distinct from knowledge does not relate to this condition (as we will see in a moment, it is standard to argue that understanding is distinct from knowledge precisely because only understanding is non-factive). Instead, Kvanvig notes two key differences between understanding and knowledge: that understanding, unlike knowledge, admits of degrees, and that understanding, unlike knowledge, is compatible with epistemic luck. Most commentators, however, have tended to focus not on these two

theses concerning the different properties of knowledge and understanding, but rather on Kvanvig's claim that understanding is (at least indirectly) factive.

For example, Elgin (*forthcoming*; cf. Elgin 1996; 2004) and Riggs (*forthcoming*) both argue that it is possible for an agent to have understanding and yet lack true beliefs in the relevant propositions. For example, Elgin (*forthcoming*) that it is essential to treat scientific understanding as non-factive. She cites a number of cases in which science has progressed from one theory to a better theory where, we would say, understanding has increased in the process even though the theories are, strictly speaking at least, *false*. A different kind of case that Elgin offers concerns scientific idealisations, such as the ideal gas law. Scientists know full well that no actual gas behaves in this way, yet the introduction of this useful fiction clearly improved our understanding of the behaviour of actual gasses. (For a defence of Kvanvig's view in the light of these charges, see Pritchard (2007*a*: §5) and Kvanvig (*forthcoming a; forthcoming b*)).

A very different sort of challenge to Kvanvig's treatment of understanding comes from Brogaard (2005). She argues that Kvanvig's claim that understanding is of greater value than knowledge is only achieved because he fails to give a rich enough account of knowledge. More specifically, Brogaard claims that we can distinguish between objectual and propositional knowledge just as we can distinguish between objectual and propositional understanding. Propositional understanding, argues Brogaard, no more requires coherence in one's beliefs than propositional knowledge, and so the difference in value between the two cannot lie here. Moreover, while Brogaard grants that objectual understanding does incorporate a coherence requirement, this again fails to mark a value-relevant distinction between knowledge and understanding because the relevant counterpart—objectual knowledge (i.e., knowledge of a subject matter)—also incorporates a coherence requirement. So provided that we are consistent in our comparisons of objectual and propositional understanding on the one hand, and objectual and propositional knowledge on the other, Kvanvig fails to make a sound case for thinking that understanding is of greater value than knowledge.

A further line of criticism against Kvanvig concerns his claim that knowledge is distinct from understanding on the grounds that only the latter admits of degrees and is compatible with epistemic luck (to keep matters simple, we will focus on propositional understanding in what follows). As Pritchard (*forthcoming b*) notes, the import of the former claim is moot since it does not follow from this claim that there are cases in which knowledge is possessed and yet the corresponding understanding is not, or that there are cases in which understanding is possessed but the corresponding knowledge is not. This point becomes especially important once one notices that the relationship between understanding and epistemic luck may well not be quite as Kvanvig supposes.

Stephen Grimm (2006), for example, argues that understanding is just as incompatible with epistemic luck as knowledge is. In contrast, Pritchard (*forthcoming b*) argues that *both* Grimm and Kvanvig are wrong on this score, in that while understanding is compatible with a certain kind of epistemic luck—'environmental' luck of the sort described earlier—that knowledge is incompatible with, it is incompatible with a second kind of epistemic luck—'intervening' luck—that knowledge is incompatible with. Pritchard's diagnosis for why this might be the case is that understanding, unlike knowledge, is a form of cognitive

achievement, for recall we saw him arguing earlier that cognitive achievements, unlike knowledge, are compatible with environmental epistemic luck as well. If this is right, then it enables Pritchard to offer a more explicit account of why understanding is distinctively valuable in the way that Kvanvig alleges. After all, if all achievements are finally valuable, and understanding is itself a type of cognitive achievement, then understanding will be finally valuable too, unlike knowledge on this view.

6. Other Accounts of the Value of Knowledge

John Hawthorne (2004) has recently argued that knowledge is valuable because of the role it plays in practical reasoning. More specifically, Hawthorne (2004: 30) argues for the principle that one should use a proposition p as a premise in one's practical reasoning only if one knows p. Hawthorne primarily motivates this line of argument by appeal to the lottery case. This concerns an agent's true belief that she holds the losing ticket for a fair lottery with long odds and a large cash prize, a belief that is based solely on the fact that she has reflected on the odds involved. Intuitively, we would say that such an agent lacks knowledge of what she believes, even though her belief is true and even though her justification for what she believes—assessed in terms of the likelihood, given this justification, of her being right—is unusually strong. Moreover, were this agent to use this belief as a premise in her practical reasoning, and so infer that she should throw the ticket away without checking the lottery results in the paper for example, then we would regard her reasoning as problematic.

Lottery cases therefore seem to show that justified true belief, no matter how strong the degree of justification, is not enough for acceptable practical reasoning—instead, knowledge is required. Moreover, notice that we can alter the example slightly so that the agent does possess knowledge while at the same time having a *weaker* justification for what she believes (where strength of justification is again assessed in terms of the likelihood, given this justification, that the agent's belief is true). If the agent had formed her true belief by reading the results in a reliable newspaper, for example, then she would count as knowing the target proposition and can then infer that she should throw the ticket away without criticism. It is more likely, however, that the newspaper has printed the result wrongly than that she should win the lottery. This sort of consideration seems to show that knowledge, even when accompanied by a relatively weak justification, is better (at least when it comes to practical reasoning) than a true belief that is supported by a relatively strong justification but does not amount to knowledge. If this is the right way to think about the connection between knowledge possession and practical reasoning, then it seems to offer at a potential response to at least the secondary value problem.

In response to Hawthorne's claim about the pivotal role of knowledge in practical reasoning, Matthew Weiner (*forthcoming*) has argued that knowledge is not important in itself for practical reasoning. More specifically, he argues that knowledge is what he calls a "Swiss Army Concept", in the sense that when we ascribe knowledge we thereby ascribe several valuable sub-concepts—Weiner lists truth, justification, persistence, stability of justification, and safety (i.e., that one could not have easily been wrong). Each of these sub-concepts could be valuable to us, depending upon which standpoint on our practical reasoning we take, but on no standpoint is knowledge of particular value. Thus, claims Weiner, the value of knowledge

relates to how ascribing knowledge is a shorthand way of ascribing a number of valuable sub-concepts, each of which may be of particular value in our practical reasoning depending upon what standpoint on our practical reasoning we take. But knowledge has no special value in itself, at least as regards practical reasoning.

For example, one standpoint that could be taken on your practical reasoning is to care about whether things turn out well for you. From this standpoint, argues Weiner, it is truth that is especially important. Consider again a case in which one reasons from one's true belief that one owns a losing lottery ticket to the conclusion that one should not bother checking the lottery result in the paper. Although this reasoning might seem generally suspicious, as we noted above, from this specific perspective it is unobjectionable— after all, things do turn out well for you in this case because, since your premises are true, you rightly save yourself the trouble of finding out what the local newspaper says about the lottery result.

In contrast, from other standpoints, such at that of caring that your reasoning not be vulnerable to criticism, a different result will be generated. (In the case under consideration, for example, it would be remiss not to check the lottery result in a local newspaper given that there is a chance that you have won a large cash prize, and given also that making such a check would not be unduly onerous). From different perspectives, then, the very same practical inference could be assessed differently. But from no perspective, argues Weiner, does it matter that one knows one's premises.

Whether or not Weiner is right about this, it is an intriguing possibility that the value placed on knowledge might derive from its being a Swiss Army Concept in this way. For it highlights the fact that how we understand the concept of knowledge can have important ramifications for how we go about determining the special value, if any, of knowledge.

A second author who thinks that our understanding of the concept of knowledge can have important ramifications for the value of knowledge is Edward Craig (1990). Craig's project begins with a thesis about the value of the concept of knowledge. Simplifying somewhat, Craig hypothesises that the concept of knowledge is important to us because it fulfils the valuable function of enabling us to identify reliable informants. The idea is that it is clearly of immense practical importance to be able to recognise those from whom we can gain true beliefs, and that it was in response to this need that the concept of knowledge arose. As with Hawthorne's theory, this proposal, if correct, could potentially offer a resolution of at least the secondary value problem.

What is particularly interesting about Craig's approach for our present purposes is that he claims that the concept of knowledge has evolved over time away from its original function through a process Craig calls "objectification". In essence, the process of objectification occurs because the need to eliminate error that is built-into the concept of knowledge becomes 'stretched' to accommodate increasingly demanding error-possibilities as we become intellectually more sophisticated. This is why, according to Craig, we have ended up with a concept of knowledge that sometimes denies knowledge to those who are clearly good

informants (e.g., when some far-fetched error-possibility is made salient), even though the original function was to enable us to identify reliable informants.

This proposal that the concept of knowledge may have changed over time so that what we now call 'knowledge' may sometimes perform a different function to the one that our original concept of knowledge was supposed to track is clearly of central importance to debates about the value of knowledge, as Craig's account of objectification indicates. After all, if we make the plausible assumption, with Craig, that at least the original function of knowledge was to pick out some property of agents that was valuable to us, then on this picture of an 'evolving' concept we can both account for the fact that our present-day conception of knowledge seems to be of particular value even though there are cases in which knowledge is ascribed where it is not of any distinctive value. If this is right, then we should not look to the necessary and sufficient conditions for knowledge in order to determine why it seems to be of distinctive value to us, but rather merely to its original function. (For critical discussion of Craig's view, see Feldman (1997)).

7. The Value of True Belief

So far, in common with most of the contemporary literature in this regard, we have tended to focus on the value of knowledge relative to other epistemic standings. A related debate in this respect, however—one that has often taken place largely in tandem with the mainstream debate on the value of knowledge—has specifically concerned itself with the value of true belief and I will close by considering this issue.

Few commentators treat truth or belief as being by themselves valuable (though see Kvanvig 2003: ch. 1), but it is common to treat true belief as valuable, at least instrumentally. True beliefs are clearly often of great practical use to us. The crucial *caveat* here, of course, concerns the use of the word 'often'. After all, it also often the case that a true belief might actually militate against one achieving one's goals, as when one is unable to summon the courage to jump a ravine and thereby get to safety, because one knows that there is a serious possibility that one might fail to reach the other side. In such cases it seems that a false belief in one's abilities—e.g., the false belief that one could easily jump the ravine—would be better than a true belief, if the goal in question (jumping the ravine) is to be achieved.

Moreover, some true beliefs are beliefs in trivial matters, and in these cases it isn't at all clear why we should value such beliefs at all. Imagine someone who, for no good reason, concerns herself with measuring each grain of sand on a beach, or someone who, even while being unable to operate a telephone, concerns herself with remembering every entry in a foreign phonebook. Such a person would thereby gain lots of true beliefs but, crucially, one would regard such truth-gaining activity as rather pointless. After all, these true beliefs do not seem to serve any valuable purpose, and so do not appear to have any instrumental value (or, at the very least, what instrumental value these beliefs have is vanishingly small). It would, perhaps, be better—and thus of greater value—to have fewer true beliefs, and possibly more false ones, if this meant that the true beliefs that one had concerned matters of real consequence.

At most, then, we can say that true beliefs often have instrumental value. What about final (or intrinsic) value? One might think that if the general instrumental value of true belief was moot then so too would be the intuitively stronger thesis that true belief is generally finally valuable. Nevertheless, many have argued for such a claim.

One condition that seems to speak in favour of this thesis is that as truth seekers we are naturally curious about what the truth is, even when that truth is of no obvious practical import. Accordingly, it could be argued that from a purely epistemic point of view, we do regard all true belief as valuable for its own sake, regardless of what further prudential goals we might have (e.g., Goldman 1999: 3; Lynch 2004: 15-16; Alston 2005: 31). Curiosity will only take you so far in this regard, however, since we are only curious about certain truths, not all of them. To return to the examples given a moment ago, no fully rational agent is curious about the measurements of every grain of sand on a given beach, or the name of every person in a random phonebook (i.e., no rational person wants to know these truths independently of having some prudential reason for knowing them).

Still, one could argue for a weaker claim and merely say that it is *prima facie* or *pro tanto* finally good to believe the truth (cf. David 2005; Lynch *forthcoming*), where cases of trivial truths such as those just given are simply cases where, *all things considered*, it is not good to believe the truth. After all, we are familiar with the fact that something can be *prima facie* or *pro tanto* finally good without being all-things-considered good. For example, it may be finally good to help the poor and needy, but not all-things-considered good given that helping the poor and needy would prevent you from doing something else which is at present more important (such as saving that child from drowning).

At this point one might wonder why it matters so much to (some) epistemologists that true belief is finally valuable. Why not instead just treat true belief as often of instrumental value and leave the matter at that? The answer to this question lies in the fact that many want to regard truth—and thereby true belief—as being in some sense the fundamental epistemic goal. Accordingly, if true belief is not finally valuable—and only typically instrumentally valuable—then this seems to downplay the status of the epistemological project.

There are a range of options here. The conservative option is to contend that truth is the fundamental goal of epistemology and also contend that true belief is finally valuable—at least in some restricted fashion. Marian David (2001; 2005) falls into this category. In contrast, one might argue that truth is the fundamental goal while at the same time claiming that true belief is *not* finally valuable. Sosa (see especially 2004, but also 2000; 2003) seems (almost) to fall into this camp, since he claims that while truth is the fundamental epistemic value, we can accommodate this thought without having to thereby concede that true belief is finally valuable.

Another axis on which this debate can be configured is in terms of whether one opts for an epistemic-value monism or an epistemic-value pluralism—that is, whether one thinks there is only one fundamental epistemic goal, or several. Kvanvig (e.g., 2005) would be an example of someone who endorses epistemic-

value pluralism, since he thinks that there are a number of fundamental epistemic goals, with each of them being of final value. Crucial to Kvanvig's argument is that there are some epistemic goals which are not obviously truth-related—he cites the examples of having an empirically adequate theory, making sense of the course of one's experience, and inquiring responsibly. This is important because if the range of epistemic goals that Kvanvig identified were all truth-related, then it would prompt the natural response that such goals are valuable only because of their connection to the truth, and hence not fundamental epistemic goals at all.

Presumably, though, it ought also to be possible to make a case for an epistemic-value pluralism where the fundamental epistemic goals were not finally valuable (or, at least, *à la* Sosa, where one avoided taking a stance on this issue). More precisely, if an epistemic-value monism that does not regard the fundamental epistemic goal as finally valuable can be made palatable, then there seems no clear reason why a parallel view that opted for pluralism in this regard could not similarly be given a plausible supporting story.

Foundationalist Theories of Epistemic Justification

First published Mon Feb 21, 2000; substantive revision Wed Mar 23, 2005

Foundationalism is a view about the structure of justification or knowledge. The foundationalist's thesis in short is that all knowledge and justified belief rest ultimately on a foundation of noninferential knowledge or justified belief.

A little reflection suggests that the vast majority of the propositions we know or justifiably believe have that status only because we know or justifiably believe other different propositions. So, for example, I know or justifiably believe that Caesar was an assassinated Roman leader, but only because I know or justifiably believe (among other things) that various historical texts describe the event. Arguably, my knowledge (justified belief) about Caesar's death also depends on my knowing (justifiably believing) that the texts in question are reliable guides to the past. Foundationalists want to contrast my inferential knowledge (justified belief) about Caesar with a kind of knowledge (justified belief) that doesn't involve the having of other knowledge (justified belief). There is no standard terminology for what we shall henceforth refer to as noninferential knowledge or justification.[1]

For convenience, in what follows we will concentrate on foundationalism about justification. Everything said about justified belief will apply mutatis mutandis to foundationalist views about knowledge. On the "classical" analysis of knowledge, the core of the concept of knowledge *is* justified true belief and the foundational structure of knowledge simply derives from the foundational structure or justification. It should be noted, however, that the presupposition that the structure of knowledge parallels the structure of justification is controversial. Indeed, in a highly influential book, Timothy Williamson (2000) argues that knowledge is unananalyzable and is a concept that should be employed in understanding a host of other interesting epistemic concepts, including the concept of evidence. In short, his view is that our evidence simply consists in everything we know. Justification may have foundations but only because we end a regress of justification with propositions that are known—the evidential foundation on which all justified belief rests is knowledge (186). A discussion of Williamson's view would take us too far afield, however, and in what follows I will continue to suppose that our understanding of knowledge is parasitic upon our understanding of justification, and not vice versa.

It is surely fair to suggest that for literally thousands of years the foundationalist's thesis was taken to be almost trivially true. When an argument was implicitly or explicitly offered for the view it was most often the now famous regress argument. It is important, however, to distinguish two quite different regress arguments for foundationalism—the epistemic regress argument and the conceptual regress argument.

---

1. The Regress Arguments for Foundationalism

Suppose I claim to be justified in believing that Fred will die shortly and offer as my evidence that Fred has an untreatable and serious form of cancer. Concerned, you ask me how I discovered that Fred has the cancer and I respond that it is just a hunch on my part. As soon as you discover that I have no reason at all to suppose that Fred has the cancer, you will immediately conclude that my whimsical belief about Fred's condition gives me no justification for believing that Fred will soon die. Generalizing, one might suggest the following principle:

To be justified in believing P on the basis of E one must be justified in believing E.

Now consider another example. Suppose I claim to be justified in believing that Fred will die shortly and offer as my justification that a certain line across his palm (his infamous "lifeline") is short. Rightly skeptical, you wonder this time what reason I have for believing that palm lines have anything whatsoever to do with length of life. As soon as you become satisfied that I have no justification for supposing that there is any kind of probabilistic connection between the character of this line and Fred's life you will again reject my claim to have a justified belief about Fred's impending demise. That suggests that we might expand our Principle of Inferential Justification (PIJ) to include a second clause:

Principle of Inferential Justification:
To be justified in believing P on the basis of E one must not only be (1) justified in believing E, but also (2) justified in believing that E makes probable P.

With PIJ one can present a relatively straightforward *epistemic* regress argument for foundationalism. If all justification were inferential then for someone S to be justified in believing some proposition P, S must be in a position to legitimately infer it from some other proposition E_1. But E_1 could justify S in believing P only if S were justified in believing E_1, and if all justification were inferential the only way for S to do that would be to infer it from some other proposition justifiably believed, E_2, a proposition which in turn would have to be inferred from some other proposition E_3 which is justifiably believed, and so on, *ad infinitum*. But finite beings cannot complete an infinitely long chain of reasoning and so if all justification were inferential no-one would be justified in believing anything at all to any extent whatsoever. This most radical of all skepticisms is absurd (it entails that one couldn't even be justified in believing it) and so there must be a kind of justification which is not inferential, i.e., there must be noninferentially justified beliefs which terminate regresses of justification.

If we accept the more controversial second clause of PIJ, the looming regresses proliferate. Not only must S above be justified in believing E_1, S must also be justified in believing that E_1 makes likely P, a proposition he would have to infer (if there are no foundations) from some other proposition F_1, which he would have to infer from F_2, which he would have to infer from F_3, and so on *ad infinitum*. But S would

also need to be justified in believing that F_1 does in fact make likely that E_1 makes likely P, a proposition he would need to infer from some other proposition G_1, which he would need to infer from some other proposition G_2, and so on. And he would need to infer that G_1 does indeed make likely that F_1 makes likely that E_1 makes likely P, and so on. Without noninferentially justified beliefs, it would seem that we would need to complete an infinite number of infinitely long chains of reasoning in order to be justified in believing anything!

The above argument relies on the unacceptability of a vicious *epistemic* regress. But one might also argue, more fundamentally, that without a *concept* of noninferential justification, one faces a vicious *conceptual* regress. What precisely is our *understanding* of inferential justification? What makes PIJ true? It is at least tempting to answer that PIJ is an analytic truth. Part of what it *means* to claim that someone has inferential justification for believing some proposition P is that his justification consists in his ability to infer P from some other proposition E_1 that is justifiably believed. But if anything like this is a plausible analysis of the concept of inferential justification, we face a potential vicious conceptual regress. The analysis of inferential justification presupposes an understanding of justified belief. We need to introduce a concept of noninferential justification in terms of which we can then recursively define inferential justification.

Consider an analogy. Suppose a philosopher introduces the notion of instrumental goodness (something's being good as a means). That philosopher offers the following crude analysis of what it is for something to be instrumentally good. *X* is instrumentally good when *X* leads to something *Y* which is good. Even if we were to accept this analysis of instrumental goodness, it is clear that we haven't yet located the conceptual source of goodness. Our analysis of instrumental goodness presupposes an understanding of what it is for something to be good. In short we can't understand what it is for something to be instrumentally good until we have some prior (and more fundamental) understanding of what it is for something to be intrinsically good. The conceptual regress argument for foundationalism puts forth the thesis that inferential justification stands to noninferential justification as instrumental goodness stands to intrinsic goodness.

2. The Analysis of Noninferential Justification

If foundationalists are united in their conviction that there must be a kind of justification that does not depend on the having of other justified beliefs, they nevertheless disagree radically among themselves as to how to understand noninferential justification. In the latter part of this century, the rise of externalist epistemologies has generated even more fundamentally different versions of foundationalism. It will not be possible to survey all of the strikingly different analyses that have been offered of noninferential justification. In what follows we will examine a few of the more prominent versions of classical and contemporary externalist foundationalisms.

2.1 Noninferential Justification as Infallible Belief

Descartes is often taken to be the paradigm of a classical foundationalist. Determined to build knowledge on appropriate and secure foundations he seemed to want to identify foundational knowledge with infallible

belief. Implicitly or explicitly others seemed to follow his lead by restricting noninferentially justified beliefs to beliefs that cannot be mistaken. Thus, for example, when Price (1950) introduced the notion of sense data, knowledge of which would be included in his foundations of empirical knowledge, he contrasted sense data and their nonrelational properties with other sorts of things about which one could be mistaken, implying again that the way to find the correct foundations of knowledge is to eliminate from one's beliefs system all those beliefs that could be false. Following Lehrer (1974, p. 81)) we might formulate the following definition of infallible belief:

S's belief that *P* at *t* is infallible if *S*'s believing *P* at *t* entails[2] that *P* is true.

As Lehrer and others have pointed, it is far from clear that this concept of infallible belief has much relevance to an attempt to understand the epistemic concept of noninferential justification. The first and most striking problem involves necessary truths. Every necessary truth is entailed by every proposition, and thus if I happen to believe a necessary truth, *P*, that I believe *P* will entail that *P* is true. Thus by the above definition my belief that *P* will be infallible whenever *P* is a necessary truth even if *P* is far too complicated for me to prove and I believe it solely on a whim.

Furthermore, a foundation of knowledge and justified belief restricted to infallible beliefs (as defined above) would arguably be far too flimsy to support any sort of substantial epistemic edifice. There are a few contingent propositions that are trivially entailed by the fact that they are believed. My belief that I exist entails that I exist, that I have at least one belief, that someone has beliefs, that experience (broadly construed) exists, etc. But once we get past these sorts of "self-referential" propositions, propositions whose very subject matter encompasses the fact that they are believed, it is hard to come up with uncontroversial examples of infallible beliefs. Ayer (1956, p. 19) argues that as long as the belief that *P* is one state of affairs and *P*'s being the case is an entirely different state of affairs (not including as a constituent the former) there can be no logical absurdity in the supposition that the former could occur without the latter.

Although it doesn't add much to the logical force of the argument, one might employ our hunches about how the brain might work to rhetorically bolster the argument. Consider a standard candidate for an infallible empirical belief, my belief that I am in pain now. It is surely possible that the region of the brain causally responsible for producing the belief that I am in pain is entirely different from the region of the brain causally responsible for producing the pain. There may be a causal connection between the occurrence of the "pain" brain event and the occurrence of the "belief" brain event, or vice versa, but even if the causal connection holds it will be a contingent fact that it does. It hardly seems that the neurophysiologist could discover these (or any other) causal connections purely *a priori*. But if the brain state responsible for my belief that I am in pain is wholly different from the brain state responsible for the pain, and if the connections between them are merely nomological, then it is in principle possible to produce the one without the other. The belief will not entail the truth of what is believed.

2.2 Noninferential Justification as Infallible Justification

It may be that classical foundationalists start off on the wrong foot if they seek foundations in logical relations between the mere fact that someone believes some proposition and the proposition's being true. Noninferential justification is, after all, a kind of justification and if the impossibility of error is essential to noninferential justification, it may be more plausible to locate the source of infallibility in a special kind of justification available in support of a belief. Let us say that *S*'s belief is infallibly justified at *t* when *S*'s justification for believing *P* at *t* relevantly entails the truth of *P*. We need to qualify the entailment as relevant to circumvent the problems discussed earlier. Whenever I have any justification at all for believing a proposition that turns out to be necessarily true, that justification will entail the necessary truth. But we do not want just any sort of justification to yield infallibly justified belief even if the object of that belief is a necessary truth.

What is the difference between relevant and irrelevant entailment? This is a question notoriously difficult to answer, but intuitively it should have something to do with the fact that would make true the proposition entailed and the fact that would make true the proposition that entails it. More specifically, we could say that *P* relevantly entails *Q* only if the fact that would make *P* true is at least a constituent of the fact that would make *Q* true. This suggestion can be considered at best only preliminary since we will obviously need a more detailed account of facts and their constituents. That I have grey hair entails that someone has grey hair, but is my having grey hair a constituent of the fact that is someone's having grey hair? There is certainly a sense in which it is something one can point to in answer to the question "What makes it true that someone has grey hair?" One cannot appropriately point to my having grey hair as something that makes it true that two plus two equals four.

Consider again my belief that I'm in pain (when I am). If such a belief is noninferentially justified, in what does the justification for that belief consist. Surely not in the mere fact that the proposition is believed. What is it that distinguishes this belief from my belief about Caesar's assassination. Some foundationalists want to locate the noninferential justification in the truth-maker for the proposition believed. What justifies me in believing that I'm in pain when I am is the mere fact that I'm in pain. But again, what is it about my being in pain as opposed to Caesar's being assassinated which makes it appropriate to claim that my being in pain justifies me in believing that I'm in pain while Caesar's having been assassinated doesn't justify me in believing that Caesar was assassinated.

It is tempting to think that the foundationalist is better off appealing to some special *relation* that I have to my pain which makes it unnecessary to look to other beliefs in order to justify my belief that I'm in pain. It is the fact that I have a kind of *access* to my pain that no-one else has that makes my belief noninferentially justified while others must rely on inference in order to discover that I am in this state. This takes us to another classical version of foundationalism, the acquaintance theory. Perhaps the best known proponent of an acquaintance theory is Bertrand Russell,[3] but it takes little imagination to read the view into most of the British empiricists. Roughly the view is that what justifies *S* in believing that he is in pain when he does is the fact that *S* is directly and immediately acquainted with his pain in a way in which he is not directly and immediately acquainted with any contingent facts about Caesar, the physical world, the future, and so on. On a correspondence conception of truth, one might add that to be fully justified in

believing a proposition to be true one must be acquainted not only with the fact that makes the proposition true but the relation of correspondence that holds between the proposition and the fact.

In one of the most influential arguments against foundationalism, Wilfrid Sellars (1963, 131-32) argued that the idea of the given in traditional epistemology contains irreconcilable tensions. On the one hand, to ensure that something's being given does not involve any other beliefs, proponents of the view want the given to be untainted by the application of *concepts*. On the other hand, the whole doctrine of the given is designed to end the regress of justification, to give us secure foundations for the rest of what we justifiably infer from the given. But to make sense of making inferences from the given the given must have a truth value. The kind of thing that has a truth value involves the application of concepts or thought, a capacity not possessed (we may presume) by at least lower-order animals.

If there is a solution to the dilemma presented by Sellars (and others) it is to emphasize that acquaintance is not by itself an epistemic relation. Acquaintance is a relation that other animals might bear to properties and even facts, but it also probably does not give these animals any kind of justification for *believing* anything, precisely because these other animals probably do not have beliefs. Without thought or propositions entertained there is no truth, and without a bearer of truth value in the picture there is nothing to be justified or unjustified. The acquaintance theorist can argue that one has a noninferentially justified belief that P only when one has the thought that P and one is acquainted with both the fact that P, the thought that P, and the relation of correspondence holding between the thought that P and the fact that P. On such a view no single act of acquaintance yields knowledge or justified belief, but when one has the relevant thought (entertains the relevant proposition), the three acts together constitute noninferential justification. When everything that is constitutive of a thought or a proposition's being true is immediately before consciousness, there is nothing more that one could want or need to justify a belief. The state that constitutes noninferential justification is a state that contains as constituents both the bearer of truth-value and the truth-maker.[4]

When an acquaintance with the fact that P is part of what constitutes my noninferential justification for believing P, there is a trivial sense in which my noninferential justification is infallible. I can't be directly acquainted with the fact that P while I believe P falsely. There is, however, nothing to prevent an acquaintance theorist from allowing that one can be noninferentially justified in believing P by virtue of being directly acquainted with a fact very similar to, but ultimately different from the fact that P (the fact that makes P true). Such an acquaintance theory could allow for the possibility of noninferentially justified but false belief that P.[5]

3. Objections to Classical Foundationalism

Once the received view, classical foundationalism has come under considerable attack in the last few decades. We have already considered the very influential objection raised by Sellars to the idea of there being a "given" element in experience. It is crucial that the foundationalist discover a kind of *truth* that can be known without inference. But there can be no bearers of truth value without judgment and judgment

involves the application of concepts. But to apply a concept is to make a judgment about class membership, and to make a judgment about class membership always involves relating the thing about which the judgment is made to other paradigm members of the class. These judgments of relevant similarity will minimally involve beliefs about the past, and thus be inferential in character (assuming that we can have no "direct" access to facts about the past). A reply to this objection would take us far afield indeed. Perhaps it will suffice to observe that the objection relies on a number of highly controversial claims about the nature of judgment, most of which the classical foundationalist should and would reject.

The direct acquaintance theorist does presuppose the intelligibility of a world that has "structure" independent of any structure imposed by the mind. Without nonlinguistic facts that are independent of the thoughts and judgments that represent them, one could not make sense of a relation of acquaintance between a person and a fact, a relation that grounds noninferential justification. More radical contemporary rejections of foundationalism may well involve dissastification with the foundationalist's implicit commitment to a strong realistic correspondence conception of truth. Since Kant there has always been a strong undercurrent of anti-realism running through philosophy. The metaphor is that of the mind imposing structure on reality. And there is an intuitively plausible sense in which one can genuinely wonder whether it makes sense to ask about the number of colors that are exemplified in the world independently of some framework provided by color *concepts*. But despite the periodic popularity of extreme anti-realism, it is surely absurd to suppose that it is even in principle possible for a mind to force a structure on a literally unstructured world. There are indefinitely many ways to sort the books in a library and some are just as useful as others, but there would be no way to begin sorting books were books undifferentiated. If a rejection of foundationalism relies on an extreme form of anti-realism so much the worse for the anti-foundationalist.

Just as some anti-foundationalists reject the conception of truth underlying classical foundationalist accounts of noninferential justification, so others profess to be bewildered by some of the fundamental concepts employed in defining noninferential justification. The acquaintance theorist tends to have relatively little to say by way of analyzing what direct acquaintance is. To be sure one can try to give someone a feel for what one is talking about by contrasting one's awareness of pain with the temporary distraction caused by an engrossing conversation. It is tempting to suppose that for a short time the pain was still present but the person with the pain was no longer aware of the fact that the pain exists. This awareness, the acquaintance theorist will argue, is obviously something over and above mere belief in the existence of the pain, as one can believe that one is in a mental state (say a subconscious mental state) without being aware of that state. Like most theories foundationalism will, however, ultimately rest its intelligibility on an appeal to a sui generis concept that defies further analysis. Just as one needs to end epistemic regresses with foundational justification, the foundationalist will argue, so one needs to end conceptual regresses with concepts one grasps without further definition.

Laurence BonJour (1985) raised another highly influential objection to all forms of classical foundationalism (an objection raised before he joined the ranks of foundationalists). The objection presupposed a strong form of what we might call access internalism. Put very superficially the access

internalist argues that a feature of a belief or epistemic situation that makes a belief noninferentially justified must be a feature to which we have actual or potential access. Moreover, we must have access to the fact that the feature in question is probabilistically related to the truth of what we believe. So suppose some foundationalist offers an account of noninferential justification according to which a belief is noninferentially justified if it has some characteristic *X*. BonJour then argues that the mere fact that the belief has *X* could not, even in principle, justify the believer in holding the belief. The believer would also need access to (justified belief that!) the belief in question has *X* and that beliefs of this sort (*X* beliefs) are likely to be true. At least one of these propositions could only be known through inference, and thus the putative noninferential justification is destroyed.

BonJour presented the objection on the way to developing a coherence theory of empirical justification. But it ultimately became obvious that the objection to foundationalism, if good, was too strong. Given the structure of the argument it should become evident that the coherence theory (and any other theory) would be equally vulnerable to the argument. Just replace "*X*" with some complicated description of beliefs cohering with each other. That might suggest to the classical foundationalist that strong access internalism is a view to be avoided.

The Principle of Inferential Justification used to generate the regress argument for foundationalism is itself controversial. It is important to note that either clause of the principle can be used by itself to generate the allegedly vicious epistemic and conceptual regress for the philosopher who rejects foundations. It is the two clauses combined that are supposed to present the anti-foundationalist with an infinite number of vicious regresses. A number of philosophers (among them foundationalists) would argue that the second clause of PIJ confuses levels of epistemic questions. It is far too strong to require someone to have a justified belief in a probabilistic connection between available evidence and the conclusion reached on the basis of that evidence. Such a requirement is at best plausible for having second-level justification for believing that one has an inferentially justified belief. In responding to a challenge presented to one's having an inferentially justified belief in *P* on the basis of *E* one might find oneself searching for justification to support the claim that *E* makes probable *P*, but that is only because in the context of the challenge one is trying to make good (i.e., justify) the claim that one has a justified belief. A similar claim might be made with respect to clause 1) of the principle, although it is more difficult to generate the supporting intuition.

In any event, the careful foundationalist is certainly not *confused* about level-distinctions. The foundationalist who supports PIJ is claiming that a necessary condition for someone's having an inferentially justified belief in *P* based on *E* is that the person have both a justified belief in *E* and a justified belief in the proposition that *E* makes *P* probable. It is simply not enough that *E* is true or that *E* does in fact make probable *P*. Our original examples used to support PIJ would seem to reinforce that conclusion. Even if there happened to be some bizarre connection between palm lines and length of life, for example, the person who has no reason to believe that such a connection exists has no justification for conclusions reached about length of life based on this anatomical feature of people.

Huemer (2002) objects to using examples like the palm-line inference to argue for the second clause of the principle of inferential justification. While most share the intuition that we would need additional information in order to legitimately infer truths about the length of a person's life from knowledge of palm lines, all that really shows is that we wouldn't view the inference in question as legitimate in the first place. Even palm readers don't think that they can make predictions about a person's life based on information about palm lines and that information alone. But that doesn't show that when we have an argument whose premises really do support its conclusion we still need to be aware of the connection in order to justifiably believe the conclusion based on the premises. But while Huemer's point is plausible and the foundationalist trying to argue for the second clause of the principle of inferential justification should heed Huemer's warning to be careful in the use of examples, it is not clear that one can't find plausible examples of inferences from premises to conclusion where the premises do make probable, even entail, the conclusion but the resulting belief is unjustified because the person who reached the conclusion had no awareness of the relevant connection between premise and conclusion. All we need to do is consider a person who infers P from E where E logically entails P, but where the entailment is far too complex for the person to see or even understand. Surely the belief that P is unjustified if the person who reaches conclusion couldn't "see" how the available evidence entais the conclusion.

There are, of course, other responses to the charge of vicious regress facing anti-foundationalists. The coherence theorist rejects the foundationalist's presupposition that justification is linear. Each belief is justified by virtue of its coherence with the rest of what one believes but one avoids the appearance of vicious circularity by insisting that one needn't *first* have justification for believing the other propositions in one's belief system. The coherence theorist's response to the argument for foundationalism is, of course, only as plausible as the coherence theory of justification (See coherence theories of justification).

Peter Klein (1998) may be the lone supporter of a view he calls infinitism. The infinitist accepts the need to be *able* to supply non-circular justification for believing what we do, but argues that given the complexity of the human mind and its capacity to entertain and justifiably believe an infinite number of propositions, there is nothing vicious about the relevant regresses we face. There is no reason to suppose that we would be unable to justify every proposition we believe by appeal to some other different proposition which we justifiably believe. Infinitism is a view that should be seriously considered, particularly once one realizes that one not only can but does have an infinite number of justified beliefs (e.g., that 2 is greater than 1, that 3 is greater that 1, and so on.). It is not clear, however, that even if the infinitist can cope with the epistemic regress argument foundationalism, he has a response to the conceptual regress argument discussed earlier. Klein will argue, however, that one needn't define inferential justification recursively by relying on a base clause that invokes the concept of noninferential justification. Rather one can employ a base clause the invokes an ununanalyzed generic understanding of justification.

Although anti-foundationalists are not always eager to admit it, I suspect that the primary dissatisfaction with classical foundationalism lies with the difficulty the view has avoiding radical skepticism. On infallible belief, infallible justification, or direct acquaintance theories of foundational justification, there is precious little included in the foundations of knowledge. Most classical foundationalists reject the idea that

one can have noninferentially justified beliefs about the past, but the present disappears into the past in the blink of an eye. How can one even hope to get back the vast body of knowledge one pre-philosophically supposes one has, if one's epistemic base is so impoverished. If the second clause of the Principle of Inferential Justification were accepted, the problem is even more serious. One might be able to convince oneself that one can know noninferentially the principles of deductive reasoning, but deduction will not take one usefully beyond the foundations of knowledge and justified belief. As Mill (1906, p. 126) argued, there is a very real sense in which one doesn't advance one's knowledge significantly employing a form of reasoning that takes one only to conclusions that were implicitly contained in the conjunction of one's premises. To advance beyond foundations we will inevitably need to employ non-deductive reasoning and according to PIJ that will ultimately require us to have noninferential (direct) knowledge of propositions describing probability connections between evidence and conclusions. It is not absurd on the face of it to suppose that one can have noninferential *a priori* knowledge of probabilistic connections, but it is perhaps an understatement to suppose that the view is not popular.[6]

4. Externalist Versions of Foundationalism

The epistemic landscape has changed dramatically in the last quarter of a century with the rise of externalist epistemologies. It is notoriously difficult to define clearly the controversy between internalists and externalists in epistemology.[7] It is sometimes taken to be a controversy over whether or not one can identify epistemic properties with "internal" states of believers. Others seem to think that the controversy centers over the question of whether one requires certain sorts of access (or potential access) to the states or properties that constitute having justification. Certainly, paradigm externalists would reject the second clause of the principle of inferential justification. According to virtually all externalists, one can arrive at a justified belief in P by inferring it from E without being aware of any sort of evidential connection between E and P.

While the externalist defends radically different views than those of classical foundationalists, the structure of knowledge and justification that emerges from such theories is still often a foundationalist structure. We might first illustrate the point by examining the view defended by the most prominent of the externalists, Alvin Goldman's reliabilism.[8]

The fundamental idea behind reliabilism is strikingly simple. Justified beliefs are reliably produced beliefs. Justified beliefs are worth having because justified beliefs are probably true. Goldman initially distinguished, however, two importantly different sorts of justified beliefs—those that result from belief-independent processes and those that result from belief-dependent processes. The former are beliefs that are produced by "software" of the brain that takes as its "input" stimuli other than beliefs; the latter are beliefs produced by processes that take as their input at least some other beliefs. So, for example, it is possible that we have evolved in such a way that when prompted with certain sensory input we immediately and unreflectively reach conclusions about external objects. And we may live in a world in which beliefs about the external world produced in this way are usually true (or would usually be true if enough of them were generated).[9] Such beliefs will be justified by virtue of being the product of reliable belief-independent

processes. They can in turn be taken as input for reliable belief-dependent processes in order to generate still more justified beliefs. A belief-dependent process is reliable if its output beliefs are usually (or would usually) be true if the relevant input beliefs are true, and the output beliefs of reliable belief-dependent processes are justified provided that the input beliefs are justified.[10]

The above is but the crudest sketch of Goldman's early reliabilism—he later modified it to deal with a number of objections. But the sketch is enough to bring out the foundationalist structure inherent in a reliabilist account. The reliabilist actually accepts the first clause of PIJ, but avoids both the epistemic and conceptual regresses by embracing a kind of justified belief that does not owe its justification to the having of other different justified beliefs. That the reliabilist is concerned with avoiding the conceptual regress is clear from the fact that the analysis offered is explicitly recursive. The base clause of the recursive analysis in effect captures the concept of a noninferentially justified belief.

I have illustrated the way in which an externalist account of justified belief can exemplify a foundationalist structure by examining one of the most prominent versions of externalism, reliabilism. But other versions of externalism are also implicitly or explicitly committed to a version of foundationalism, or, at the very least, give an account of justification that would enable one to distinguish noninferential from inferential justification, direct from indirect knowledge. Consider, for example, a crude version of the so-called causal theory of knowledge according to which one knows a proposition when one believes it and the belief is caused (in the "right" way) by the very fact that makes true what is believed. Obviously, on such an account one can distinguish causal chains leading to the belief in question that involve intermediate beliefs from those that do not, and using this distinction one can again define a distinction between direct and indirect knowledge.[11]

Externalist versions of foundationalism are probably attractive to many because they often allow at least the possibility of a much expanded foundational base of justified beliefs. The reliabilist's noninferentially justified beliefs, for example, might be produced by processes that are not even very reliable. Unlike the Cartesian, the reliabilist's distinction between noninferentially and inferentially justified belief has nothing to do with how probable it is that the belief in question is true. If nature has been co-operative enough to insure the evolution of cognitive agents who respond to their environmental stimuli with mostly true beliefs then there might be an enormous store of foundational knowledge upon which we can draw in arriving at inferentially justified conclusions. On most externalist accounts of noninferentially justified belief there are literally no *a priori* constraints on what might end up being noninferentially justified.

A full evaluation of externalist versions of foundationalism is far beyond the scope of this article. The very ease with which the externalist can potentially broaden the foundational base of noninferentially justified belief is, ironically, one of the primary concerns of those philosophers unhappy with externalist epistemology. Many internalists are convinced that externalists are simply re-defining epistemic terms in such a way that they lose the kind of meaning that the philosopher wants them to have in order to ask the kind of penetrating philosophical questions that are the peculiar product of a kind of philosophical curiosity. When a philosopher starts looking for justification in support of a belief, the internalist will

argue, the philosopher is interested in achieving a state in which a kind of philosophical curiosity is satisfied. That philosopher wants epistemic justification to provide a kind of assurance of truth. If I'm wondering whether or not I have justification to believe that God exists, I'm hardly going to think that my question has been answered when I'm told by the reliabilist that I might have a reliably produced belief that God exists or when I'm told by the causal theorist that my belief that God exists might be caused by the very fact that God exists. As far as satisfying intellectual curiosity, exemplifying reliably-produced belief or belief caused by the right fact is no more useful than having true belief. If I were to stipulate a technical sense of foundational Knowledge* according to which I foundationally know that *P* when I believe truly that *P* and my belief isn't caused by any other belief, there may well be all sorts of truths I "know", but will having such knowledge do me any good as far as putting me in a state that satisfies my philosophical curiosity?

Coherentist Theories of Epistemic Justification

First published Tue Nov 11, 2003; substantive revision Tue Sep 18, 2007

Coherentism is a view about the structure of justification or knowledge. The coherentist's thesis is normally formulated in terms of a denial of its contrary foundationalism. Coherentism thus claims, minimally, that not all knowledge and justified belief rest ultimately on a foundation of noninferential knowledge or justified belief.

This negative construal of coherentism occurs because of the prominence of the regress problem in the history of epistemology, and the long-held assumption that only foundationalism provides an adequate, non-skeptical solution to that problem. After responding to the regress problem by denying foundationalism, coherentists normally characterize their view positively by replacing the foundationalism metaphor of a building as a model for the structure of knowledge with different metaphors, such as the metaphor which models our knowledge on a ship at sea whose seaworthiness must be ensured by repairs to any part in need of it. Coherentists typically hold that justification is solely a function of some relationship between beliefs, none of which are privileged beliefs in the way maintained by foundationalists, with different varieties of coherentism individuated by the specific relationship among beliefs appealed to by that version.

1. The Regress Problem

When we are justified in believing a claim, we often are so justified because our belief is based on other beliefs. Yet, it is not an adequate defense of a belief merely to cite some other belief that supports it, for the supporting belief may have no epistemic credentials at all — it may be a belief based on mere prejudice, for example. In order for the supporting belief to do the work required of it, it must itself pass epistemic muster, standardly understood to mean that it must itself be justified. If so, however, the question of what justifies this belief arises as well. If it is justified on the basis of some yet further belief, that belief, too, will have to be justified; and the question will arise as to what justifies it.

Thus arises the regress problem in epistemology. Skeptics maintain that the regress cannot be avoided and hence that justification is impossible. Infinitists endorse the regress as well, but argue that the regress is not vicious and hence does not show that justification is impossible. Foundationalists and coherentists agree that the regress can be avoided and that justification is possible. They disagree about how to avoid the regress. According to foundationalism, the regress is found by finding a stopping point for the regress in terms of foundational beliefs that are justified but not wholly justified by some relationship to further beliefs. Coherentists deny the need and the possibility of finding such stopping points for the regress. Sometimes coherentism is described as the view that allows that justification can proceed in a circle (as long as the circle is large enough), and that is one logically possible version of the view (though it is very hard to find a defender of this version of coherentism). The version of coherentism that is more popular, however, objects in a more fundamental way to the regress argument. This version of coherentism denies that justification is linear in the way presupposed by the regress argument. Instead, such versions of coherentism maintain that justification is holistic in character, and the standard metaphors for coherentism are intended to convey this aspect of the view. Neurath's boat metaphor — according to which our ship of beliefs is at sea, requiring the ongoing replacement of whatever parts are defective in order to remain seaworthy–and Quine's web of belief metaphor–according to which our beliefs form an interconnected web in which the structure hangs or falls as a whole — both convey the idea that justification is a feature of a system of beliefs.

To see exactly where this conception of justification takes a stand on the regress problem, a formulation of the standard skeptical version of the regress argument will be helpful. To formulate such an argument, we need to use the idea of an inferential chain of reasons. Such an inferential chain traces the inferential dependence of a given belief, including in it as first link the belief in question, as second link whatever reason justifies it, as third link whatever epistemically supports the reason in question, and so on. The skeptical argument then proceeds as follows:

1. No belief is justified unless its chain of reasons
   o is infinitely long,
   o stops, or
   o goes in a circle.
2. An infinitely long chain of reasons involves a vicious regress of reasons that cannot justify any belief.

3. Any stopping point to terminate the chain of reasons is arbitrary, leaving every subsequent link in the chain depending on a beginning point that cannot justify its successor link, ultimately leaving one with no justification at all.
4. Circular arguments cannot justify anything, leaving a chain of reasons that goes in a circle incapable of justifying any belief.

As noted, coherentists are ordinarily characterized as maintaining that premise 4 of this argument is false. Though such a view would count as a version of coherentism, standard coherentism has no quarrel with 4, but instead rejects 1 because it presupposes that justification is non-holistic. Premise 1 assumes that justification is linear rather than holistic in virtue of characterizing justification in terms of inferential chains of reasons, and it is this feature of the regress problem to which typical coherentists object.

In sum, then, coherentism can be negatively characterized as the view that, first, agrees with foundationalism that there is no regress of justification that is infinite (thereby rejecting both skepticism and infinitism) and, second, disagrees with foundationalism that justification depends on having an inferential chain of reasons with a suitable stopping point. This negative point can be maintained either by denying that the chain has a stopping point, thereby endorsing a linear version of coherentism, or by denying the assumption that justification requires the existence of an inferential chain of reasons, thereby endorsing a holistic viewpoint. Since the primary examples of coherentism in the history of the view are holistic in nature, I will focus in the remainder of this entry on this version of the view.

2. The Positive Account

Coherentists often defend their view by attacking foundationalism, implicitly relying on the implausibility of infinitism and skepticism. They attack foundationalism by arguing that no plausible version of the view will be able to supply enough in the way of foundational beliefs to support the entire structure of belief. This attack takes two forms. First, coherentists argue against the very idea of a basic belief, maintaining that it is always a sensible question to ask, "Why do you believe that (i.e., what reason can you give me for thinking that is true)?" Second, coherentists attack the idea that the kind of foundation developed will be adequate to support the structure. If, as is usual, foundationalists limit foundational beliefs to those about our experience in the specious present, it is hard to see how such a limited foundation can support the entire edifice of beliefs, including beliefs about the past and future, about the vast array of scientific opinion both about the observable realm and the unobservable, and about the abstract domain of mathematical and logical truth and the truths of morality. Foundationalists may, of course, introduce epistemic principles of justification that license whatever chain of reasons they wish to endorse from the foundations to the rest of the edifice of belief, but the resulting theory will look more and more ad hoc as new epistemic principles are offered whenever the threat of skepticism looms regarding a kind of belief not defensible by standard inductive and deductive rules of inference.

Regardless of the persuasiveness of these challenges to foundationalism, coherentists must and do go beyond negative philosophy to provide a positive characterization of their view. A bit of taxonomy and some specific examples will allow us to see how the required positive characterization is provided by coherentists. A useful taxonomy for coherentism can be provided by distinguishing between subjective and

objective versions of coherentism. At a purely formal level, a version of coherentism results from specifying two things: first, the things that must cohere in order for a given belief to be justified, and second, the relation that must hold among these things in order for the belief in question to be justified. In the realm of the logical space of coherentism, both features can be given subjective or objective construals.

2.1 The Things Over Which Coherence is Defined

Consider first the items that need to cohere. As noted already, coherentists typically adopt a subjective viewpoint regarding the items that need to cohere, maintaining that the system on which coherence is defined is the person's system of beliefs. Coherence could be defined relative to other, more objective systems, however. Social versions of coherentism may define coherence relative to the system of common knowledge in a given society, for example, and religious versions may define coherence relative to some body of theological doctrine. These latter two systems are objective in that the obtaining of the system in question implies nothing about the person whose belief is being evaluated. For this reason, they tend to be rather implausible, since they deny the perspectival character of justification, according to which whether or not one's beliefs are justified depends on facts about oneself and one's own perspective on the world. Versions that combine subjective and objective features are also possible. For example, a theory might begin with the system of a person's beliefs, and supplement it with additional claims that any normal person would believe in that person's situation. It is true, however, that standard versions of coherentism are subjective about the items relative to which coherence is defined.

Even if this aspect of the view is subjective, however, belief is not the only subjective item to which a theorist might appeal, leaving one to wonder what explains the uniform agreement among coherentists that coherence should be defined relative to the class of beliefs. The reasons for this uniformity fall into two categories. One kind involves the claim that the only other possibly relevant mental states are experiential states (appearance states, sensation states), and that such states cannot be reasons at all since they lack propositional content(see Davidson 1989). This viewpoint has little plausibility to it, however. It may be true that there are some experiential states without content (perhaps the experience of pain is an experiential state without content), but it is equally true that some have content. It can appear to a person that it is raining, and the mental state involved has as content the proposition that it is raining.

A more plausible way to pursue this kind of argument is to maintain that if experiential states play a role in justification, they'll have to be able to play that role whether or not they are the kind of state that has propositional content. So, if some lack content and cannot be reasons on account of lacking content, then experiential states cannot play a role at all.

The difficulty with this line of argument is the conception of reasons it involves. It is true that if an experience has no content, then it cannot be in virtue of its content that it provides a reason. Even so, it is far from obvious that a reason has to be one in virtue of its content, for if we attend to ordinary defenses people give of their beliefs, they often cite their experience as a reason. One can question whether they are merely explaining their beliefs rather than justifying them, but when that distinction is clarified, they'll still cite their experience as their reason ("Why are you grimacing?" "Because my leg hurts." "Why do you

think your leg hurts?" "Because I can feel it." "Well, your experience may explain why you believe that your leg hurts, but I'm not asking for an explanation of your belief, I'm asking you to provide a reason for thinking that your belief that your leg hurts is correct; can you give me such a reason?" "Yes, because I can feel it hurting…")

The second category of defense for the idea that coherence is a relation on beliefs involves an argument to the effect that other mental states are either irrelevant to the question of the epistemic status of a belief (e.g., affective states such as hoping, wishing, fearing, and the like) or are insufficient for generating positive epistemic status (e.g., states such as sensation states or appearance states) — there is, after all, the issue of what to make of the sensory input, and that issue takes us beyond the sensation state itself (see Lehrer 1974, esp. p. 188). The former point is unproblematic, but the latter point fails to imply the claim in question. Arguing that an appeal to experiential states is insufficient for justification in no way shows that an appeal to such states is not necessary for an adequate account of justification.

There is, however, a deeper motivation behind coherentists' aversion to defining coherence over a subjective system that includes experiential states. The worry is that appealing to experiential states in any way will result in a version of foundationalism. The understanding of foundationalism which results from the regress argument involves two features. The first is an asymmetry condition on the justification of beliefs — that inferential beliefs are justified in a way different from the way in which non-inferential beliefs are justified — and the second is an account of intrinsic or self-warrant for the beliefs which are foundationally warranted and which support the entire structure of justified beliefs. There are various proposals for how this latter commitment of foundationalism is to be formulated, but we can already see the outline of an argument for requiring that coherence not be defined over a system that includes experiential states. For if a theory were to include such states in the class of things with which a belief must cohere in order to be justified, the above considerations might seem to suggest that such a theory would have to involve some notion of intrinsic warrant or self-warrant. Some justification or warrant would be possessed by a belief, but not in virtue of some warrant-conferring relationship to any other belief. Hence, it might seem, this relation between the appearances and related beliefs would have to generate at least some positive degree of warrant for such beliefs, even if that warrant were not sufficient for full justification. Even if not sufficient for full justification, though, the theory would appear typically foundationalist in that it includes some notion of positive warrant not dependent on any relationship to other beliefs.

This argument is quite persuasive, but is ultimately flawed. The distinctive feature of foundationalism, in the context of the relationship between appearances and beliefs, is that this relation between appearances and beliefs is taken to be one which imparts positive epistemic status (perhaps only in the absence of defeaters). So, for example, if a version of foundationalism appeals to the appearance that it is raining as that which undergirds the foundational warrant for the belief that it is raining, that theory must maintain that the appearance supplies some positive warrant for the belief. It is this warrant-conferring requirement that allows coherentism to escape the above argument, for it is open to coherentists to deny that appearances impart, or tend to impart (even in the absence of defeaters), any degree of positive epistemic status for related beliefs. The coherentist can maintain, instead, that appearances are necessary (in the usual

situations) for those beliefs to have some degree of positive epistemic status, but in no way sufficient in themselves for any degree of positive epistemic status. Coherentists can go on to identify what would be sufficient in conjunction with the relation to appearances in typically coherentist fashion, focusing on the way in which any one of our beliefs is related to an entire system of information in question. The resulting theory would be one in which experience plays a role, but not the kind of role that is distinctive of foundationalism.

Another way to make this same point is to recall that coherentism is not committed to the view that coherence is a relation on the system of the person's beliefs. For one thing, coherence might be a relation on an objective body of information, perhaps in the form of coherence with some body of common knowledge (or, more plausibly, by supplementing a system of beliefs with information any normal person would believe). So when coherentists defend a subjective version of the items over which coherence is defined, there cannot be some definitional requirement on the view that coherence must be a relation on a system of beliefs. That conclusion could be drawn only if there were a sound argument that showed that any appeal to experience would turn a theory into a version of foundationalism. Since the argument for that conclusion is flawed as explained above, coherentism proper need not prohibit the subjective system over which coherence is defined from containing experiential states.

2.2 The Relation of Coherence

The second positive feature required of coherentism is a clarification of the relation of coherence itself, and here again we find an important distinction between subjective and objective approaches. The most popular objective approach is explanatory coherentism, which defines coherence in terms of that which makes for a good explanation. On such a view, hypotheses are justified by explaining the data, and the data are justified by being explained by our hypotheses. The central task for such a theory is to state conditions under which such explanation occurs.

BonJour (1985) presents a different objective account of the coherence relation, citing the following five features in his account:

1. logical consistency;
2. the extent to which the system in question is probabilistically consistent;
3. the extent to which inferential connections exist between beliefs, both in terms of the number of such connections and their strength;
4. the inverse of the degree to which the system is divided into unrelated, unconnected subsystems of belief; and
5. the inverse of the degree to which the system of belief contains unexplained anomalies. (pp. 95,98)

These factors are a good beginning toward an account of objective coherence, but by themselves they are not enough. We need to be told, in addition, what function on these five factors is the correct one by which to define coherence. That is, we need to know how to weight each of these factors to provide an assessment of the overall coherence of the system.

Even such a specification of the correct function on these factors would not be enough. One obvious fact about justification is that not all beliefs are justified to the same degree, so once we know what the overall coherence level is for a system of beliefs, we will need some further account of how this overall coherence level is used to determine the justificatory level of particular beliefs. It would be easy if the justificatory level simply matched the overall coherence level for the system itself, but this easy answer conflicts with the fact that not all beliefs are justified to the same degree.

One way to address this problem is to distinguish between beliefs and strength of belief or degrees of belief. We believe some things more strongly or to a greater degree than other things. For example, I believe there is a cup of coffee on my desk much more strongly than I believe that I visited my parents in 1993, even though I believe both of those claims. Using the concept of a degree of belief, a coherentist may be able to identify what degree of belief coheres with a system of (degrees of) belief, and thereby explain how some beliefs are more justified than others. The explanation would be that one belief is more justified than another just in case a greater degree of belief coheres with the relevant system for one of the two beliefs.

The best-known example of a theory that employs the language of degrees of belief is also a useful example of a subjective account of the coherence relation. Such a subjective account can be developed by identifying a subjective theory of evidence that determines whether and when a person's belief, or degree of belief, is justified. A beautiful and elegant theory of this sort is a version of probabilistic Bayesianism. The version in question identifies justified beliefs with probabilistic coherence, so that a (degree of) belief is justified if and only if it is part of a system of beliefs against which no dutch book can be made. (A dutch book is a series of fair bets which are such that, if accepted, are guaranteed to produce a net loss.) In addition, this version of Bayesianism places a conditionalization requirement on justified changes in belief. Conditionalization requires that when new information is learned, one's new degree of belief match one's conditional degree of belief on that information prior to learning it. So if p is the new information learned, one should change one's degree of belief in q so that it matches one's degree of belief in q given p (together with everything else one knows) prior to learning q. The idea is that each person has an internal, subjective theory of evidence at a given time, in the form of conditional beliefs concerning all possible future courses of experience, so that when new information is acquired, all one needs to do is consult one's prior conditional degree of belief to determine what one's new degree of belief should be. Further, it is this subjective theory of evidence that defines the relation of coherence on the system of beliefs in question: coherence obtains when a belief conforms to the subjective theory of evidence in question, given the other items in the set of things over which coherence is defined.

More generally, subjective versions of the coherence relation can be thought of in terms of the specification of a theory of evidence that is fully internal to the believer. One obvious way for the theory of evidence to be fully internal is for the theory of evidence to be contained within the belief system itself, as is true on the Bayesian theory above. There are other options, however. A subjective theory could appeal to dispositions to believe rather than to actual beliefs, or to something like one's deepest epistemic standards for trying to get to the truth and avoid error. Foley (1986) develops such a view in service of a type of

foundationalist theory, understanding one's deepest standards in terms of the views one would hold given time to reflect without limitation and interference, and subjective coherentists could adopt much of this account in service of their view.

This broader characterization of the options open to subjective versions of the coherence relation carries the additional cost of appealing to the concept of what is internal to a believer, a notion that is none too clear (see the related entry justification, epistemic, internalist vs. externalist conceptions of). In broad terms, there are two important ways of thinking about what is internal here, one emphasizing whether the feature in question is somehow "in the head", and the other emphasizing whether the feature is accessible to the believer on the basis of reflection alone. Unconscious beliefs would count as internal in the first sense, but not in the second; one's own existence is internal in the second sense, but presumably not in the first.

When offering a taxonomy of subjective versus objective characterizations of the coherence relation, it is not necessary to prefer one of these characterizations of what is internal. Instead, we can allow either to be used to specify a subjective account. Doing so places a greater burden on what kinds of arguments could be given for preferring one account of the coherence relation to another, and here the arguments will proceed in two stages. The first stage will address whether one's account of the coherence relation should be objective or subjective. On the side of an objective construal are the manifold intuitions in which we describe views as unjustified even though they are, from the point of view of the believer, the best view to hold. For example, we would say that cultic beliefs, such as the belief that accepting a blood transfusion is a terrible thing to do, are unjustified; and our judgment is not altered by learning that the believer in question was raised in the cult and can't be held responsible for knowing better. On the side of a subjective construal are the arguments for access internalism, according to which the fact that some people can't be held responsible for knowing better is a clear sign that their beliefs are justified, for justification is a property whose presence is detected by careful reflection. Another argument for subjective accounts relies on the new evil demon problem. Descartes' evil demon problem threatens the truth of our beliefs, for the demon makes the beliefs of the denizens of that world false. The new evil demon problem involves the concept of justification rather than truth, threatening theories that require objective likelihood of truth for a belief to be justified. For beliefs in demon worlds are false and likely to be so, but seem to have the same epistemic status as our beliefs do, since, after all, they could be us!

Recently, a new argument has appeared for subjective accounts of justification and, by extension, for subjective accounts of the coherence relation, if coherentism is the preferred theory of justification. This argument appeals to the idea that an adequate theory of knowledge needs to account both for the nature of knowledge and for the value of knowledge. This issue arose first in Plato's dialogue between Meno and Socrates, in which Meno originally proposes that knowledge is more valuable than true belief because it get us what we want (his particular example is finding the way to Larissa). Socrates points out that true belief will work just as well, a response that befuddles Meno. When he finally replies, he expresses perplexity regarding two things. He first wonders whether knowledge is more than true belief, and he also questions why we prize knowledge more than true belief. The first issue is one concerning the nature of knowledge,

and the second concerning the value of knowledge. To account for the nature of knowledge requires minimally that one offer a theory of knowledge that is counterexample-free. To account for the value of knowledge requires an explanation of why knowledge is more valuable than its (proper) parts, including true belief and justified true belief (for more on why knowledge is more than justified true belief, see knowledge, analysis of). Such an explanation would seem to require showing two things: first, that justified true belief is more valuable than true belief; and second, that justified true belief plus whatever further condition is needed to produce a counterexample-free account of the nature of knowledge is more valuable than justified true belief on its own. These requirements show the need for a conception of justification that adds value to true belief, and it is difficult for objective theories of justification to discharge this obligation. In the context of objective accounts of the coherence relation, such an account would be governed by a formal constraint to the effect that satisfying that account would increase one's chances of getting to the truth, and theories of justification guided by such a constraint are prime examples of theories that find it difficult to explain why justified true belief is more valuable than mere true belief. The problem they encounter is called "the swamping problem." It occurs when values interact in such a way that their combination is no more valuable than one of them separately, even though both factors are positively valuable. Examples that provide relevant analogies to the epistemic case include: beautiful art is no more valuable in terms of beauty for having been produced by an artist who usually produces beautiful artwork; functional furniture has no more functional value for coming from a factory that normally produces functional furniture. Just so, true beliefs are no more valuable from the epistemic point of view — the point of view defined in terms of the goal of getting to the truth and avoiding error — by having the additional property of being likely to be true.

Adopting a subjective theory allows one to avoid the swamping problem. The swamping problem arises for theories that characterize the teleological concept of justification in terms of properties whose presence makes a belief an effective means for getting to the goal of believing the truth and avoiding error. Subjective theories may also characterize the relationship between justification and truth in terms of a means/ends relationship, but they reject the requirement that something is a means to an end only if it is an effective means to that end, i.e., only if it increases the objective chances of that goal being realized. Subjectivists advert to the deepest and most important goals in life as examples, for such goals are rarely ones for which we have much idea of which means will be effective. Consider, for example, the goal of securing some particular person as a spouse, or the goal of raising psychologically healthy, emotionally responsible children. In each case, there are well-known ways in which achieving these goals can be sabotaged, and so we try not to proceed in that fashion. The problem is that there are too many ways that have worked for other people in securing similar goals, with no good way of assessing which of these ways would be effective in the present case. Doing nothing will certainly not work, but among the various actions available, we can only choose and hope for the best.

Subjectivists say the same for beliefs. They maintain that what is objectively a good ground for a belief is no more transparent to us than is how to maximize happiness over a lifetime. We learn by trial and error on what to base our beliefs, in much the same way as we fumble along in trying for fulfilling existence. In

doing our best in the pursuit of truth, subjectivists hold, we generate justification for our beliefs, even if all we have is hope that our grounds for belief make our beliefs likely to be true.

Whether these arguments on behalf of subjectivism in the theory of knowledge are weighty enough to overcome the strong intuitions on behalf of more objective accounts is not yet settled, though there is something approaching a consensus that subjectivism cannot quite be right in spite of the arguments in its favor. To the extent that the arguments are deemed plausible, a burden is created for relieving the tension that exists between the attractions of objective accounts and the arguments for subjective accounts. One move to reconcile this conflict is to posit different senses of the term 'justified' and its cognates. There are costs to such a move, however. One cost is that subjectivists and objectivists are confused, thinking they are disagreeing when they are not. In ordinary cases when a term has more than one meaning, competent speakers of the language are not confused in this way. Another cost is that ambiguity must be posited without any linguistic clues to its existence, and ambiguities that linguists would not discover but can only be discovered by philosophers are suspect for that reason.

3. Problems for Coherentism

Besides these family disputes within the coherentist clan, there are various problems that threaten to undermine every version of coherentism. The focus here will be on three problems that have been widely discussed: problems related to the non-linear character of coherentism, the input problem, and the problem of the truth connection.

3.1 Problems Related to the Basing Relation

The non-linear approach adopted by the most popular versions of coherentism raises concerns that coherentism is incompatible with a proper account of the basing relation. In brief, an account of the basing relation is needed to explain the difference between a situation where a person has good evidence for a belief, but believes it for other reasons, and a situation where has person holds the belief because of, or on the basis of, the evidence. The idea behind an appeal to the basing relation is that if the explanation of a person's belief does not appeal to the evidence for the belief, then the belief itself is not justified (even if the person has good evidence for the belief and thus the content of the belief is, in some sense, justified for that person). In the former case, where the belief is based on the evidence for it, we will say that the belief is doxastically justified; when there is good evidence for the belief, but the belief is held on other grounds, we will say that the belief is only propositionally justified.

The difficulty is that this way of drawing the distinction makes it appear that holistic coherentism can only use the distinction if, somehow, the entire belief system of a person explains the holding of each belief that is a part of the system since, it would seem, a belief needs to be based on that which justifies it if the belief is to be properly based. If coherentism is at its best in its holistic guises, then coherentism succumbs because it is unable to distinguish properly based from improperly based beliefs (see Pollock 1985). If one goes so far as to maintain the stronger position that coherentism can only be a holistic theory, then coherentists may find themselves in the position of having to maintain that all warranted beliefs are

properly basic. For if holistic coherentists cannot draw a distinction between properly and improperly based beliefs, every belief will have automatically survived all requisite tests for warrant just by cohering with the relevant system. If a belief is properly based when it has survived all appropriate scrutiny, then all warranted beliefs will be properly basic, according to coherentism (see Plantinga 1993).

Another way to voice this complaint is to find in the belief system a set of beliefs that can be inferentially related in an appropriate way, thereby allowing for the final step of the inference to be justified. It doesn't follow, however, that any inferential path using the same set of beliefs is a justifying one, simply because one such path is. So suppose there are two paths through the same set of five beliefs, one allowing for justification and the other not allowing for it. Let the contents of the beliefs be *p*, *q*, *r*, *s*, and *t*. Further, let each belief imply the next in sequence, i.e., *p* implies *q*, *q* implies *r*, and so forth. Assume as well that *p*, *q*, *r*, and *s* are all justified for the person in question. If so, a person can come to justifiably believe *t* by inferring from *p* to *q* to *r* to *s* and then to *t*. Suppose, however, that there are no other inferential relationships here besides the ones already assumed. If the order of inference were from *p* to *s* to *r* to *q* and then to *t*, believing *t* would not be justified. If holistic coherentism can only explain proper basing in terms of whatever justifies the belief, then holistic coherentism will be in trouble since in the case in question there is no difference in the system of beliefs in question. The only difference is in the order of inference, and this difference need imply no difference in belief.

One resource for a coherentist to use in replying to this concern about the basing relation is to distinguish between that which justifies a belief and that which is epistemically relevant to the epistemic status of belief, using this distinction to challenge the assumption that proper basing must be characterized in terms of that which justifies a belief. Consider a very abstract example. Suppose we have evidence *e* for *p*. This evidence can be defeated by further information we have, and this defeater might itself be undermined by even further information, information that would reinstate justification for *p*. Furthermore, there is no limit to the complexity that might be involved in this sequence of defeaters and reinstaters. Suppose, then, that the sequence of defeaters and reinstaters is significantly complex, e.g., suppose there are 20 levels of defeaters and reinstaters. From the perspective of a linear view, what must the person base a belief that *p* on in such a case in order for that belief to be justified? It would be unrealistic to assume that all 20 levels play a causal role in the belief, for it is not necessary to consider explicitly the sequence of defeaters and reinstaters in order to be justified in believing *p*. All that is necessary is that there be a reinstater for every level of defeat. If so, however, even a linear theorist will give an account of the basing relation on which it is acceptable to base a belief on something other than that which justifies the belief, all-things-considered.

Such a theorist may still maintain that one must base the belief on something that imparts prima facie justification (the kind of justification that will be all-things-considered justification if there is a reinstater for every defeater). What matters to the present discussion, however, is that even for non-holists there can be parts of a system of beliefs that are relevant to the justificatory status of a belief and yet which need not play a role in the proper basing of a justified belief. If, on the one hand, everything involved in the all-things-considered justification of a belief has to play a role in the basing relation, then every theory will be susceptible to unrealistic assumptions about the basing relation, for it is implausible to think that known

rebutted defeaters enter into any kind of causal or deliberative process of belief formation and hence are not suitable candidates for helping to explain the presence of the resulting belief. For example, if I build a room with a blacklight in it, but include a device to block the light from shining on anything less than six feet off the floor, then I can know the color of my daughter's shirt without this information about room construction entering into the story of belief formation — I needn't consciously think of that information or engage in any inference guided by it, and that information need to be part of the cause of my belief. If, on the other hand, a belief can be properly based by being based on only part of the all-things-considered justification for the belief, then holists are free to clarify the basing relation in non-holistic terms as well. They can say that a belief is properly based when its presence is explained by features relevant to the all-things-considered justificatory status of a belief, even if these features themselves do not constitute an all-things-considered justification of the belief.

A simple example of such a feature illustrates how this idea would work in a holistic setting. On a holistic theory, every particular belief is insufficient for warrant on its own. Even so, a given belief might be an essential ingredient of the larger system on which coherence is defined, where that system is one of the systems under which a target belief in question could be justified. In such a case, the belief is relevant to the epistemic status of the target belief, even though it imparts no warrant to the target belief. Beliefs with such special epistemic relevance can be used to clarify what is required for a belief to be properly based without violating the holistic requirement that no such beliefs impart any degree of warrant by themselves.

3.2 The Isolation Objection

A second major problem for coherentism is the isolation objection, also called "the input problem," which Laurence BonJour formulates as follows:

Coherence is purely a matter of the internal relations between the components of the belief system; it depends in no way on any sort of relation between the system of beliefs and anything external to that system. Hence if, as a coherence theory claims, coherence is the sole basis for empirical justification, it follows that a system of empirical beliefs might be adequately justified, indeed might constitute empirical knowledge, in spite of being utterly out of contact with the world that it purports to describe. Nothing about any requirement of coherence dictates that a coherent system of beliefs need receive any sort of input from the world or be in any way causally influenced by the world (BonJour 1985, p. 108).

The input problem concerns the relationship between a system of beliefs and the external world. It underlies a multitude of counterexamples to coherentism on which we take a person at a given time with a coherent system of beliefs whose system of beliefs meshes well with their experience of the world at that given time. We then freeze this coherent system of beliefs, and vary the person's experience (so that the person still thinks, e.g., he's climbing a mountain when he's really at an opera house experiencing a performance of La Boheme), thereby isolating the system of beliefs from reality. The result is that coherentism seems to be a theory that allows coherence to imply justification even when the system of beliefs is completely cut off from individuals' direct experience of the world around them.

The standard response by coherentists is to try to find a way to require some effect of experience in a belief system, perhaps in the form of spontaneous beliefs (BonJour 1985). Such attempts are not very promising, and lead to the impression that the only way to deal with the input problem is to transform coherentism into a version of foundationalism. That is, the harder coherentists try to find some ineliminable effect of experience on a belief system, the more their theory hinges on finding a role for experience in the story of justification; and when foundationalism is conceived as the kind of theory that allows such a role, then the efforts of coherentists to find such a role for experience look more like acquiescence to the inevitability of affirming foundationalism. For if the only way to avoid the isolation objection is to insist that a belief system must be responsive to experience in order for the beliefs involved to be justified, and if any appeal to experience commits one to foundationalism, then coherentism succumbs to the isolation objection.

As noted above, however, there is nothing in coherentism proper that requires coherence to be defined solely as a relation on beliefs. It is a mere artifact of the history of the view that coherentists always claim such, and whatever the force of the isolation objection against standard versions of coherentism, it disappears as a problem unique to coherence theories once experience is allowed to play a role in a coherentist theory.

3.3 Problems Related to the Truth Connection

A longstanding objection to coherentism can be expressed by noting that a good piece of fiction will display the virtue of coherence, but it is obviously unlikely to be true. The idea is that coherence and likelihood of truth are so far apart that it is implausible to think that coherence should be conceived of as a guide to truth at all, let alone the singular such guide that justification is supposed to constitute.

This concern over the truth connection is sometimes put in the form of the alternative systems objection, according to which there is always some coherent system to fit any belief into, so that if a person were to make sufficient changes elsewhere in the system, any belief could be justified. This particular version of the worry involves too many distractions from the fundamental problem, however. For one thing, it appeals to the idea of making vast changes to one's system of beliefs, but beliefs are not the sort of thing over which we typically can exert control. Furthermore, there is no reason to think that only one system of beliefs can be justified, so rather than constituting an objection to coherentism, this particular formulation of the problem in question looks more like a pleasantly realistic consequence of any adequate theory of justification.

Hidden behind the explicit language of the alternative systems objection, however, is a deeper concern relying on the idea that justification is somehow supposed to be a guide to truth, and mere coherence is not a likely indicator of truth. The deeper concern will have to be formulated carefully, however, for once we see the proper response to the isolation objection above, it is far from clear how coherentism suffers from any failure on this score that would not equally undermine foundationalism. For one way of thinking about the isolation objection is in terms of the idea that coherent systems of belief can be completely cut off from reality, in the same way that a good piece of fiction can be, and once such severance occurs, likelihood of

truth must go as well. As we have seen, however, nothing about coherentism proper forces it to succumb to this problem (as long as finding a role for experience in the story of justification blocks the objection, as it must if foundationalism can escape the objection), and if coherentists are able to find a role for experience in their theory, then coherentism cannot be criticized for failure to provide a suitable guide to truth anymore than foundationalism can.

Moreover, there are problems with casual formulations of the truth concern. First, such casual formulations can run into difficulty explaining how one can be justified in believing a scientific theory rather than believing merely the conjunction of its empirical consequences. Since the theory implies its empirical consequences, the conjunction will, in ordinary cases, have a higher probability than the theory (since it is a theorem of the probability calculus that if A entails B, then the probability of A is less than or equal to the probability of B). Second, casual formulations of the truth concern ordinarily fall prey to the new evil demon problem discussed earlier. Inhabitants of demon worlds would appear to have roughly the same justified beliefs that we have (since they could be us), but their beliefs have little chance of being true. So any formulation of the truth concern that insists that justification must imply likelihood of truth will have to find an answer to the new evil demon problem. Further, one of the fundamental lessons of the lottery and preface paradoxes has been held to be that justified inconsistent beliefs are possible. (The lottery paradox begins by imagining a fair lottery with a thousand tickets in it. Each ticket is so unlikely to win that we are justified in believing that it will lose. So we can infer that no ticket will win. Yet we know that some ticket will win. In the preface paradox, authors are justified in believing everything in their books. Some preface their book by claiming that, given human frailty, they are sure that errors remain, errors for which they take complete responsibility. But then they justifiably believe both that everything in the book is true, and that something in it is false, from which a contradiction can be easily derived.) The paradoxes are paradoxical because contradictory beliefs cannot be justified, but inconsistent beliefs, even when the inconsistency is known, are not the same thing as contradictory beliefs (the challenge, of course, is to find a principled way to stop the inconsistency from turning into a contradiction). If justified inconsistent beliefs are possible, and it surely seems that they are, then a system of beliefs can be justified even if the entire system has no chance whatsoever of being true.

This possibility of justified inconsistent beliefs has been held to constitute a refutation of coherentism (see, e.g., Foley 1986), but some coherentists have demurred (e.g., Lycan 1996). One idea is to partition a system of beliefs and only apply the requirement of consistency within partitions of the system, not to the entire system itself. If consistency applies only with partitions, then, presumably, that is also where coherence does its work, leaving us with a coherence theory that is less than globally holistic. A further issue is how the partitioning is to be accomplished, and in the absence of an account of how to do so, it remains undetermined whether the possibility of justified inconsistent beliefs is compatible with coherentism.

It is fair to say that the issue of the truth connection has not been resolved for coherentism. In a way, this fact should not be surprising since the issue of the truth connection is a fundamental issue in epistemology as a whole, and it affects not only coherentism but its competitors as well.