NOÛS

Evidentialist Reliabilism

JUAN COMESAÑA
University of Arizona
comesana@email.arizona.edu

1 Introduction

In this paper I present and defend a theory of epistemic justification that combines two competing approaches in contemporary epistemology: reliabilism and evidentialism. Evidentialist reliabilism, as I will call the view, incorporates the best features of reliabilism and evidentialism, while at the same time avoiding their most serious problems. Moreover, evidentialist reliabilism relies on its evidentialist features to overcome some of the problems of reliabilism, and on its reliabilist features to overcome some of the problems of evidentialism. It is, thus, not a chimera, but a well motivated hybrid view.

Roughly speaking, reliabilism is the view that all that matters for the justification of a belief is that the belief be produced by a mechanism that tends to produce *true* beliefs. This pure form of reliabilism must face several problems. Perhaps the most serious problem is that reliabilism seems to imply that a "blind" reliable mechanism—that is, a mechanism that, as a matter of fact, is sensitive to the external environment, but that produces beliefs without the benefit of any conscious reflection, in the subject, of this sensitivity—is good enough to produce justification. According to evidentialism, on the other hand, all that matters for justification is *fitness with evidence*. Without an independent account of fitness, however, evidentialism amounts to little more than a statement of our pre-theoretic judgments about what beliefs are justified. Reliability without evidence is blind, evidence without reliability is empty.

In sections 2 and 3 I present evidentialism and reliabilism, together with some of the problems that these theories must face. In sections 4 through 7

I develop evidentialist reliabilism and explain how it avoids those problems. I then consider, in section 8, some objections to evidentialist reliabilism, and argue that they can be satisfactorily answered. Lastly, in section 9 I specify the senses in which evidentialist reliabilism is an internalist theory and the senses in which it is an externalist theory.

2 Evidentialism

2.1 The View

Evidentialism, as formulated by Conee and Feldman (1985), is the thesis that the justificatory status of a belief for a subject is determined by the evidence that the subject has for that belief. More precisely and more generally:

Evidentialism: Doxastic attitude D toward proposition p is epistemically justified for S at t if and only if having D toward p fits the evidence S has at t.[1]

Three things have to be clarified in order to have a fully intelligible characterization of evidentialism: the notion of *evidence* itself, what it is for a subject to *have* some evidence, and what it is for a doxastic attitude towards a proposition to *fit* some evidence that the subject has. We will see in a moment what Conee and Feldman have to say about what evidence is and what it is to have evidence. As to what it means to say that a doxastic attitude fits some evidence, that is one of the central questions of this paper, and it is addressed in detail in sections 2.2 and 5.

Evidentialism is usually marketed as an internalist theory. Conee and Feldman (2001) characterize internalism as "the view that a person's beliefs are justified only by things that are internal to the person's mental life":

Somewhat more precisely, internalism as we characterize it is committed to the following two theses. The first asserts the strong supervenience of epistemic justification on the mental:

S The justificatory status of a person's doxastic attitudes strongly supervenes on the person's occurrent and dispositional mental states, events, and conditions.

The second thesis spells out a principal implication of S:

M If any two possible individuals are exactly alike mentally, then they are alike justificationally, e.g., the same beliefs are justified for them to the same extent.[2]

Let's for the time being follow Conee and Feldman in calling a theory that satisfies M "mentalist," although the claim that mentalism thus characterized

captures an interesting internalist thesis is doubtful, and will be examined further in section 9.[3]

Is evidentialism as defined above a mentalist theory? No, it isn't. To see that it isn't, consider this possible answer to the first two questions about evidentialism that we asked above: *evidence* is constituted by *facts*, and subjects *have* as part of their evidence the fact that p if and only if they *know* that p.[4] Let's call this view "the knowledge account of evidence." The fact that the knowledge account of evidence is compatible with evidentialism shows that evidentialism by itself is not committed to mentalism, for two subjects who are exactly alike mentally could know different things.[5]

But although evidentialism as defined above is not committed to mentalism, Conee and Feldman are. Their mentalism commits them to the supervenience of justificatory status on the mental, and their evidentialism commits them to the necessary equivalence of facts about justificatory status and evidential facts.[6] These two commitments entail a third one, the supervenience of evidential facts on mental facts. Let us call a main implication of that thesis "evidential mentalism:"

> **Evidential mentalism**: If any two possible individuals are exactly alike mentally, then they are exactly alike evidentially.

That two subjects are exactly alike evidentially means not only that they have the same evidence, but also that their evidence justify them in adopting the same attitudes towards the same propositions. That is, fixing the mental fixes not only what evidence subjects *have*, but also what it is evidence *for*. It is because it is incompatible with evidential mentalism that the knowledge account of evidence is not mentalist. Aside from committing themselves to evidential mentalism, Conee and Feldman do not say much more about what constitutes evidence and what it is to have evidence.[7]

One important thing to notice about evidentialism is that, to put it in a widely used terminology, it is a theory of *propositional* justification only, not of *doxastic* justification—that is, it is a theory about what it is for a doxastic attitude towards a proposition to be justified for a subject, whether the subject adopts that attitude or not, not a theory of what it is for a subject's existing doxastic attitude to be justified. And notice that it won't do to just say that an attitude towards a proposition is justified in the doxastic sense for a subject just in case the subject has that attitude and the attitude is justified in the propositional sense for the subject—subjects can, for instance, believe for the wrong reasons propositions that are nevertheless justified for them. Conee and Feldman recognize the incomplete character of evidentialism in this respect, and they add a theory of "well-foundedness" to their theory of justification in order to deal with doxastic justification. Their characterization of well-foundedness is the following:

WF   S's doxastic attitude D at t toward proposition p is well-founded if and only if:

   (i) having D toward p is justified for S at t; and
  (ii) S has D toward p on the basis of some body of evidence e, such that
    (a) S has e as evidence at t;
    (b) having D toward p fits e; and
    (c) there is no more inclusive body of evidence e' had by S at t such that having D toward p does not fit e'.[8]

It is important to realize that, without this theory of well-foundedness, evidentialism is an incomplete epistemological theory, for at least two reasons. First, doxastic justification is at least as important an epistemological notion as propositional justification. Second, and relatedly, a theory that doesn't accommodate doxastic justification will have little to say about knowledge: I may well believe a proposition that is justified for me, but if I believe it for some bad reason, that belief of mine will not amount to knowledge.[9]

2.2 A Problem for Evidentialism

Even when complemented with a theory of well-foundedness, evidentialism is still incomplete in a crucial respect. Consider the following question:

Q: What necessary and jointly sufficient conditions must evidence e and proposition p satisfy for it to be the case that adopting doxastic attitude D towards p fits e?[10]

Any answer to Q can be thought of as a completion of the following schema:

Doxastic attitudes: Necessarily, adopting D towards p fits e if and only if . . .

And that schema, in turn, can be seen as the conjunction of these other three schemas:

Belief: Necessarily, believing that p fits e if and only if . . .
Disbelief: Necessarily, disbelieving that p fits e if and only if . . .
Suspension of judgment: Necessarily, suspending judgment about p fits e if and only if . . .

I said in the previous section that in order to have a fully intelligible version of evidentialism we need to answer three questions: what is *evidence*, what is it to *have* evidence, and what is it for a doxastic attitude towards a proposition to *fit* some evidence. Q is a more precise version of that third question. We saw in the previous section that evidential mentalism constrains the possible answers to the first two questions. It also constrains the possible

answers to Q. The following partial answer to Q (a completion of the belief schema), for instance, is incompatible with evidentialist mentalism:

Causal fit: Necessarily, believing that p fits e if and only if the fact that p caused e.[11]

To see that causal fit is incompatible with evidential mentalism, consider two possible subjects who are exactly alike mentally, but such that in only one of them is the experience as of something red in front of him caused by the fact that there is something red in front of him. Under the assumption that experiences count as evidence and that our subjects have no other evidence relevant to whether there is something red in front of them, causal fit has it that only one of them is justified in believing that there is something red in front of him, but evidentialist mentalism has it that either both of them are justified or neither of them is. More generally, evidentialist mentalism is incompatible with any answer to Q that makes it a *contingent matter* whether adopting D towards p fits E.

So, even though evidentialism *by itself* is completely silent regarding Q, evidentialist mentalism does have *something* to say about it. Ideally, however, we would like a *full*, *correct* and *informative* answer to Q. Of course, full and informative but incorrect answers are easy to come by. Consider, for instance, the following one: necessarily, adopting D towards p fits e if and only if snow is white. Full and correct but non-informative answers are also easy to come by: necessarily, adopting D towards p fits e if and only if anyone who has evidence e is justified in adopting D towards p. Correctness can also be bought at the cost of informativeness by relinquishing the implicit constraint that the answer must be *finite*. We could (?), for instance, give an infinitely long answer to Q that starts as follows:

Necessarily, adopting D towards p fits e if D = belief, p = there is something red in front of one, and e = an experience as of something red in front of one.[12]

That is, we could, in principle, give a list of "epistemic principles" that state what attitude we are justified in adopting towards which propositions given that we have certain evidence. To then upgrade this list to an answer to Q we would have to add the claim that the *only* way in which adopting an attitude towards a proposition is justified is if it falls under one of the principles in the list. We could (perhaps) even abbreviate the list by listing epistemic-principle-schemas, such as the following one:

Necessarily, adopting D towards p fits e if D = belief, p = there is something X in front of one, and e = an experience as of something X in front of one.[13]

Even with the use of epistemic-principle-schemas, though, the list is likely to be infinite. In that case, it would be hard to argue that the list provides

us with a *full* answer to Q (one that we can understand, anyway). But let us leave that concern aside, and let us suppose that the list of epistemic principles gives us a full and correct answer to Q—the answer would still not be informative. Notice that the complaint here is not that we haven't been given a naturalistically acceptable reduction of the epistemic notion of fit. The complaint is, rather, that we haven't been told *why* precisely those and not other principles belong in the list.[14] Compare Q with this other question:

> **Q***: What are the necessary and jointly sufficient conditions that an action has to satisfy for it to be just?

If the best that we can do in answering Q* is just to give a list of all just actions, then we don't have an *informative* answer to Q*, however full and correct the list may be. Q* doesn't ask which ones are the just actions, it asks the Socratic question, what do all the just actions have in common? It may well turn out that there is no answer to the Socratic question, and that the best that we can do is to offer that list—but this is to *admit* that Q* has no answer, and to offer a consolation answer instead. Analogously, Q doesn't ask which ones are the epistemic truths, it asks the Socratic question, what do all the epistemic truths have in common? It may well turn out that there is no answer to *this* Socratic question either—but, again, to just give a list of all the epistemic truths is to admit that there is no answer to Q.

Even if we cannot have a full answer to Q, *partial*, correct and informative answers would be welcome. For instance, if we could correctly and informatively complete the belief, disbelief or suspension of judgment schemas, that would be progress. Also progress would be to uncover correct and informative necessary conditions, even if they are not jointly sufficient, as would be to uncover correct and informative sufficient conditions, even if they are not necessary.

It is compatible with evidentialism that we cannot have even a *partial* correct and informative answer to Q, because it is compatible with evidentialism (even when complemented with mentalism) that epistemic facts are brute facts. But it is undeniable that settling for brute epistemic facts should not be done lightly. Perhaps we can do better. I will come back to Q in section 5. Let us now turn to what many consider to be the main rival of evidentialism: reliabilism.

3 Reliabilism

3.1 The View

The *locus classicus* for reliabilism as a theory of epistemic justification is of course Goldman (1979). Let's start with a rough version of a reliabilist theory of epistemic justification:

> S's belief that p is justified if and only if that belief was produced by a reliable belief-forming process.

One of the main reasons why that formulation of reliabilism can only be considered a rough approximation is that it ignores the distinction between belief-dependent and belief-independent processes introduced in Goldman (1979). Belief-dependent processes (that is, processes whose inputs include beliefs) need not be unconditionally reliable in order to justify, they only need to be such that they would tend to produce mostly true beliefs if the beliefs that are part of their inputs were true. A belief produced by a belief-dependent process is justified just in case the process is thus conditionally reliable and the input beliefs are themselves justified. We can capture this in the following recursive formulation of reliabilism:

Reliabilism: S's belief that p is justified if and only if:

1. S's belief that p was produced by a belief-independent process that is reliable; or
2. S's belief that p was produced by a belief-dependent process, the beliefs on which the process depends are justified, and the process is conditionally reliable.[15]

Reliabilism is marketed as a version of externalism—indeed, as the *paradigmatic* externalist theory. It is not as obvious as it might seem, however, that this is accurate. The issue of whether reliabilism is compatible with internalism will be taken up later, in section 9.

3.2 Problems for Reliabilism

There are three main problems that reliabilism must face. The first is the objection that reliability is not *necessary* for justification, the second is the objection that reliability is not *sufficient* for justification, and the third is the objection that the notion of reliability, as used in reliabilism, is *unintelligible*.

The objection that reliability is not necessary for justification can be traced back to Cohen (1984), and it has been called the "new evil demon problem" by Sosa (1991). The objection starts by having us consider victims of a Cartesian evil demon. Everyone agrees that these victims don't have knowledge (if only because most of their beliefs are false, but not *only* because of that),[16] but, Cohen argues, we should admit that they can be just as justified as we are. To be sure, some of these victims engage in wishful thinking, hasty generalization, and other epistemic vices—*their* beliefs are not justified. But some other victims are careful reasoners and reasonable "perceivers": they take their experience at face value when there is no conflicting information, they commit as few fallacies as we do, etc. We better say that the beliefs of *these* victims are justified—for what is the relevant difference between them and us? But reliabilism seems to have the consequence that the beliefs of the

victims of an Evil demon can never be justified, because what the demon does is to make the belief-forming processes of his victims nearly completely *un*reliable.

The objection that reliability is not sufficient for justification also has its root in a number of examples. This time the objector alleges that there are cases where the belief-forming process is reliable and yet we would not say that the subject in question has justified beliefs. The *locus classicus* of this objection to reliabilism is BonJour (1980). The cases that BonJour presents in that paper all concern subjects that have clairvoyant powers. The most famous one is the following:

> Norman, under certain conditions that usually obtain, is a completely reliable clairvoyant with respect to certain kinds of subject matter. He possesses no evidence or reasons of any kind for or against the general possibility of such a cognitive power, or for or against the thesis that he possesses it. One day, Norman comes to believe that the President is in New York City, though he has no evidence for or against this belief. In fact, the belief is true and results from his clairvoyant power, under circumstances in which it is completely reliable.[17]

Norman, BonJour says, doesn't know that the President is in New York City because, despite reliabilism, his belief is not justified.

The third objection to reliabilism that we will consider does not consist in counterexamples. It is presupposed in advancing counterexamples that it is clear what verdict reliabilism gives with respect to particular cases, but the objection in question now (known as the "generality problem" for reliabilism) is that this presupposition is false. In a nutshell, the problem is that any *token* process that produces a belief will be an instance of indefinitely many *types* of processes. But it makes sense to talk of the reliability of a process only if we are talking about types (because only types are repeatable entities that can have a number of outputs over which a reliability ratio can be construed), and the different types to which any token process belongs will differ widely with respect to their reliability. Which of the indefinitely many types to which any given token process belongs has to be reliable in order for the belief produced by that process to be justified?

As we will see, the last two of these three problems can be dealt with by moving from reliabilism to evidentialist reliabilism. The new evil demon problem is not resolved just by this move, however, so I will deal with it briefly first.

4 Indexical Reliabilism

The new evil demon problem for reliabilism, remember, is that reliabilism apparently has the consequence that the beliefs of a victim of an evil demon cannot be justified because they are not reliably produced—but, intuitively, they *can* be justified if the victim in question is a careful epistemic agent.

There are now in the literature a number of ways of dealing with this objection. One of the most popular strategies is the postulation of different senses of 'justification', one in which the victims are justified and the other one (the reliabilist one) in which they are not.[18] A different strategy consists in saying that, in order to provide justification, the process would have to be reliable in "normal worlds"—i.e., those worlds which largely resemble how the subject takes the actual world to be.[19]

This is not the place to discuss these strategies in detail, but neither of them is ultimately plausible. With respect to the first one, ambiguities should not be postulated without necessity. It may be that, at the end of the day, the boring truth is that there are different concepts of justification, some of them more or less answering to evidentialist conceptions and others to reliabilist conceptions; but we are nowhere near the end of the day. With respect to the second strategy, it simply isn't reliabilist anymore: if a belief is justified just in case it is produced by processes that would be reliable in worlds that are in general outlines like the subject *takes* the actual world to be, then what determines whether a belief is justified is what the subject *believes* (justifiably, presumably, which raises other problems)—and this is as internalist a position as they come.[20]

But there are ways of dealing with the new evil demon problem which do not consist either in simply postulating ambiguity or in giving the game away to the internalist. The following is my favorite one:[21]

Indexical Reliabilism: S's belief that p is justified if and only if:

1. S's belief that p was produced by a belief-independent process that is *actually* reliable; or
2. S's belief that p was produced by a belief-dependent process, the beliefs on which the process depends are justified, and the process is conditionally *actually* reliable;

where we understand 'actually' to have a two-dimensional semantics.[22] According to indexical reliabilism, then, there are two different propositions associated with any attribution of justification: the "diagonal proposition," which says that the belief was produced by a method that is reliable in the world where it is believed, and the "horizontal proposition," which says that the belief was produced by a process that is reliable in whichever world the proposition is considered.[23] In order to amount to knowledge, a belief must be diagonally justified, but even subjects for whom knowledge is impossible (like victims of an evil demon) can be justified in the horizontal sense. The converse is also true: indexical reliabilism leaves it open that there could be subjects whose beliefs are diagonally but not horizontally justified, and so it leaves it open that there can be knowledge without (horizontal) justification (more on this in sections 6 and 8.3).

Indexical reliabilism, then, has the consequence that victims of an evil demon can have justified beliefs, if those beliefs are produced by processes that are reliable in our world, and in the same sense that our beliefs are justified. This is not the place to argue in detail the merits of indexical reliabilism,[24] but notice that it is a solution to the new evil demon problem that, while granting that victims of an evil demon can be just as justified as we are, still retains whatever externalist character was present in reliabilism (more about this in section 9). Moreover, we don't need to say that there are two senses of 'justified,' just as we wouldn't say that there are two senses of 'here' even if we are convinced that a two-dimensional semantics is the best account of that term—two persons in different cities both of whom say "Clinton is here" express the same diagonal proposition and a different horizontal proposition, but there is no temptation to posit ambiguity in 'here.' Indexical reliabilism is thus different from, and better than, solutions to the new evil demon problem that either postulate ambiguity or give the game away to the internalist.

5 What is Fit?

Let us return now to the problems with reliabilism and evidentialism whose proper solution demands a combination of the views. Let us start with the problem for evidentialism that I identified in section 2.2: the problem of answering Q.

We can categorize the possible answers to Q as follows. First, there is the pessimistic answer that we would like to be false:

Brutal fit: Q doesn't have a full, correct and informative answer, because epistemic facts of the form "adopting D towards p fits E" are brute facts.

Evidentialism is compatible with brutal fit.[25] As I said in section 2.2, there are reasons to try to resist brutal fit, even if at the end of the day we should retreat to it.

Second, there is a class of answers that satisfy the following restriction:

Mental fit: Q can be given a correct, full and informative answer in exclusively mental terms.[26]

It should be noticed here that evidentialism doesn't entail mental fit. One clear way to see this is by noticing, once again, that evidentialism is compatible with brutal fit, which is itself *in*compatible with mental fit, from where it follows that evidentialism is compatible with the negation of mental fit.[27]

One interesting theory that satisfies mental fit can be found in the work of Richard Foley (although Foley himself wouldn't use this terminology).[28] For Foley, a doxastic attitude fits some evidence that a subject S has if and only if S would, on reflection, endorse the epistemic principle linking that doxastic attitude to that evidence. If we are somewhat liberal about what to

count as a mental state, Foley's theory satisfies mental fit. Foley's theory, however, faces a serious problem which can be traced back to the fact that it incorporates an answer to Q that satisfies mental fit. Consider a subject that would endorse seriously misguided epistemic standards, such as the standard that believing that there is a black cube in front of one fits the evidence that one has when one faces a round snowball under standard viewing conditions. If such a subject goes ahead and believes, in those conditions, that there is a black cube in front of him, then he is justified according to Foley's theory, but this is the intuitively incorrect result. To be sure, such a belief has something going for it: it is not subject to criticism by the subject himself. But that only shows that being immune to this kind of self-criticism is not coextensive with epistemic justification.[29]

I believe that objections analogous to this one will apply to any answer to Q that satisfies mental fit. Intuitively put, the fundamental problem with answers that satisfy mental fit is that for a doxastic attitude towards a proposition to fit certain evidence, that evidence must be *a good reason* to take that attitude towards that proposition—and, whereas it is plausible to suppose that evidence itself is constituted by mental states, it is not plausible to suppose that what makes evidence *good* is itself constituted by mental states. This is just the barest sketch of an argument,[30] but it does suffice for considering seriously a different answer to the question about what fit consists in.

Finally, then, consider anti-mental fit:

> **Anti-mental fit**: Q can be given a correct, full and informative answer, but not in mental terms exclusively.

As we shall see, evidentialist reliabilism provides an answer to Q that satisfies anti-mental fit. In particular, in developing evidentialist reliabilism we will appeal to the following full and informative answer to Q:

> **Reliabilist fit**
>
> **Belief**: Necessarily, believing that p fits e for subject S if and only if:
>
>   1. e doesn't include any beliefs of S and the connection between S's having e and p is actually reliable; or
>   2. e includes beliefs of S, all of these beliefs are justified, and the connection between S's having e and p is conditionally actually reliable.
>
> **Disbelief**: Necessarily, disbelieving that p fits e for subject S if and only if:
>
>   1. e doesn't include any beliefs of S and the connection between S's having e and not-p is actually reliable; or

2. e includes beliefs of S, all of these beliefs are justified, and the connection between S's having e and not-p is conditionally actually reliable.

Suspension of judgment: Necessarily, suspending judgment with respect to p fits e for subject S if and only if neither believing nor disbelieving that p fits e for subject S.

Given the desirability of avoiding the postulation of primitive epistemic facts and the shortcomings of mental fit, it is prudent to give anti-mental fit a fair hearing. In particular, reliabilist fit seems to me to be a promising complement to evidentialism.[31]

6 Proto-Evidentialist Reliabilism

Reliabilist fit is also promising in that, as I will now argue, it can help deal with one of the problems for reliabilism that we identified before: the challenge to the sufficiency of reliability for justification.

Now, it might be thought that BonJour's case of Norman is already dealt with by indexical reliabilism: after all, clairvoyance is not *actually* reliable, and so Norman is not justified. But remember that indexical reliabilism presupposes a two-dimensional semantics for 'actual': the sentence "clairvoyance is actually reliable" (uttered in a context where Norman is being discussed) expresses two propositions, a true diagonal proposition (because clairvoyance is reliable in Norman's world), and a false horizontal proposition (because clairvoyance is not reliable in our world). Norman's clairvoyant beliefs have what it takes to amount to knowledge (and they do amount to knowledge provided that they are un-gettierized), but they are still not justified in a very familiar sense of justification: they are not formed by methods that are reliable in our world. However, it can plausibly be claimed that there is *no sense* in which Norman's beliefs are justified, and so the mere fact that, according to indexical reliabilism, there is *a* sense in which they are not is not enough to answer the objection.

Notice that one crucial feature of BonJour's example is that Norman has no *evidence* for or against his clairvoyant powers, or regarding the whereabouts of the President—the belief just "pops up" in his head.[32] This is at least one of the reasons why we have the intuition that Norman is not justified in believing that the President is in New York City: because he doesn't base that belief on any evidence.

If the problem that BonJour identified with reliabilism has to do with its failure to account for the dependency of justification on evidence, then the following revision of indexical reliabilism fixes that problem:

Proto-evidentialist reliabilism: S's belief that p is justified if and only if that belief was produced by a process P which includes some evidence e and:

1. e doesn't include any beliefs of S and P is actually reliable; or
2. e includes beliefs of S, all of these beliefs are justified, and P is conditionally actually reliable.

Proto-evidentialist reliabilism has the consequence that Norman's belief is not justified, because it was produced by a process that doesn't include any evidence.

Proto-evidentialist reliabilism has advantages over both reliabilism and evidentialism. Unlike reliabilism, proto-evidentialist reliabilism handles BonJour-style counterexamples. Unlike evidentialism, proto-evidentialist reliabilism incorporates a non-pessimistic answer to Q. As we will see in the next section, the remaining problem for reliabilism can be solved by further integration of the two theories.

7 Evidentialist Reliabilism

The remaining problem for reliabilism was the challenge to the intelligibility of the notion of reliability as it is used in the theory—the challenge that has come to be known as the "generality problem" for reliabilism. The generality problem is the problem of specifying which process-type has to be reliable in order for the belief produced by a token of that type to be justified.

Proto-evidentialist reliabilism makes some headway towards answering that question, because types of process that don't include evidence should not be considered. But proto-evidentialist reliabilism still faces a generality problem, for there will be many types of process that include evidence under which any given token process of belief-production falls, and it is not guaranteed that they will all have the same reliability ratio. For instance, the token process that produced my belief that there is a computer in front of me falls under all of these types of process, all of which include evidence: perceptual process, visual process, visual process while sober, visual process while sober and not wearing socks under bad lightning conditions, etc.

Intuitively, the answer to the generality problem lies in noticing that some of the types are descriptions of the psychological process that actually produced the belief, whereas other types include irrelevant information, or leave aside relevant information. For instance, any type which includes information about the kind of socks that I am wearing while forming my belief is likely to include irrelevant information, whereas any type that doesn't include information about features of my experience that I take notice of in forming my belief is likely to leave out relevant information. This is the kind of answer to the generality problem advocated by Alston (1995). Conee and Feldman (1998) have resisted Alston's solution. Instead of evaluating their reasons for resisting, however, I will argue that everyone who thinks that epistemology is possible (and, therefore, Conee and Feldman in particular) must accept that something like Alston's solution to the generality problem has to work.[33]

As we remarked regarding evidentialism, an epistemological theory that has nothing to say about doxastic justification is *thereby* incomplete. Thus, if epistemology is possible, then it is possible to have a theory of doxastic justification. As we have seen, evidentialism is not incomplete because it contains a theory of well-foundedness. That theory, remember, makes an essential appeal to the notion of *basing* a belief on certain evidence: roughly, when the subject *bases* his belief on evidence that justifies it, his belief is well-founded. Moreover, this is not a parochial feature of evidentialism, for any theory that accounts for doxastic justification must make (implicit or explicit) use of the basing relation.

Now, the evidentialist theory of well-foundedness has it that every justified belief will be based on some evidence.[34] If so, then the token process that produced that belief will always be an instance of a type of process of the form *producing a belief that p based on evidence e*—for instance, *producing a belief that there is something gray in front of me based on the fact that it looks as if there is something gray in front of me*, etc. That type, I propose, is the one whose reliability is relevant for the justification of the belief.

More explicitly, then, the solution (based on the evidentialist definition of well-foundedness) is the following:

Evidentialist Reliabilism: A belief that p by S is justified if and only if:

1. S has evidence e;
2. the belief that p by S is based on e; and either
   (a) e doesn't include any beliefs and the type *producing a belief that p based on evidence e* is actually reliable; or
   (b) e includes other beliefs of S, all of those beliefs are justified and the type *producing a belief that p based on evidence e* is conditionally actually reliable.

Evidentialist reliabilism is no longer *proto*-evidentialist because it fully incorporates the evidentialist's insight that what evidence the subject has plays a very important role in determining whether a given belief is justified for that subject. Proto-evidentialist reliabilism only requires the existence of some evidence or other in the belief-forming process, whereas for evidentialist reliabilism the justification of the belief depends on whether the subject *used* that evidence to form the belief. Indeed, evidentialist reliabilism can be seen as a form of evidentialism where the overtly epistemic notion of *fit* is replaced by the notion of *reliability*.

Evidentialist reliabilism is a form of evidentialism, then—but does it satisfy evidentialist mentalism? Yes and No. Because it incorporates *indexical* reliabilism, evidentialist reliabilism accounts for *two* notions of justification—what I have called "horizontal" justification (which is conferred by processes that are reliable in the actual world) and what I have called "diagonal"

justification (conferred by processes that are reliable wherever the subject is). Horizontal justification does satisfy evidentialist mentalism, because whenever a process is reliable in the actual world, *that it is reliable in the actual world* is a necessary truth.[35] On the other hand, diagonal justification just is the kind of justification involved in straight, non-indexical reliabilism and, as such, *does not* satisfy evidentialist mentalism, because it is a contingent matter whether a process is reliable wherever the subject is.[36]

Evidentialist reliabilism has two advantages over evidentialism. First, like proto-evidentialist reliabilism, evidentialist reliabilism replaces talk of evidence *fitting* certain propositions with talk of evidence being reliably connected to the truth of certain propositions. This allows us to give a full and *informative* answer to Q (I defend the correctness of the reliabilist answer to Q in sections 8.2 and 8.3). Second, evidentialist reliabilism has the advantage of being upfront about the theory's reliance on the basing relation, rather than relegating it to what might look like an afterthought.

For better or worse evidentialist reliabilism is a theory of doxastic justification and, unlike the evidentialist characterization of well-foundedness, it is a theory of doxastic justification that doesn't rely on an independently defined notion of propositional justification. Someone attracted to evidentialist reliabilism has a number of ways of dealing with propositional justification, two of which stand out. First, he may just be silent about it; after all, there is an asymmetry between propositional and doxastic justification in that, whereas an epistemological theory *must* say something about doxastic justification on pain of being incomplete, it is possible to be a skeptic about the existence of propositional justification. For doxastic justification is necessary for knowledge, and so any epistemological theory is incomplete without a characterization of it; propositional justification would be necessary for knowledge only if it were necessary for doxastic justification, but, as evidentialist reliabilism makes clear, it is not necessary to appeal to an independently defined notion of propositional justification in order to characterize doxastic justification. The second way in which the evidentialist reliabilist may deal with propositional justification is by reversing the evidentialist's order, and characterizing propositional justification in terms of doxastic justification—perhaps on the model of the following schema: S is propositionally justified in believing that p if and only if there is some evidence e that S has such that the type *producing a belief that p based on e* is an actually reliable process.[37]

Views that are in some ways similar to evidentialist reliabilism have been defended in the literature. The closest one is Alston (1988), to which the present article obviously owes a great deal. Alston argues that a belief is justified just in case it is based on internally available grounds that make the content of the belief objectively probable. There are two main differences between Alston's overall position and mine. First, Alston doesn't argue for the view on the same basis as I do—namely, that it helps to solve problems with both evidentialism and reliabilism taken by themselves. The fact that

evidentialist reliabilism incorporates the best features of each view is, to my mind, the best motivation for it. Second, Alston argues for a perspectival version of internalism regarding grounds. In the terminology of this paper, Alston argues that the evidence should be "fairly easily" available to the subject. I advocate no such condition on evidence.[38]

Let's recapitulate. Evidentialist mentalism claims that whether a belief is justified for a subject is determined entirely by the evidence that the subject has (where the evidence is construed mentalistically), and that a subject's belief is well-founded just in case it is based on evidence that fits adopting that attitude and there is no further evidence that the subject has and that doesn't fit belief. The main problem with evidentialism is its silence regarding Q—its compatibility with brutal fit. Reliabilism claims that a subject's belief is justified (well-founded, in the evidentialist terminology) just in case it is produced by a reliable belief forming process. Reliabilism must face three problems: the new evil demon problem (it seems to entail that the beliefs of victims of an evil demon are not justified), BonJour's problem (it seems to entail that subjects whose beliefs are produced by processes that are reliable where they are but not around here cannot have justified beliefs) and the generality problem (the problem of specifying which of the indefinitely many types under which any token process of belief formation falls has to be assessed for reliability). Evidentialist reliabilism solves all four problems. Having thus argued that there are reasons to adopt evidentialist reliabilism, I turn now to the question whether there are reasons to reject it.

8 Objections and Replies

I have extolled the virtues of combining evidentialism and reliabilism, but it might be thought that, together with their virtues, evidentialist reliabilism inherits some of the problems of the views it combines. In this section I aim to briefly indicate how some of the more serious problems can be dealt with.

8.1 Knowledge Without Evidence?
Because of its evidentialist heritage, evidentialist reliabilism entails that every justified belief is based on some evidence. But this seems to rule out the possibility of knowledge without evidence—for instance, it seems to rule out the possibility of innate knowledge. It might be that there really isn't any innate knowledge, but that seems to be a clearly empirical question, not one that can be settled by philosophical fiat.

We should note that, for the objection to work, we must assume that knowledge entails justification. Without this assumption the objection is without any merit, because both evidentialism and evidentialist reliabilism are theories of *justification* and, as such, have no direct consequences about knowledge. Under the assumption that knowledge entails justification, it is

of course true that evidentialist reliabilism is incompatible with knowledge without evidence. However, that doesn't mean that evidentialist reliabilism is incompatible with the existence of innate knowledge, not even if we assume that knowledge entails justification.

Consider first evidentialism. Recall that evidentialism is surprisingly non-committal on the nature of evidence. Even when complemented with evidentialist mentalism, evidentialism imposes only the restriction that mental twins must be evidential twins. Therefore, evidentialism is incompatible with the possibility of innate knowledge only if the possibility of innate knowledge entails that there can be mental twins that are not evidential twins. I don't see how the possibility of innate knowledge enters into any conflict at all with this restriction.

What about the evidentialist conception of well-foundedness, which evidentialist reliabilism incorporates? According to WF, S knows that p only if *there is some evidence* on which S bases his belief that p. Even if the possibility of innate knowledge doesn't conflict with evidentialist mentalism, isn't it incompatible with WF, and thus with evidentialist reliabilism?

Evidentialist reliabilism is certainly incompatible with *some conceptions* of innate knowledge, conceptions according to which it entails that there can be justification without evidence. For instance, take the conception of innate knowledge according to which S has innate knowledge that p if and only if God endowed S with an innate belief that p. According to that conception of innate knowledge, there can be justification without evidence. But, while it is plausible to require of a theory of justification that it be compatible with the possibility of innate knowledge, it is not plausible to require that it be compatible with *any theory* of innate knowledge. Evidentialist reliabilism is, for instance, compatible with the following not implausible conception of innate knowledge: S has innate knowledge that p if and only if S has the innate capacity to understand p, and understanding p is sufficient for knowing p. Under that conception of innate knowledge, it is easy to specify the mental state that constitutes the evidence which justifies S in believing that p: it is the fact that S understands that p.

Evidentialist reliabilism, then, is compatible with credible theories of innate knowledge. That is all that should be required of an epistemic theory of justification in this regard.

I turn now to objections to the evidentialist reliabilism based on its reliabilist heritage.

8.2 The Problem of Necessary Propositions

It is a datum for epistemic theorizing that justification and truth don't coincide: there can be false but justified beliefs and true but unjustified beliefs.[39] Evidentialist reliabilism respects that datum. But it may be a further datum that even necessarily false beliefs can be justified, and even necessarily true beliefs can be unjustified. For instance, if a well-known mathematician tells

me that Fermat's last theorem has been proven false, then I am justified in believing, on the basis of the mathematician's testimony, that Fermat's last theorem is false, and I am justified in disbelieving Fermat's last theorem, even if Fermat's last theorem is necessarily true. As formulated, however, evidentialist reliabilism entails that necessary falsehoods cannot be justifiably believed, and necessary truths cannot be unjustifiably believed—any process of the type *believing that p based on e*, where p is a necessarily false proposition, will have minimal reliability, whereas if p is a necessarily true proposition, it will have maximal reliability, no matter what e stands for.

Several things could be said about this problem. I will mention three possible answers on behalf of evidentialist reliabilism. All three answers seem to me to deserve attention, although I present them in what it seems to me to be an order of increasing plausibility.

First, the evidentialist reliabilist could appeal to idealization. The problem of necessary propositions is intimately related to the problem of logical omniscience for epistemic logics and probabilistic approaches to rationality. Epistemic logics and probabilistic approaches to rationality usually assume that rational agents believe all the logical consequences of what they believe, and, therefore, believe (or have a degree of belief 1) in all (logically) necessary truths and disbelieve (or have a degree of belief 0) in all (logically) necessary falsehoods.[40] In their defense, these theorists have said that epistemic logics, as well as probabilistic theories of rational degrees of belief, presuppose an *idealized* epistemic agent who *is* logically omniscient. All real agents, these theorists agree, fall far short of the ideal, and to that extent they fail to be fully rational. But insofar as this irrationality has as its cause limitations inherent to the human nature, being irrational in this sense need not be taken to be a *fault* of the agent. The evidentialist reliabilist could obviously adopt this strategy, and claim that, although believing a necessary falsehood or disbelieving a necessary truth is never a justified attitude, if these unjustified attitudes are adopted due to limitations inherent to human nature then being thus unjustified need not be taken to be a fault of the agent.[41]

Second, it is important to notice that, from the point of view of an important and influential view of the metaphysics of belief, it is not clear that there is such a thing as belief in necessary propositions. According to a conception of belief championed by Stalnaker,[42] belief is a state of mind whose aim is to correctly represent the world, and to have a belief is to eliminate some but not all of the ways the world could have been. Thus, to believe that snow is white is to eliminate the possibilities that snow is green, red, etc. An omniscient believer eliminates all but one of the possibilities—the actual one. On this picture, it is not clear that belief in a necessary proposition is possible. For someone who believes a necessary falsehood eliminates *every* possibility, whereas someone who believes a necessary truth eliminates *no* possibility. But what does this approach have to say about cases of *apparent* belief in necessary propositions? What do I believe, for instance, when (as

we would naturally put it) I believe that Fermat's last theorem is false, and I disbelieve Fermat's last theorem? The approach need not have anything *systematic* to say here, but the following is a possibility that might work in some cases: I believe that 'No three positive integers a, b, and c can satisfy the equation $a^n + b^n = c^n$, for any integer value of n greater than two' expresses the necessarily false proposition (and *this* is a contingent belief). Of course, if this approach to the metaphysics of belief is correct and there are no beliefs in necessary propositions, then there is no problem of necessary propositions for evidentialist reliabilism.

But suppose that neither the idealization move nor the Stalnakerian approach to belief are correct. Is there anything else that the evidentialist reliabilist could say? There is. Notice why the problem of necessary falsehoods arises for evidentialist reliabilism: because, according to it, the type relevant to determining whether a belief that p is justified includes p itself. Not all versions of reliabilism share this feature. Indeed, most versions of reliabilism individuate types so that the same type can produce belief in different propositions. For instance, according to Sosa's virtue epistemology,[43] a belief is justified just in case it issues from a reliable faculty of belief formation, and a faculty is reliable only relative to a field of propositions:

> Subject S believes proposition p at time t out of intellectual virtue only if there is a field of propositions F, and there are conditions C, such that: (a) p is in F; (b) S is in C with respect to P; and (c) S would most likely be right if S believed a proposition X in field F when in conditions C with respect to X.[44]

If the type to be assessed for reliability includes reference to a *field* of propositions, some of which are true and some of which are false, then it is no longer the case that processes that issue in necessary propositions have only extreme reliability measures. Evidentialist reliabilism can avail themselves of this move, as follows:

> **Evidentialist Reliabilism**: A belief that p by S is justified if and only if:
>
> 1. there is a field of propositions F to which p belongs;
> 2. S has evidence e;
> 3. the belief that p by S is based on e; and either
>   - (a) e doesn't include any beliefs and the type *producing a belief that q (where q is in F) based on evidence e* is actually reliable; or
>   - (b) e includes other beliefs of S, all of those beliefs are justified and the type *producing a belief that q (where q is in F) based on evidence e* is conditionally actually reliable.

Of course, this solution to the problem of necessary propositions involves further work, for we must now give conditions under which some propositions belong to the same field. This means that a variety of the generality

problem is back with us. But it is important that the version of the general-ity problem that must now be faced is considerably weakened: I can see no argument that there is no principled way of specifying fields.

Evidentialist reliabilism, then, has many resources to deal with the prob-lem of necessary propositions.

8.3 Is Evidence-Based Reliability Sufficient for Justification?

I am not concerned now with the cases presented by BonJour. Those cases show that reliability in itself is not sufficient for justification, and thus they are adequately dealt with by evidentialist reliabilism, which requires *evidence-based* reliability. But is evidence-based reliability sufficient for justification? In this section I defend evidentialist reliabilism from potential counterexam-ples to this sufficiency claim.

Potential counterexamples to the sufficiency of evidence-based reliability for justification arise from the possibility of subjects with *alien* doxastic prac-tices who inhabit worlds that are friendly to those practices. By a "doxastic practice" I mean a transition from evidence to a given doxastic attitude, such as the doxastic practice of believing that there is something red in front of one when facing a red apple. By an *alien* doxastic practice I mean a doxastic practice that is not generally implemented by normal humans. There are two reasons why a doxastic practice might be alien: first, it may be that the evidence in question in the doxastic practice is not of a kind that humans ever have; second, it may be that, even though the evidence in question is sometimes had by humans, the doxastic attitude involved in the doxastic transition is not the one that normal humans would have when faced with that evidence. Now, some alien doxastic practices are sure to be reliable, and so evidentialist reliabilism entails that some alien doxastic practices provide justification. To see whether this is the right thing to say about those alien doxastic practices, let us take a look at a couple of particular examples.

For an example of the first kind of alien doxastic practice, one where the evidence in question is not of a kind that is ever had by humans, we can modify BonJour's Norman case. Let us suppose that the way clairvoyance works for Norman is as follows: on some occasions, he has a phenomenal experience that is very similar to seeing something, except that a big "C" is superimposed over the scene. Whenever Norman "sees" a scene with a superimposed "C," he is strongly inclined to accept a proposition depicted in that experience. For instance, Norman has an experience as of seeing the President in front of the Empire State building and, although he realizes that he is not really *seeing* the President, he finds that the experience makes the proposition that the President is in New York City irresistible, and so believes that proposition on the basis of that evidence.

Michael Bergmann has provided us with an example of the second kind of alien doxastic practice, one where the evidence in question is had by normal humans, but where the doxastic response to that evidence is not one that

normal humans would have when faced with that evidence.[45] Bergmann asks us to imagine possible subjects who, upon grabbing a billiard ball in their hands, experience sensations that are qualitatively the same as the olfactory sensations that we typically have when we smell a meadow full of flowers. Moreover, whenever they have an experience of this kind, the doxastic response of Bergmann's subjects is to believe that there is a smallish hard round object in their hands.

Evidentialist reliabilism does have the consequence that both Norman as well as Bergmann's subjects are diagonally justified—because their doxastic practices are reliable in their worlds. What is the intuitive verdict about these cases? It is not clear. It is, for instance, by no means as clear as the verdict in the original Norman case against simple reliabilism.[46] This unclarity actually favors evidentialist reliabilism, because although Norman and Bergmann's subjects are diagonally justified, they are not horizontally justified, and thus evidentialist reliabilism can capture the ambivalence we feel towards them.

As remarked before, we should want to leave it open that subjects can come to know (and, thus, be in some sense justified) even if they inhabit environments where the *only* way to learn from experience is to use processes that are unreliable around here. In evidentialist reliabilism, this requirement is captured in the verdict that both Norman and Bergmann's subjects are diagonally justified. Any theory that entails that alien doxastic practices can *never* give rise to justification and knowledge is guilty of epistemic chauvinism.

On the other hand, there is something undeniably *weird* about alien doxastic practices; something, moreover, *epistemically* weird. There is surely *something* right to the claim that, for instance, believing that there is a smallish round object in your hand is not the correct doxastic response to an experience that is qualitatively the same as the one we have when we smell a meadow full of flowers, and any epistemic theory should try to capture what is right about that claim.[47] Evidentialist reliabilism does capture this claim, for it has the consequence that neither Norman nor Bergmann's subjects are horizontally justified. But what if Bergmann's subjects are actual (in which case their doxastic practices would no longer be alien, of course)? In that case, doesn't evidentialist reliabilism entail that there is no sense in which they are not justified? If Bergmann's subjects are actual, then their doxastic practice is reliable in their world as well as in ours—because they are the same world. But we can still capture the sense that there is something weird about their epistemic practice—if *we* implemented it, then we would not be justified, because that doxastic practice is not reliable *for us*.[48]

Far from being refuted by the possibility of reliable alien doxastic practices, then, evidentialist reliabilism is uniquely positioned to capture both what is justified and what is unjustified about those practices. But let me end the discussion of these cases by noting what form the theory would have if we wanted it to have the consequence that there is absolutely *no sense* in which

alien doxastic practices can give rise to justification. One way of securing that consequence is to weaken evidentialist reliabilism so as to make it posit evidence-based reliability as only a necessary, but not sufficient, condition on justification. This would, of course, bring us back to the problem that evidentialism faces: we would now have a theory compatible with the existence of brute epistemic facts. But, as remarked in our discussion of Q, even having a *partial* correct and informative answer to it beats having nothing to say. Thus, even if evidentialist reliabilism needed to be demoted to providing only a necessary condition for justification, it would still be superior to evidentialism by itself as far as answering Q goes.

8.4 Is Evidence-Based Reliability Necessary for Justification?

I am not concerned now with the new evil demon problem—I take it that evidentialist reliabilism provides an adequate answer to it. But are there cases of doxastic practices that give rise to justification even though they are actually unreliable?

Consider the following case. Hilary doesn't know how elms and beeches look—he can visually discriminate between elms and beeches, but not *as* elms and beeches. He consults Jerry, a famous arborist. Unbeknownst to Hilary, his friends have contacted Jerry and convinced him to play a prank on Hilary. While facing an elm, Jerry tells Hillary that beeches look like *that*. So, whenever Hilary faces an elm and has the characteristic experience, he believes that he sees a beech, and he bases this belief on that characteristic experience and his recollection of what Jerry said.

Now, it is clear that Hilary's belief is justified. But, the objection goes, his belief is not reliable in the way required by evidentialist reliabilism. After all, whenever Hilary believes that he sees a beech on the basis of how it looks and his recollection of what Jerry said his belief is false. How could it possibly be reliable?

There are two things to say about this case. First, even if it were true that, as it stands, evidentialist reliabilism has the consequence that Hilary's belief is unjustified, a modification of evidentialist reliabilism along the lines suggested at the end of section 8.2 might well help here. That is, even if the type in *believing that there is a beech in front of one based on an elm-like experience and an expert's testimony* were unreliable, that wouldn't mean that the type *believing a proposition relevantly like the proposition that there is a beech in front of one based on an elm-like experience and an expert's testimony* is also unreliable.

But, second, it is not at all clear that the first type is unreliable. It is true that the proposition in question—that beeches look *this* way—is not only false, but robustly so (false as a matter of robust biological fact, say). But that fact would be relevant to the reliability of the type in question only if it affected the connection between an expert's saying that beeches look like elms and beeches's looking like elms, and it doesn't. There are, of course,

possible worlds where beeches do look like elms, and in all of those the type is reliable. Moreover, in those worlds where beeches don't look like elms (like our world), relatively few experts would say that they do. Therefore, the type *believing that there is a beech in front of one based on an elm-like experience and an expert's testimony* is reliable.

I have argued that evidentialist reliabilism adequately captures the truth in both evidentialism and reliabilism while avoiding their respective problems. In the present section I have also answered four main objections to the resulting theory. Before concluding, I address the question whether evidentialist reliabilism is an internalist or an externalist theory.

9 Internalism and Externalism

As remarked before in this paper, evidentialism is taken to be a paradigm case of an internalist theory, and reliabilism is taken to be a paradigm case of an externalist theory. Matters are more complicated than that, however. In this section I will be making a number of internalism/externalism distinctions, and I will classify evidentialism, reliabilism and evidentialist reliabilism in terms of these distinctions. For each of the distinctions, the externalist side will be understood to be equivalent to the negation of the internalist side.

Given an evidentialist framework, there are two possible *loci* for the distinction between internalism and externalism. One may be an internalist or an externalist about *evidence*, and one may be an internalist or an externalist about fit. Second, there are (at least) two ways of being an internalist with respect to either evidence or fit. One can be an *accessibilist* or a *mentalist*. Accessibilists believe that the relevant facts (either facts about evidence or facts about fit) are accessible in a privileged way to the subject that is justified. Mentalists believe that the relevant facts are about mental states of the subject. When we take into account that one can be either an accessibilist or a mentalist with respect to either evidence or fit, we have (assuming that the choices are logically independent of each other) sixteen different positions on the internalism/externalism debate.[49] Before examining in which respects evidentialist reliabilism is internalist and in which respects it is externalist, let us ask the same question about both evidentialism and reliabilism.

Let us start with evidentialism. As we said in section 2.1, the theory is marketed as a version of internalism, and more precisely as a version of mentalism. But that classification overlooks the first of the internalism/externalism distinctions just presented. Evidentialism, when complemented with evidential mentalism, is of course mentalist about evidence. But, being silent on the issue, it is compatible with both mentalist and anti-mentalist conceptions of fit.[50]

It should also be noted that the official characterization of mentalism given by Conee and Feldman is not a good classificatory principle. The characterization, remember, is the following:

> **Mentalism**: If any two possible individuals are exactly alike mentally, then they are alike justificationally, e.g., the same beliefs are justified for them to the same extent.

That principle would count as mentalist a theory which admitted *non-mental* factors as justificationally relevant, as long as those non-mental factors obtained necessarily. For, if a non-mental factor obtains necessarily, then there cannot be two possible individuals who are alike mentally and are not alike justificationally, even if the obtaining of that non-mental factor is relevant to their justification. For instance, the indexical-reliabilist notion of horizontal justification counts as mentalist under this definition—but it clearly isn't a notion that is internalist in any interesting sense. Thus, it is better to characterize mentalism as we did above, as the claim that factors of the relevant kind *are* mental, instead of doing it in terms of a supervenience thesis.

What is the relation between evidentialism and accessibilism? Conee and Feldman appealed to mentalism as a way of demarcating between internalism and externalism precisely because they wanted to remain neutral on the question whether we have privileged access to our evidence. So, officially, evidentialism is compatible with both accessibilism and anti-accessibilism about evidence—and, again, in virtue of being silent about fit, it is compatible with both accessibilism and anti-accessibilism about that as well.

So, evidentialism is mentalist about evidence, neutral regarding accessibilism about evidence, and compatible with either internalism or externalism (of either kind) about fit. What about reliabilism? Reliabilism isn't itself concerned with evidence, so there is no direct implication about whether it is internalist or externalist about it. Still, it is plausible to identify the evidence in a reliabilist theory as the inputs to the belief-forming processes. In that case, a reliabilist who (like Goldman (1986)) insisted that the inputs are mental states will be a mentalist reliabilist about evidence, and one who insisted that the inputs reach to the external world will be an anti-mentalist reliabilist about evidence. Reliabilism is, of course, anti-mentalist about fit, for the reliability of a belief-forming process is not a mental state of the relevant (or any other) subject.

What is the relation between reliabilism and accessibilism? With respect to evidence, an anti-mentalist kind of reliabilism will also be anti-accessibilist. Whether a reliabilism that is mentalist about evidence is also accessibilist about it will depend, as it was the case with evidentialism, on whether we have privileged access to the mental states that constitute our evidence. It might seem obvious that reliabilism is anti-accessibilist about fit. After all, isn't the reliability of our belief-forming processes a paradigmatic example of the kind of things to which we *do not* have special access? The situation is not as straightforward as it might seem. It is plausible to suppose that the reliability of our belief-forming processes is something that is presupposed in everything that we know *a posteriori*. If that is so, and if we become aware

of this presupposition, then in order to know anything at all *a posteriori* we would also have to know that our belief-forming processes are reliable. But it is hard to see how we could know this *a posteriori*. If we want to resist the skeptical conclusion that we do not know it, the only option seems to be to say that we know *a priori* that our belief-forming methods are reliable.[51] If the fact that something is knowable *a priori* is enough for us to have privileged access to the fact that it obtains, then we have here a line of reasoning according to which reliabilism has to be accessibilist about fit.

What about evidentialist reliabilism, then? Given its evidentialist heritage, it is mentalist and compatible with both accessibilism and anti-accessibilism about evidence, and given its reliabilist heritage, it is anti-mentalist about fit. Like reliabilism, whether evidentialist reliabilism is accessibilist or anti-accessibilist about fit depends on the interesting question of whether we can know *a priori* that our belief-forming methods are reliable.

10 Conclusion

Almost since the distinction first appeared in the literature, there have been attempts to reconcile internalism and externalism. As remarked before, many of those attempted reconciliations have been postulations of ambiguity. The view proposed here doesn't have that form: I do not say that, in one sense, to be justified is to believe in accordance with one's evidence and, in another sense, to be justified is to have reliable beliefs. Rather, the kind of reconciliation that I have argued for integrates evidence and reliability: to be justified in the only epistemologically important sense is to believe in accordance with one's evidence, and one's beliefs accord with one's evidence if and only if that evidence is reliably connected to the truth of those beliefs.[52]

Notes

[1] Conee and Feldman (1985), p. 83. As Conee and Feldman recognize (Conee and Feldman (1985), p. 102), the fit in question is *epistemic* fit. There may be some attitudes towards propositions that, for instance, *morally* fit the evidence that the subject has, even if the subject is not epistemically justified in adopting that attitude.

[2] Conee and Feldman (2001), pp. 233–4.

[3] A different conception of internalism is *accessibilism*, the view according to which all factors that are justificatorily relevant for a subject are accessible to that subject. More on this in section 9.

[4] Williamson (2000) advances an evidentialist theory along these lines.

[5] Assuming that Williamson is wrong in claiming that knowledge is a mental state.

[6] The attribution of that last commitment rests on the safe assumption that evidentialism is advanced as a necessary truth.

[7] Feldman (1988) argues for the claim that S has p available as evidence at t if and only if S is currently thinking of p, which suggests that evidence is constituted by propositions and that subjects have evidence by believing propositions. In other parts of that same paper, however, Feldman recognizes that the claim is both too restrictive and too liberal: it is too restrictive in that it doesn't leave room for experiential or memorial states to count as evidence (perhaps

"intuitional" states, such as its seeming to you that $2 + 2 = 4$, should also be added to the list); and it is too liberal in that it allows for any belief, no matter how irrationally held, to count as evidence.

[8] Conee and Feldman (1985), p. 93.

[9] Of course, if propositional justification is necessary for doxastic justification, then (given that doxastic justification is necessary for knowledge) propositional justification is necessary for knowledge, and so evidentialism would have something to say about knowledge even in the absence of a full characterization of doxastic justification. But notice that the claim that propositional justification is necessary for doxastic justification is not itself part of the evidentialist definition of propositional justification.

[10] The question is a generalization of the one asked by Feldman (1995).

[11] It should be noted, however, that causal fit is *not* incompatible with evidentialism itself. Feldman (2006) seems to miss this when he says "evidentialism holds that necessarily, people who have the same evidence are justified in believing the same things."

[12] I state the principle for the sake of illustration—I don't mean to commit myself to its truth.

[13] As before, I don't mean to commit myself to the truth of this principle. In fact, in this case the principle is almost certainly false.

[14] This complaint, which is intimately related to what Sosa has called the "problem of scatter" for evidentialism (see, for instance, Sosa (1991), p. 128), arises in a clear form for Chisholm (1977).

[15] Both reliabilism and the theory I will defend, evidentialist reliabilism, are (like evidentialism but unlike the evidentialist characterization of well-foundedness) theories of *prima facie* justification.

[16] After all, those inside The Matrix don't know that they have hands even if they do have hands.

[17] BonJour (1980), p. 21 of the version in Kornblith (2001).

[18] See, for example, Goldman (1988).

[19] A position advocated by Goldman (1986).

[20] In particular, it is a position that is internalist in all of the senses specified in section 9.

[21] The solution presented in the text is inspired by Sosa (1993). Other solutions that don't posit ambiguity and do not give the game away to the internalists include Bergmann (2004) (although Bergmann does abandon reliabilism) and Majors and Sawyer (2005).

[22] Two-dimensional semantics have been developed, in different ways, by Kaplan (1979); Kaplan (1989), Stalnaker (1978), Evans (1977), and Davies and Humberstone (1981). For a comparison of these different frameworks and a review of some of the applications of two-dimensional semantics, see Chalmers (2006). It is important to note that the application of two-dimensional semantics to the new evil demon problem is completely independent of the project of using the framework to explain either some form of the *a priori* (as in Chalmers (1996)) or the nature of philosophical analysis (as in Jackson (1998)).

[23] I am following Stalnaker's terminology for the two-dimensional framework in Stalnaker (1978).

[24] I have done that in Comesaña (2002).

[25] Indeed, Conee (1992) seems to adhere to brutal fit when he says that "[i]t should be acknowledged that no (. . .) reduction seems to be in the offing for the relation of giving evidence" (p. 254).

[26] I intend mental fit and brutal fit to be mutually exclusive, and so I am assuming that the mental facts appealed to in any answer that satisfies mental fit are not epistemic facts.

[27] Another way of seeing that evidentialism doesn't entail mental fit is by noticing that evidentialist reliabilism, the view developed in this paper, is an evidentialist theory incompatible with mental fit.

[28] See, e.g., Foley (1987) and Foley (1993).

[29] For a development of this kind of objection to Foley's theory, see Feldman (1989). Foley replies in Foley (1989). In the example that I am considering, the subject is in a normal environment, where, typically, only round snowballs give one experiences as of a round snowball. The example would be different if the subject were in an environment where black cubes typically produced such an experience—although there would still be *something* wrong with the subject's reaction even in that case. I consider cases of that kind in section 8.3.

[30] I develop the argument in detail in Comesaña (2005).

[31] In what follows I concentrate on belief to the exclusion of the other doxastic attitudes, but characterizations of disbelief and suspension of judgment can be easily produced by following the model of reliabilist fit.

[32] In particular, this has to either be one of the first cases of clairvoyantly acquired beliefs for Norman or he has to have forgotten about the previous ones—otherwise Norman would have inductive evidence for the truth of his clairvoyantly acquired beliefs, and thus those beliefs would be justified on anyone's account.

[33] For a detailed version of the argument, see Comesaña (2006).

[34] This has been challenged, for instance by Plantinga (1996). See Conee and Feldman (2001) for what I take to be a correct reply to Plantinga's specific worries. For related worries, see also Sosa (2007), Lecture Three. See section 8.1 for further consideration of the possibility of justification without evidence.

[35] In other (perhaps unhelpful) words, if process P is reliable in the actual world, then "Process P is reliable in the actual world" is true in every world "considered as counterfactual."

[36] Again, in other words, "Process P is reliable in the actual world" is not (in general) true in every possible world considered as actual, not even if process P is reliable in the actual world.

[37] One question that any reliabilist theory has to answer is the following: what does it mean to say that a process of belief formation is reliable? For the purposes of this paper I assume that the usual answer, namely that a process is reliable to the extent that it tends to produce true beliefs, is correct. It is not crystal clear, however, what it means to say that a process tends to produce true beliefs. One suggestion with which I am sympathetic, but which I cannot develop here, is that we should understand reliability in terms of what is usually called, misleadingly, "epistemic probability," or, only marginally less misleadingly, "logical probability" (see Comesaña (2009)). The idea, pioneered by Keynes (1921) and Carnap (1950), is that, just as there is an objective relation of entailment between "Socrates is a man and all men are mortal" and "Socrates is mortal," so too there is an objective relation of probabilification between "Socrates is a man and most men are less than 7 feet tall" and "Socrates is less than 7 feet tall." If this relation really does exist, then we can say that a process of the type *believing that p based on e* is reliable if and only if e probabilifies p to a sufficiently high degree. It is misleading to call this relation "epistemic" probability because the name suggests that probabilification should be understood in terms of what justifies what, whereas the position that I am suggesting is precisely the contrary: if e justifies belief in p, then that is *because* e probabilifies p. (It should also be noticed that this probabilification relation is arguably necessary, and so appealing to it gives the reliabilist another solution to the new evil demon problem.) Of course, the project of characterizing logical probability has met with considerable skepticism, beginning with Ramsey (1926). Many of those objections are not relevant to the *existence* of the relation of logical probability, and I believe that those that are can be satisfactorily dealt with (for relevant discussion, see Stove (1986), part two), but, as I said, this is not the place to do it.

[38] Goldman (1986) defends the idea that the inputs to the belief-forming process must include mental states, but he appeals to a guidance-deontological conception of justified belief that I would not endorse. Even Conee (1988) argues that "[w]hat can be external to the mind of a person whose belief is justified (...) are epistemological facts about what evidence provides the person's justification and about the nature of the epistemic link of the belief to its justifying evidence" and "[w]hat must be internal (...) is evidence that does in fact suffice to justify the belief" (p. 50).

References

Alston, W. (1988), "An Internalist Externalism," *Synthese* 74, pp. 265–283.

Alston, W. (1995), "How to Think About Reliability," *Philosophical Topics* 23, pp. 1–29.

Bergmann, M. (2004), "Externalist Justification Without Reliability," *Philosophical Issues* 14, pp. 35–60.

Bergmann, M. (2006), *Justification without Awareness*, Oxford University Press.

BonJour, L. (1980), "Externalist Theories of Empirical Knowledge," *Midwest Studies in Philosophy* V, pp. 53–73.

Carnap, R. (1950), *Logical Foundations of Probability*, University of Chicago Press.

Chalmers, D. (1996), *The Conscious Mind*, Oxford University Press.

Chalmers, D. (2006), "Two-Dimensional Semantics," in E. Lepore and B. Smith (eds.), *The Oxford Handbook of Philosophy of Language*, Oxford University Press.

Chisholm, R. (1977), *Theory of Knowledge*, Prentice-Hall, Englewood Cliffs, NJ, 2nd. edition.

Christensen, D. (2004), *Putting Logic in its Place*, Oxford University Press.

Cohen, S. (1984), "Justification and Truth," *Philosophical Studies* 46, pp. 279–295.

Comesaña, J. (2002), "The Diagonal and the Demon," *Philosophical Studies* 110, pp. 249–266.

Comesaña, J. (2005), "We Are (Almost) All Externalists Now," *Philosophical Perspectives* 19(1), pp. 59–76.

Comesaña, J. (2006), "A Well-Founded Solution to the Generality Problem," *Philosophical Studies* 129(127–47).

Comesaña, J. (2009), "What Lottery Problem for Reliabilism?" *Pacific Philosophical Quarterly* 90(1), pp. 1–20.

Conee, E. (1988), "The Basic Nature of Epistemic Justification," *The Monist* Reprinted in Conee and Feldman (2004), pp. 37–52.

Conee, E. (1992), "The Truth Connection," *Philosophy and Phenomenological Research* Reprinted with an "Afterword" in Conee and Feldman (2004), pp. 242–258.

Conee, E. and Feldman, R. (1985), "Evidentialism," *Philosophical Studies* 48, reprinted with an "Afterword" in Conee and Feldman (2004), pp. 83–107.

Conee, E. and Feldman, R. (1998), "The Generality Problem for Reliabilism," *Philosophical Studies* (89), pp. 1–29.

Conee, E. and Feldman, R. (2001), "Internalism Defended," in H. Kornblith (ed.), *Epistemology: Internalism and Externalism*, Blackwell, pp. 231–260.

Conee, E. and Feldman, R. (2004), *Evidentialism*, Oxford University Press, New York.

Davies, M. and Humberstone, I. (1981), "Two Notions of Necessity," *Philosophical Studies* 58, pp. 1–30.

Evans, G. (1977), "Reference and Contingency," *The Monist* 62, pp. 161–89.

Feldman, R. (1988), "Having Evidence," in D. Austin (ed.), *Philosophical Analysis*, Kluwer, reprinted with an "Afterword" in Conee and Feldman (2004), pp. 219–241.

Feldman, R. (1989), "Foley's Subjective Foundationalism," *Philosophy and Phenomenological Research* 50(1), pp. 149–58.

Feldman, R. (1995), "Authoritarian Epistemology," *Philosophical Topics* Reprinted in Conee and Feldman (2004), pp. 111–134.

Feldman, R. (2006), "Naturalized Epistemology," in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/archives/fall2006/entries/epistemology-naturalized/. First published in 2001.

Foley, R. (1987), *The Theory of Epistemic Rationality*, Harvard University Press.

Foley, R. (1989), "Reply to Alston, Feldman and Swain," *Philosophy and Phenomenological Research* 50(1), pp. 169–188.

Foley, R. (1993), *Working Without A Net*, Oxford University Press.

Goldman, A. (1979), "What is Justified Belief?" in G. Pappas (ed.), *Justification and Knowledge*, Reidel, Dordrecht, pp. 1–23.

Goldman, A. (1986), *Epistemology and Cognition*, Harvard University Press, Cambridge, MA.

Goldman, A. (1988), "Strong and Weak Justification," *Philosophical Perspectives* 2, pp. 51–69.

Goldman, A. (1992), "Epistemic Folkways and Scientific Epistemology," in *Liaisons: Philosophy Meets the Cognitive and Social Sciences*, MIT Press.

Hawthorne, J. (2002), "Deeply Contingent A Priori Knowledge," *Philosophy and Phenomenological Research* pp. 247–269.

Hintikka, J. (1962), *Knowledge and Belief*, Cornell University Press.

Jackson, F. (1998), *From Metaphysics to Ethics*, Oxford University Press.

Kaplan, D. (1979), "Dthat," in P. Cole (ed.), *Syntax and Semantics*, Academic Press.

Kaplan, D. (1989), "Demonstratives," in H. W. J. Almog, J. Perry (ed.), *Themes from Kaplan*, Oxford University Press, Oxford.

Keynes, J. (1921), *A Treatise on Probability*, Macmillian.

Kornblith, H. (ed.) (2001), *Epistemology: Internalism and Externalism*, Blackwell.

Majors, B. and Sawyer, S. (2005), "The Epistemological Argument for Content Externalism," *Philosophical Perspectives* 19, pp. 257–280.

Plantinga, A. (1996), "Respondeo ad Feldman," in J. Kvanvig (ed.), *Warrant in Contemporary Epistemology: Essays in Honor of Plantinga's Theory of Knowledge*, Rowman and Littlefield, pp. 357–361.

Ramsey, F. (1926), "Truth and Probability," In Ramsey (1931), pp. 156–198.

Ramsey, F. (1931), *The Foundations of Mathematics and Other Logical Essays*, Routledge and Kegan Paul.

Sosa, E. (1991), *Knowledge in Perspective*, Cambridge University Press, New York.

Sosa, E. (1993), "Proper Functionalism and Virtue Epistemology," *Noûs* 27(1), pp. 51–65.

Sosa, E. (2007), "Experimental philosophy and philosophical intuition," *Philosophical Studies* (132), pp. 99–107.

Stalnaker, R. (1978), "Assertion," in P. Cole (ed.), *Syntax and Semantics, vol. 9: Pragmatics*, Academic Press.

Stalnaker, R. (1984), *Inquiry*, Bradford Books, MIT Press.

Stove, D. C. (1986), *The Rationality of Induction*, Oxford University Press.

Sutton, J. (2007), *Without Justification*, Massachussetts Institute of Technology.

Weatherson, B. (2005), "Scepticism, Rationalism and Empiricism," in T. S. Gendler and J. Hawthorne (eds.), *Oxford Studies in Epistemology*, Oxford University Press, volume 1, pp. 311–31.

White, R. (2006), "Problems for Dogmatism," *Philosophical Studies* 131(3), pp. 525–557.

Williamson, T. (2000), *Knowledge and Its Limits*, Oxford University Press.