The Case for Infallibilism

Julien Dutant*

* University of Geneva, Switzerland: julien.dutant@lettres.unige.ch
http://julien.dutant.free.fr/

Abstract. Infallibilism is the claim that knowledge requires that one satisfies some infallibility condition. I spell out three distinct such conditions: epistemic, evidential and modal infallibility. Epistemic infallibility turns out to be simply a consequence of epistemic closure, and is not infallibilist in any relevant sense. Evidential infallibilism is unwarranted but it is not an satisfactory characterization of the infallibilist intuition. Modal infallibility, by contrast, captures the core infallibilist intuition, and I argue that it is required to solve the Gettier problem and account for lottery cases. Finally, I discuss whether modal infallibilism has sceptical consequences and argue that it is an open question whose answer depends on one's account of alethic possibility.

1 Introduction

Infallibilism is the claim that knowledge requires that one satisfies some infallibility condition. Although the term and its opposite, "fallibilism", are often used in the epistemology literature, they are rarely carefully defined. In fact, they have received such a variety of meanings that one finds such contradictory statements as the following:

> "Fallibilism is virtually endorsed by all epistemologists"[1]
> "Epistemologists whose substantive theories of warrant differ dramatically seem to believe that the Gettier Problem can be solved only if a belief cannot be at once warranted and false, which is [infallibilism]. Such is the standard view."[2]

Faced with such statements, one is tempted to dismiss any discussion about infallibilism as a merely terminological dispute. However, I will argue here that there is a non-trivial infallibility condition that best captures the core infallibilist intuition, and that knowledge is subject to such a condition. Thus some substantial version of infallibilism is right. But it should be noted that that is compatible with "fallibilism" being true under some definitions of the term, for instance if "fallibilism" is understood as the (almost trivial) thesis that it is logically possible that most of our beliefs are false.[3]

---

1 Reed (2002: 143). Similar statements are made by Cohen (1988: 91), Siegel (1997: 164), Brueckner (2005: 384), and Dougherty and Rysiew (forthcoming).
2 D. and F. Howard-Synder and Feit (2003).
3 Lehrer (1990: 47) uses "fallibilism" in that way.

Latin Meeting in Analytic Philosophy Genova 20-22 September 2007
CEUR-WS Proceedings ISSN 1613-0073. pp.59-84

Infallibilism is often rejected from the start because it is thought to lead to scepticism. For instance, Baron Reed says that:

"Fallibilism is the philosophical view that conjoins two apparently obvious claims. On one hand, we are *fallible*. We make mistakes – sometimes even about the most evident things. But, on the other hand, we also have quite a bit of *knowledge*. Despite our tendency to get things wrong occasionally, we get it right much more of the time."[4]

Of course, infallibilism would straightforwardly lead to skepticism if it was the claim that one has knowledge only if one has never had any false belief. But nobody has ever defended such a strong infallibility condition. It must nevertheless be granted that some skeptical arguments have been based on infallibility requirements.[5] But as I will argue here, the relation between infallibilism and skepticism is not straightforward.

I will discuss three notions of infallibilism that can be found in the literature. In section 2, I will define *epistemic* infallibilism and show that it is just a consequence of closure, that it is neutral with respect to scepticism, and that it should not be called "infallibilism" at all. In section 3. I will define *evidential* infallibilism and agree with most contemporary epistemologists that it straightforwardly leads to scepticism and that it should be rejected; however, I will argue that it fails to provide a satisfactory account of the infallibilist intuition. In section 4. I will define *modal* infallibilism, and argue that it captures the core infallibilist intuition, that it should be accepted in order to solve the Gettier problem and to account for our ignorance in lottery cases, and that whether it leads to scepticism depends on further non-trivial issues about the metaphysics of possibility and the semantics of modals. I conclude by saying that modal infallibilism might turn out to warrant scepticism, but that it is an open question whether it does so, and that at any rate we will have to live with it.

2 Epistemic infallibilism and epistemic closure

2.1 Certainty and ruling out possibilities of error

The basic task of epistemology is to say what knowledge requires above true belief. True belief is clearly insufficient for knowledge: if one guesses rightly whether a flipped coin will land on heads, one does not thereby know that it will land on heads.[6]

Descartes famously argued that knowledge required absolute certainty:

"my reason convinces me that I ought not the less carefully to withhold belief from what is not entirely certain and indubitable, than from what is manifestly false"[7]

---

4 Reed (2002: 143).

5 See sections 2.4, 3.3 and 4.4 below.

6 Sartwell (1992) argues that knowledge is just true belief, but I will leave that discussion aside.

7 (*Med. First Phil*. I) Here Descartes only claims that *rational belief* requires certainty, but it is clear from the rest of the *Meditations* that Descartes takes it to be a condition on knowledge.

Several philosophers have followed Descartes in claiming that certainty was a condition on knowledge.[8] Unger has defended the certainty requirement on the basis of the infelicity on such claims as (1):

(1) He really *knows* that it was raining, but he isn't certain of it[9]

Now the certainty requirement is often taken for a requirement for infallibility.[10] However, that is far from obvious. For instance, Unger construes "certainty" as maximal confidence.[11] If certainty means that one has no doubts and is maximally confident about the truth of some proposition, then certainty is compatible with one's having a false belief, which is a case of fallibility on all accounts. But it is also clear that *psychological* certainty is not going to help with the account of knowledge: a perfectly confident lucky guess that the coin will land on heads is not knowledge either. Thus a more objective or normative notion of certainty is generally assumed, such as being *legitimately* free from doubt. Descartes had probably such a notion in mind. One way to cash out that requirement has been suggested by Pritchard: [12]

Ruling-out. One knows that *p* only if one is able to rule out all possibilities of error associated with *p*. (RO)

The thought goes as follows: if one is not able to rule out a given error possibility, then for all one knows, that possibility might obtain. Thus it would be legitimate for one to doubt whether one is free from error. Therefore one cannot legitimately be certain.

The (RO) requirement is intuitive. Suppose Bob has sighted Ann in the library, but Ann has a twin that Bob cannot distinguish from her. If Bob had sighted Ann's twin, he would not have noticed any difference. Intuitively, if he has no further evidence to rule out the possibility that he saw Ann's twin, he does not know that Ann is in the library, even if that is in fact true. The intuition can be defended by an appeal to considerations of luck: if Bob's epistemic position is compatible with either Ann being there or her not being there, then it is a matter of luck whether his belief that she is there is right. And that kind of luck prevents knowledge.[13] The situation is analogous to one's flipping a coin and Bob luckily guessing which side the coin landed on.

Moreover, the (*ruling-out*) requirement is naturally understood as an infallibility condition. Lewis defended it as such:

"It seems as if knowledge must be by definition infallible. It you claim that S knows that *p*, and yet you grant that S cannot eliminate a certain possibility in which not-*p*, it certainly seems as if you have granted that *S* does not after all know that *p*. To speak of fallible knowledge, of knowledge despite uneliminated possiblities of error, just *sounds* contradictory."[14]

---

8 Notably G.E. Moore (1959).

9 Unger (1971: 216, italics in the original). Unger argues that the modifier "really" and the stress on "knows" prevent loose use of the verb, thereby allowing us to realize that assertions like (1) are contradictory.

10 See *e.g.* Pritchard (2005: 18-21).

11 Unger (1974: 217).

12 Pritchard (2005:17).

13 See Pritchard (2005: 17-18).

14 Lewis (1996: 549).

2.2 The epistemic construal of ruling out possibilities of error

It is unclear what the (*Ruling-out*) requirement amounts to, as long as we are not clear about what "to rule out" and "possibilities of error" mean. There are two definitions of "possiblities of error":

W is a possibility of error for S with respect to *p* iff: if *W* obtained, *p* would be false. (PE)

W is a possibility of error for S with respect to *p* iff: if *W* obtained, S would not know whether *p*. (PE_2)

Possibilities of error (*Ws*) are sets of possible worlds. The conditional "if *W* obtained, *p* would be false" is meant to be a strict implication: for any world *w* in *W*, *p* is false in *w*. By contrast, the objects of knowledge (*p*s) are not construed as set of worlds but as sentence-like propositions. For instance, one may know that (*p*) Hespherus shines while ignoring that (*q*) Phosphorus shines. Yet if Hesperus is Phosphorus, and identity is necessary, the sets of worlds in which Hespherus shines is just the set of worlds in which Phosphorus shines. (A similar problem arises with any pair of true logical or mathematical propositions.) To allow one to know that Hespherus shines while ignoring that Phosphorus does, I take the objects of knowledge to be sentence-like objects associated with a set of worlds as their truth-condition.

Two remarks. First, assuming that knowledge entails truth, any possibility of error in the sense of (PE) is also a possibility of error in the sense of (PE_2). (PE_2) adds further possibilities of error: those in which *p* is true but not believed, or not known for some other reason, and those in which *p* is false but it is not known that not-*p*. Thus the infallibility condition built with (PE_2) is strictly stronger than one built with (PE). Second, (PE) implies that there is no possibility of error associated to beliefs in necessarily true propositions. That would wrongly classify a lucky guess that a five-digit number is prime as an infallible belief.[15] An obvious fix would be to include situations in which one has a different belief about whether *p* and that belief is false:

w is a possibility of error for S with respect to *p* iff: if *w* obtained, S's belief about whether *p* would be false. (PE')

For instance, the situation in which one wrongly believes that 2+2=5 would count as a possibility of error with respect to one's belief that 2+2=4. But to keep things simple I will leave the case of necessary true propositions aside for the time being. (We will return to it in section 4.1.)

Now what does "ruling out" possibilities of error consist in? A common construal is that one rules out an error possibility if and only if one knows it to be false: [16]

15 See Reed (2002).

16 See Dretkse (1981: 371): "In saying that [someone] is in a position to exclude these possibilities I mean that his evidence or justification for thinking these alternatives are not the case must be good enough to say that he knows they are not the case." See also Pritchard (2005: 25) and Stanley (2005: 130).

Epistemic Ruling Out. S rules out an error-possibility *w* iff S knows that not-*w*.

The formulation is problematic though, because the objects of knowledge are propositions (individuated in a sentence-like manner) and not sets of worlds. Suppose Bob knows that Hespherus shines but doesn't know that Hespherus is Phosphorus. Does Bob rule out a possibility *w* in which that planet does not shine? Depending on how *w* is presented (as a dark-Hespherus versus a dark-Phosphorus situation), the answer is different. So let us say that S *weakly* rules out a possibility *w* iff there is *some* mode of presentation under which S knows that *w* does not obtain:

Weak epistemic ruling out. S rules out an error-possibility *w* iff: there is *some* proposition *m* such that *m* is true iff *w* obtains, and S knows that not-*m*.

Thus Bob weakly rules out the dark-Phosphorus situation, because that situation obtains iff Hespherus does not shine, and Bob knows that Hespherus shines. In the following, I will drop the mode-of-presentation qualification whenever it can be safely ignored. So "S knows that not-*w*" should be understood as "there is some proposition *m* such that *m* iff *w*, and S knows that not-*m*".

Now the definition implies that one is *able* to rule out *w* just if one is *able* to know that not-*w*. But what does *being able* to know amounts to? The most natural way to construe that idea is Williamson's notion of *being in a position to know*:

"To be in a position to know *p*, it is neither necessary to know *p* nor sufficient to be physically and psychologically able of knowing *p*. No obstacle must block one's path to knowing *p*. If one is in a position to know *p*, and one has done what one is in a position to do to decide whether *p* is true, then one does know *p*."[17]

The characterization is somewhat vague but it will do for our present purposes. Typically, if one is in a position to know *p*, then if one asked oneself whether *p*, one would come to know that *p*. Thus one is able to rule out an error possibility iff: were one to consider that possibility, one would know that it does not obtain.

2.3 Epistemic infallibility is a consequence of epistemic closure

With these epistemic definitions of ruling out in hand, we can state two infallibility conditions, based on (PE) and (PE_2), respectively:

Basic Epistemic Ruling Out. S knows that *p* only if S is in position to know that every possibility *w* in which *p* is false does not obtain. (BERO)

Reflective Epistemic Ruling Out. S knows that *p* only if S is in position to know that every possibility in which S does not know *p* does not obtain. (RERO)

Let me first discuss (RERO). It should be pretty clear that (RERO) is roughly equivalent to the well known *KK principle* according to which one knows only if one is in position to know that one knows.[18] Let *W* be the disjunction of possibilities in

17 Williamson (2000: 95).

18 Prichard (1950: 86), Hintikka (1962), Chisholm (1977: 116). See Williamson (2000: 114) for the present formulation.

which one does not know *p*. Then not-*W* implies that one knows that *p*. Given (RERO), if one knows *p*, then for any *w* in *W*, one is in position to know that not-*w*. Given some background assumptions, one is thereby in position to know that not-*W*, and thereby in position to know that one knows *p*.[19]

Carrier's (1993) definition of infallibilism substantially amounts to (RERO). Accordingly, he takes fallibilism to be the claim that one knows without being in position to know that one knows. Now I agree that both (RERO) and (KK) are unwarranted. It seems to me possible that one knows without knowing that one knows, and at any rate that is not the kind of infallibilism I will be arguing for here. But as Reed (2002: 148) argues, Carrier's characterization of the fallibilism / infallibilism debate relies on confusing orders of knowledge. Intuitively, in order to know that *p*, one has to rule out the possibility *that p is false*. But ruling out the possibility that *one does not know p* is what is required in order to know *that one knows p*. Accordingly, (RERO) should be rejected, but it should not be taken as a characterization of infallibilism.[20]

Now let us turn to (BERO). (BERO) is in fact a consequence of another well-know epistemic principle, namely *epistemic closure*, according to which if one knows that *p* and is in position to know that *p* implies *q*, then one is in position to know that *q*. That is easily shown. Suppose S knows *p*. Let *w* be any possibility in which *p* is false. Then the proposition "*w* and not-*p*" is a mode of presentation of that possibility. Assuming S knows basic (classical) logic, namely that *p* implies that it is not the case that *w* and not-*p*, S is in position to know that if *p*, then "*w* and not-*p*" is false. By epistemic closure, S is in position to know that "*w* and not *p*" is false. By the definition of weakly ruling out, S is able to rule out "w and not *p*". So (BERO) is true.[21]

Epistemic closure is extremely plausible. Suppose that one knows (2):

(2) There are two apples on the table

---

19 The background assumptions are: (a) if one knows for each *w* that it does not obtain, then one knows that the disjunction of all *w* does not obtain. (a) assumes multiple-premise closure, which has been rejected by some: see Hawthorne (2004:46-50) for a discussion of multiple-premise closure. (b) if one is in position to know that not-*W*, where not-*W* implies that one knows *p*, then one is in position to know that one knows that *p*. That assumes a rough version of epistemic closure (namely: if one knows that *p* and *p* implies *q* then one is in position to know *q*) that might fail in various ways. For instance, one might ignore that the possibilities one rules out are *all* the possibilities of error. Thus one might fail to be in a position to know that not-*W* implies that S knows *p*. So (RERO) does not quite entail (KK). The converse holds, though: if one knows that one knows *p*, then there is a proposition which is true iff *W* (namely the proposition *that one does not know p*), such that one knows that it is false. So one weakly rules out all possibilities in which one does not know that *p*. That is why I say that the principles are merely roughly equivalent.

20 Another way to see that Carrier's definition of infallibilism is too strong is that it would classify as "fallible" someone whose beliefs are all true and could not have been mistaken in any way, but who lacks second-order beliefs about whether she knows.

21 Given that, one might be surprised that Pritchard (2005: 27) takes epistemic closure to be *weaker* that the epistemic infallibility condition. But that is due to the fact that the present infallibility condition is slightly weaker than Pritchard's, because it relies on "being in position to know" where Pritchard uses "know", and on weak ruling out.

And one knows that (2) entails (3):

(3) There is one apple on the table

Given one knows (2), (2) is true. Given that (3) is a logical consequence of (2), then it just cannot be the case that (3) is false. And one can very well know all of this. If that is so, it is hard to see how one could fail to know that (3) is true. But that is what the denial of closure would make possible.

Dretske and Nozick have famously argued that epistemic closure fails.[22] According to Dretske, for instance, the following is possible:

(4) Chris knows about some animal in a cage that it is a zebra, but he does not know that it is not a cleverly painted mule.[23]

Their rejection is based on well-motivated and powerful analyses of knowledge to which I cannot do justice here. But as Hawthorne argues, rejecting closure has many counter-intuitive consequences.[24] For instance, it seems hard to reject the two following closure principles:

Addition closure. If S knows *p*, then one is in position to know (*p* or *q*). (AC)

Equivalence closure. If S knows *p*, and if S is in position to know that *p* and *q* are equivalent, then S is in position to know *q*. (EC)

But as Hawthorne points out, that leads to the conclusion that Dretske and Nozick want to avoid. Suppose Chris knows (5):

(5) That is a zebra.

By (AC), Chris is in position to know (6):

(6) That is a zebra or that is a zebra which is not a painted mule.

But (6) is equivalent to (7), and so by (EC) Chris is in position to know (7):

(7) It is not the case that that is not a zebra and that is not a zebra which is not a painted mule.

Thus Dretske and Nozick have to give up such intuitive principles as addition closure and equivalence closure. Epistemic closure should be taken as a *prima facie* constraint on accounts of knowledge that we should not drop unless some substantial revision of our intuitive understanding of knowledge appears to be inevitable.

Let us take stock. I have argued that the following version of infallibilism is simply a consequence of epistemic closure:

(BERO) S knows that *p* only if S is in position to know that every possibility *w* in which *p* is false does not obtain.

(BERO) S knows that *p* only if S is in position to know that every possibility *w* in which *p* is false does not obtain.

Call that *epistemic infallibilism*. It is a non-trivial condition, since it excludes some accounts of knowledge (Dretske's and Nozick's). However, since most epistemologists want to hold closure, they in fact endorse epistemic infallibilism.

---

22 Dretske (1970), Nozick (1981).

23 Drestke (1970: 1016). Dretske assumes, as we will, that Chris knows that if the animal is a mule then it is not a zebra. Schaffer (2005) offers a related but distinct defense of the truth of such statements. However, Schaffer argues that knowledge is a three-place relation between a knower and *a pair* of contrastive propositions (*e.g.* S knows that *p* rather than *q*), and that once the implicit contrast propositions are restored sentence (4) turns out not to be a counter-instance to closure.

24 See Hawthorne (2004: 38-46).

But it is doubtful whether epistemic infallibilism should be called "infallibilism" at all. By all accounts, the analysis of knowledge as being just true belief (Sartwell 1992) is a fallibilist account: a lucky guess is a true belief that could easily have been false. But true belief does satisfy (BERO): if one has the true belief that *p*, then any corresponding possiblity of error *w* is false, and thus one is in position to have the true belief that *w* is false. (Similarly, true belief satisfies closure.) Given that, it is striking that some take epistemic "infallibilism" to lead to skepticism. In the next section, I will discuss some reasons what one might think so, and argue that they are misleading.

2.4 Epistemic infallibilism is neutral with respect to skepticism

As Pritchard notes, epistemic infallibilism is crucial to classical skeptical arguments such as the following:[25]

(8) I do not know that I am not a handless brain in a vat.

(9) If I do not know that I am not a handless brain in a vat, then I do not know that I have hands.

(10) Therefore, I do not know that I have hands.

Premise (9) can be derived both from epistemic infallibilism and epistemic closure. However, Moore famously showed that premise (9) can just as well be used in an *anti*-sceptical argument:

(11) I know that I have hands.

(12) If I know that I have hands, then I know that I am not a handless brain in a vat.

(13) Therefore, I know that I am not a handless brain in a vat.

Premise (12) is equivalent to premise (9). What they say is that knowledge that one has hands and knowledge that one is not a handless brain in a vat go together: either one has both, or one has none.[26] But they do not tell us which way of using the argument is the right one. And there is no way in which a sceptic can use epistemic infallibilism in order to defend premise (8): instead, she would have either to appeal to intuition or to put forward an independent argument.[27] (As we will see in section 4.5, such an argument can rely on *other* infallibility conditions.) Thus epistemic infallibilism is neutral with respect to scepticism.

Similar things can be said with respect to a slightly different formulation of the fallibilism / infallibilism distinction. Rysiew has argued one needed to be a fallibilist in order to avoid scepticism, and that fallibilism consisted in accepting such claims as (14), which he calls "concessive knowledge attributions":

(14) Chris knows that Harry is a zebra, but it is possible for Chris that Harry is a painted mule.

---

25 Pritchard (2005: 25-30).

26 Of course in some non-standards situations (*e.g.* if one is handless) one could know that one is not a handless brain in a vat without knowing that one has hands. But I am assuming a standard situation here.

27 That is pretty clear in Pritchard (2005: 24): even though Pritchard sometimes talks as if epistemic infallibilism was the source of sceptical intuitions, he does in fact appeal to considerations of subjective indistinguishability to explain why premise (8) seems plausible.

Rysiew claims that concessive knowledge attributions can be true even though, for pragmatic reasons, they are infelicitous.[28]

Now Stanley has argued that "it is possible" is naturally interpreted as *epistemic* possibility and along the following lines: [29]

Epistemic possibility as knowledge. "It is possible for S that *p*" is true iff what S knows does not entail, in a manner that is obvious to S, not-*p*. (EPK)

A similar account can be given for epistemic readings of "might":

Epistemic "might" as knowledge. "It might be that *p*" uttered by S is true iff what S knows does not entail, in a manner that is obvious to S, not-*p*. (EMK)

Thus statements like (15) would also be concessive knowledge attributions:

(15) I know that Harry is a zebra, but it might be a painted mule.

Given these understandings of "it is possible" and "might", epistemic infallibilism can be restated in the following ways:

(16) If one knows that *p*, and *p* obviously entails *q*, then it is not possible for one that *q*.

(17) If I know that *p*, and *p* obviously entails *q*, then it is false that it might be that *q*.

Accepting (16) and (17) amounts to rejecting concessive knowledge attributions. But as Stanley notes, that gives no weight whatsoever to sceptical arguments. If Chris knows that Harry is a zebra, then something that Chris knows – namely, that Harry is a zebra – obviously entails that Harry is not a painted mule. So it is not possible for Chris that Harry is a zebra.[30] So avoiding scepticism need not lead one to accept concessive knowledge attributions, nor to reject epistemic infallibilism.[31]

A second way in which one might think that epistemic infallibilism leads to scepticism is that it seems to forbid inductive knowledge. Suppose one knows of the *n* observed apples that they all have seeds. Then what one knows does not rule out that the remaining apples have seeds. So it seems that epistemic infallibilism implies that one cannot thereby know that all apples have seeds. Since a vast amount of our putative knowledge is based on inductive evidence, infallibilism would lead to scepticism.

However, the worry is misplaced as well. Let t_0 be the time just before one draws the inductive inference, and t_1 the time at which one has inferred that all apples are seed. What epistemic infallibilism implies is just this: at t_1, one knows that all apples have seeds only if one is able to rule out any possibility in which some apple does not have seeds. It does not imply that one should be able to rule them out at t_0. Thus epistemic infallibilism just says that *if that inductive inference yields knowl-*

---

28 Rysiew (2001 : 492-494).

29 Stanley (2005 : 128).

30 Stanley (2005 : 130).

31 Dougherty and Rysiew (forthcoming) have objected to Stanley's account of epistemic possiblity. They defend Rysiew's (2001) original account, according to which "it is possible for S that *p*" is true iff S's evidence does not logically entail not-*p*. On that reading, concessive knowledge attributions do not amount to a denial of *epistemic* infallibilism, but to a denial of *evidential* infallibilism, which will be discussed in section 3.

edge, then once the inference is made one is able to rule out alternatives to the general proposition. But it is silent on whether inductive inference yields knowledge.

3 Evidential infallibilism

Epistemic infallibilism is a consequence of epistemic closure and is thus a fairly plausible condition on knowledge. However, it fails to deal with our original problem, which was to say what knowledge required above true belief. That is so because epistemic infallibility is itself formulated in terms of knowledge: very roughly, it amounts to saying that one knows that *p* only if one knows that not-*p* is false. I will now turn to a seemingly more promising account of infallibilism that is widely found in the literature, which I will call *evidential infallibilism*. Evidential fallibilism is rightly thought to lead to scepticism, and is widely rejected. I will reject it as well, but I will argue that it is not a useful characterization of the infallibilist intuition either.

3.1 The evidential-infallibility condition

When infallibilism is not defined as ruling out possibilities of error, it is often defined along the following lines: [32]

One knows that *p* on the basis of evidence *e* only if *e* logically entails *p*.

What does "evidence" means here? Since evidence is taken to enter in logical relations, it has to be propositional. So even though we sometimes speak of a bloodied knife as evidence, strictly speaking the evidence would be *that the knife was bloodied*, or *that there was a bloodied knife in the defendant's car*, and so on.[33] Moreover, we are talking about evidence *one has*, and that implies at least that the relevant propositions are believed. Even though the fact that there is a bloodied knife in the defendant's car is evidence against him, it cannot be the basis of one's knowledge that he is guilty unless one is aware of it. So evidence is at least believed propositions.[34]

Morever, there are good reasons to think that propositional evidence cannot be the basis of knowledge if it is not itself *known*.[35] Familiar Gettier cases give intuitive support to that idea. Suppose that Denise has a justified but false belief that her col-

---

32 See *e.g.* Feldman (1981 : 266), Cohen (1988 : 91), Fogelin (1994 : 88-89), Jeshion (2000), Audi (1998 : 82, 84, 255, 292-294), Brueckner (2005 : 384), Stanley (2005 : 127), Dougherty and Rysiew (forthcoming).

33 My claim here is just that the evidence *mentioned in the evidential infallibility condition* is propositional. I am not arguing for the broader claim that all evidence is propositional. See Unger (1975 : 204-206), Davidson (1986), and Williamson (2000, sec. 9.5) for such arguments, and Moser (1989 : 47-125) and Millar (1991) for opposing views.

34 Strictly speaking, the foregoing argument requires only that evidence be *entertained*, not believed. But that distinction will not matter for present purposes.

35 Again, I am only claiming that the evidence *mentioned in the evidential-infallibility condition* has to be knowledge, because if it is not it cannot ground knowledge. However, see Unger (1975 : 206-214), Hyman (1999) and Williamson (2000, sec. 9.6) for arguments that all propositional evidence has to be knowledge.

league Eric has a car – for instance, Erik has had a car for ages and she saw it on many occasions, but unbeknownst to her, he just sold it. From that she infers that someone in her office has a car. As it turns out, that is true because another colleague has a car – but she is also unaware of that. Intuitively, even though her belief is the logical consequence of a justifiably believed proposition, it is not knowledge. It is just a matter of luck that she gets things right: if the other colleague had not had a car, she would have been mistaken. By contrast, if she *knew* that Erik had a car (and thus he would have one), then she could know on that basis that someone in her office has a car. (That just follows from epistemic closure.) Thus evidence can be the basis of one's knowledge only if it is a proposition one knows to be true.[36]

We can thus reformulate the infallibility condition as follows:

Evidential infallibilism. S knows p on the basis of e only if S knows e and e logically entails p. (EF)

The basing relation is hard to define. A sufficient condition is the following: if S has competently deduced p from e, then S's belief that p is based (at least) on e. But there might be subtler ways of basing, *e.g.* if one's belief that e supports one belief that p in the following sense: one would cease to believe that p if one ceased to believe that e. However, we can focus on the inference case for the present discussion.

3.2 Why evidential infallibilism is not a satisfying characterisation of infallibilism

Given the reformulation (EF), it is clear that evidential infallibilism cannot be a *general* condition on knowledge unless it is a trivial one. That results from the following dilemma:

Either one allows p to be among one's evidence for knowing that p, and then evidential infallibilism is trivially true. That is so because it is always the case that if S knows p then S knows p and p entails p.[37]

Or one does not allow p to be among the evidence for one's knowing that p, and then a regress is generated that has to stop at some knowledge that does not satisfy condition (EF). If p is not among the one's evidence for knowing that p, then one's evidence consists in some further knowledge. If that further knowledge is itself based on evidence, then it requires yet some further knowledge. And so on until we reach items of knowledge that are not based on propositional evidence, and thus to which condition (EF) does not apply.

The second horn of the dilemma relies on several assumptions. First, I am assuming that circular justification (p is known on the basis of q which is known on the basis of p) substantially amounts to the first horn, that is allowing p to be part of one's evidence for p. Second, I am assuming that if the regress if infinite, knowledge is impossible.[38]

---

36 See Williamson (2000, sec. 9.6) for further arguments for the claim that one's evidence should just be one's knowledge.

37 On such a reading, (EF) boils down to a trivial case of epistemic infallibilism, namely S knows that p only if S knows that p.

38 See Klein (1998) for the opposite view.

Consequently, people who defend evidential infallibilism on the second reading turn out to be a specific set of views: namely, the sceptics who rely on Agrippa's trilemma to argue that knowledge is impossible,[39] and the foundationalists who take all knowledge to be either part of a special set of beliefs that are not themselves based on propositional evidence (call them *foundational beliefs*) or deduced from such beliefs. Let us call the later view *deductive foundationalism*.

Granted, deductive foundationalism has been an influential view. Descartes' own view is plausibly an instance of it: knowledge is either the intuition of some basic truths, or deduced from such intuitions.[40] Yet evidential infallibilism only characterizes half of the view: namely, the part concerning the non-foundational beliefs. Consequently, the *denial* of evidential infallibilism is itself a limited view.[41] It focuses on beliefs that are based on known evidence. So the debate does not deal with general conditions of knowledge, but local ones.

One might think that the cases under discussion are nevertheless the vast majority of our putative knowledge. Many beliefs about empirical matters, it seems, are inferred from non-logically conclusive evidence. For instance, I might infer that my car is parked in front of the house on the basis of my belief that it was so five minutes ago. In that case, my premise is consistent with the conclusion's being false. However, it is unclear to what extent our beliefs are so organised. As several have pointed out, our beliefs typically fit into a *redundant* set of beliefs, in the sense that any belief of the set is the logical consequence of several other beliefs in the set.[42] For instance, my beliefs that there is an apple in front of me is related to my belief that it was there five seconds ago and did not move since, that if I looked in a mirror I would see an apple in front of me, that somebody beside me knows that there is an apple in front of me, and so on. It is unclear that our putative knowledge can be neatly organised in a pyramid-like structure, and if it cannot the beliefs that belong to redundant sets will not provide counter-instances to evidential infallibilism. As it turns out, the cases where one can clearly identify the premises on which a belief is based are very few.

For those reasons, the evidential formulation of the fallibilism/infallibilism debate is not fully satisfactory. One would like to have a characterization of the infallibilist's intuition that covers all putative kinds of knowledge. Yet it will be useful to discuss the restricted set of cases on which that debate turns, so let us turn to them.

---

39 Agrippa's trilemma against knowledge is the following argument : for any p, one knows that p only if one's belief that p is based on some further knowledge. Consequently, either one's chain of justifications has an arbitrary end, or it goes on infinitely, or it is circular. In the three cases one fails to know. See Sextus Empiricus (1933-49 : I, 95-101) and Williams (1991 : §2.4 and 2001, chap. 5).

40 It is not clear, however, whether Descartes' view amounts to saying that deduction yields knowledge only if the deduced belief is *logically entailed* by its basis. Descartes only demands that it be *evident* that the deduced belief is true if its basis is, and that condition might be slightly less stringent than logical entailement.

41 Since on the first horn of the dilemma evidential infallibilism is trivially true, we can safely assume that critics of evidential infallibilism rely on the second reading.

42 Williamson (2000 : 204), Reed (2002 : 146).

3.3 Evidential infallibilism is unwarranted

We have seen that evidential infallibilism is controversial on the following reading only: [43, 44]

Non-Trivial Evidential Infallibilism. One knows that p on the basis of e ($e \neq p$), only one knows e and e logically entails p. (EF)

Why one would want to accept evidential infallibilism? One can at least see three reasons. First, the *dialectical motivation*: suppose one claims that p. Then one can always be challenged with such questions as "Why do you think that p?", "How do you know that p?" or "What is the evidence for p?".[45] Simply restating "p" will not satisfactorily answer the question, since by asking such questions one's interlocutor implies that she suspends her belief in p. Thus one has to put forward *further* claims as propositional evidence for p. Yet those can be challenged if they do not logically entail that p: the interlocutor might object something along the lines of "Well, I grant e, but it is possible that e and not-p, so I do not see why I should believe that p". For instance, if one argues that all apples have seeds because all the apples one has observed have seeds, the interlocutor might object that even though all the apples that one has observed had seeds, it is possible that the unobserved ones do not, and thus refuses to believe that all apples have seeds. And thus the interlocutor might challenge the claim that one knows that all apple have seeds.

Second, evidential infallibilism provides a simple solution to the Gettier problem. Gettier showed that justified true belief is not sufficient for knowledge, and thus that justification was not the (only) thing that distinguishes knowledge from true belief. [46] However, as Gettier himself noted, his counterexamples can be constructed only if one's justification for p is compatible with p being false. But if one's justification is a known proposition that entails p, it is incompatible with p being false. (If one knows that e, then e is true, and if e entails p, then p is true as well). So evidential infallibility is a good candidate for being the third condition on knowledge.

---

43 As formulated, the condition concerns only those beliefs that are based on some propositional evidence. It is thus silent on putative foundational beliefs.

44 Recall (section 2.4) that Rysiew (2001) and Dougherty and Rysiew (forthcoming) take fallibilism to entail the truth of concessive knowledge attributions such as "Chris knows that it is a zebra but it is possible for him that it is a painted mule". Dougherty and Rysiew argue *contra* Stanley (2005) that the epistemic reading of "it is possible" should be rendered as follows: "it is possible for S that p" is true iff *S's evidence* does not entail that not-p. (By contrast, Stanley argues that "it is possible for S that p" is true iff *what S knows* does not entail not-p.) Thus their fallibilism is the rejection of non-trivial evidential fallibilism. As we will see below, I agree that non-trivial evidential infallibilism should be rejected (and so does Stanley). However, I leave it open whether concessive knowledge attributions should be understood as Stanley does (in which case I would agree with him that they are false, because they amount to the rejection of epistemic closure), or as Dougherty and Rysiew do (in which case I would agree with them that they are true, because they amount to the rejection of non-trivial evidential fallibilism).

45 Williamson (2000 : 188).

46 Gettier (1963).

Third, evidential infallibilism provides a good explanation of why epistemic closure holds. Whenever one knows p and one knows that p implies q, then what one knows entails that q. So, according to evidential infallibilism, one would know that q if one believed that q on the basis of one's knowledge that p and that p implies q.

However, note that the second and third points only suggest that evidential infallibility is *sufficient* for knowledge. Only the dialectical motivation suggests that it is necessary. But it is unclear that the dialectical constraints on arguments do apply to knowledge – in fact, even evidential foundationalists will have to argue in the end that regarding some claims (the foundational ones), the dialectical challenge is unwarranted.

Moreover, there are good reasons to think that evidential infallibility is not *necessary* for knowledge as evidential infallibilism requires. First, there are our intuitions on standard knowledge cases. Suppose one believes that there was a coin on the table five seconds ago on the basis of one's knowledge that there is a coin on the table now. What one knows is consistent with the falsity of that belief. However, intuitively, it seems possible for one to come to know that there was a coin on the table five second ago in such a way, at least in standard situations. Evidential infallibilists might reply that in order to come to know that, one would need to know that coins do not spontaneously appear or something similar, but it is doubtful that one needs to.

Second, evidential infallibilism has devastating sceptical consequences. As Vogel (1990) argues, evidential infallibilism rules out the possibility that we gain knowledge through induction. An inductive inference form e to p is an inference such that e does not logically entail p. If the evidential infallibility condition holds, no such inference can yield knowledge. Since a vast amount of our putative empirical knowledge is plausibly based on such inferences, evidential infallibilism entails a significant dose of skepticism.

To sum up, the arguments for evidential infallibilism are not strong, and its consequences counter-intuitive. That has lead many to reject it. I agree with them that evidential *fallibilism* is a desirable feature for an account of knowledge to have. However, I have also claimed that the evidential construal of the fallibilism / infallibilism debate is not satisfactory because it focuses on a restricted set of putative knowledge cases, and fails to account for the infallibilist's general view about knowledge.

4 Modal infallibilism

The last kind of infallibilism that I will discuss is based on an *alethic-modal* construal of possibilities of error. Roughly, the idea is that one knows only if one *could not* have been wrong about p. The possibility here is an alethic modality like logical, metaphysical or physical possibility, not an epistemic one as in closure-based infallibilism.[47] I will argue that the modal infallibility condition captures the infallibilist's intuition in its full generality, and that it is a requirement that any account of knowledge should incorporate in order to solve the Gettier problem and to account for our ignorance of lottery outcomes.

47 See section 2.4 above.

4.1 The modal construal of possibilities of error

A rough formulation of modal infallibilism is thus the following: [48]

Modal infallibilism. S knows that *p* only if S's belief that *p* could not have been wrong. (MI)

The condition has some initial intuitive appeal. If one guesses rightly that a flipped coin has landed on heads, one does not know that it has because it could have landed on tails, and one's belief would then have been wrong. However, the formulation is too restrictive. Suppose that one has seen the coin landing on heads. Then one's belief could have been wrong, in the following way: one could have failed to see it, and wrongly guessed that it landed on tails. Yet that does not prevent one from knowing that it landed heads based on one's seeing it. For *given that one had seen it*, one's belief about whether it landed on heads could not have been wrong. Thus the condition should only rule out the possibility that one be wrong given one's basis for one's belief:

S knows that *p* on the basis *b* only if S's belief that *p* on basis *b* could not have been wrong. (MI')

What is the basis of a belief? The need to take into account the basis of a belief in order to account for knowledge is widely recognized in epistemology. If one believes that there is a coin on the table because one saw it, what is relevant to that belief's being knowledge is whether one's perceptual capacities can be mistaken about coins, not whether one's guessing abilities can be mistaken about coins. I call bases for belief what is elsewhere called "justification", "evidence",[49] "methods",[50] or "cognitive processes".[51] Bases can be an inference, a memory, a perception, and so on.[52] The

---

48 Formulations of this kind can be found in Ayer (1956 : 54-56), BonJour (1985 : 26 and 1998 : 16), Alston (1992), and Lehrer (1974: 81 and 1990: 45). See also Reed (2002: 144).

49 Lewis (1996). What Lewis calls "evidence" is not the propositions that one knows or believes, but *the fact that one is in such-and-such cognitive state, i.e.* the fact that one has the experiences and memories one has. Thus it belongs to the same family of notions as what are called "bases" here.

50 Nozick (1981).

51 Goldman (1976, 1979).

52 To be slightly more precise, I take bases of belief to be events. Thus they are non-repeatable. If "cognitive processes" are understood as *types* of processes, *e.g.* remembering, what I call bases are tokens or instances of cognitive processes, *e.g.* a given memory episode. Formally, bases are represented by sets of worlds (coarse propositions): for instance, *b* can be the set of worlds in which *S had such and such memory episode at time t*. The basing relation is a causal relation: belief B is based on basis *b* iff event *b* caused state B in the normal way. The restriction "in the normal way" is needed to set aside non-standard causal links between cognitive processes and beliefs, *e.g.* one's cognitive process causing a modification on an electro-encephalogram display that one sees, which causes in turn one to have belief B.

individuation of bases is famously a vexed matter, to which I will return below.[53] For the time being, I will rely on an intuitive comprehension of the notion.

What the condition says is the following: one knows *p* on basis *b* only if it was not possible that *p* was false while S based is belief that *p* on *b*.[54] The relevant notion of possibility here is an *alethic* modality, not an epistemic one. The condition does not say that one knows *p* only if *given what one knows* it is not possible that S has *b* and *p* is false. It says that one knows *p* only if *in fact* it was not possible that S had *b* and *p* was false. Now, alethic modalities come in different species: logical possibility, metaphysical possibility, physical possibility, and so on. Our present characterization of modal infallibilism is neutral with respect to those. As we will see in the next section, that is in fact a welcomed feature.

Necessarily true propositions raise difficulties for (MI'). If I luckily guess that a five-digit number is a prime, my belief that it is a prime on the basis of guess could not have been false, simply because that number could not have failed to be prime. Yet we do not want to say that I could not have been wrong, nor that I know. One way to account for that is the following: on the same basis (guess), I could have come up with another belief (that it is *not* prime) which would have been mistaken. That is why I could have been wrong. Accordingly, the infallibility condition should be reformulated as follows:

S knows that *p* on the basis *b* only if S could not have had a wrong belief about whether *p* on basis *b*. (MI'')

But there are reasons to suspect that (MI'') will not do either. Suppose that I decided whether the number is prime on the basis of some systematic but absurd mathematical method (for instance, adding its digits and assuming that it is prime iff the sum of its digits is). Then the only belief I could have come up with on that basis about that particular number is the one I do have now, namely that it is prime, and that belief could not have been false. However, intuitively I could have been wrong because it is just a matter of luck that my method gave the right result about that number, for it gives wrong results about most other numbers. So it not sufficient to look whether the basis would have yielded a true belief about *p*. One should also look whether the basis would have yielded true beliefs with respect to *q*, *r*, …, where the latter are matters close to *p*. That is not easily done, and I will simply leave the case of necessarily true propositions aside here.[55]

4.3 Modal infallibilism captures the core infallibilist intuition

By contrast with evidential infallibilism, which is restricted to cases where one knows on the basis of propositional evidence (section 3.2), modal infallibilism is a

---

53 See Goldman (1976), Nozick (1981) for different attempts. Conee and Feldman (1998) and Williamson (2000, sec. 7.4) argue that the problem is intractable. I return to the issue in section 4.4.

54 In possible worlds terms : look at all possible worlds in which *S* has the same basis for his belief that *p* ; if in any of these worlds, *p* is false, then S does not know that *p*.

55 A more straightforward solution would be to allow oneself to *impossible* worlds and maintain (MI') as it is. But since that would require a non-standard semantics for alethic modality, I will not pursue that option here.

general condition on knowledge.[56] I will now argue that modal infallibilism captures the core infallibilist intuition. I will do so on two grounds: first, the hypothesis that Descartes endorsed modal infallibilism provides a fairly good account of his epistemology, which is the prototypical infallibilist position. Second, modal infallibilism explains how one would expect infallibilists to endorse evidential infallibilism.

The prototypical infallibilist position, namely Descartes', can be derived from modal infallibilism and a few assumptions. The crucial assumption is to take possibility to be *metaphysical* possibility: it is possible that *p* iff not-*p* is not a metaphysical truth. Now there are three kinds of knowledge in Descartes' view: (a) intuitive knowledge that I think, and that I exist, (b) intuitive knowledge of some metaphysical principles, such as that causes have at least as much reality as their effects, (c) knowledge deduced from other knowledge. Moreover, (d) beliefs that are merely based on perception are not knowledge. As we saw (section 3.2), Descartes' acceptance of evidential infallibilism would account for (c), and possibly for (d) if Descartes assumed that perceptual beliefs are based on propositional evidence, which is doubtful. By contrast, modal infallibilism accounts for all of (a) - (d). Why does Descartes take his own belief that he exists to be knowledge? A plausible answer is that he does so because such a belief could not fail to be true. In whatever situation in which I have the belief that I exist, that belief is true. So beliefs of type (a) satisfy the modal infallibility condition. Furthermore, beliefs of type (c) count as knowledge if epistemic closure holds. But as we will see in section (4.4), epistemic closure is a plausible consequence of modal infallibilism. So Descartes' acceptance of modal infallibilism would also explain (c).

Turning to (b), why does Descartes takes the intuition of metaphysical principles to be knowledge? If he was relying on evidential infallibilism, he would take himself to deduce them from further principles, but he clearly does not so. Instead he insists on the fact that one cannot doubt such principles as soon as one conceives them clearly and distinctly. But many have been puzzled why Descartes takes a purely psychological fact (clear and distinct perception) to yield knowledge. Modal infallibilism provides an explanation of why that could be so. Let *p* be one of the metaphysical principles. Descartes' view seems to be that (1) when one forms a belief about whether *p* on the basis of clear and distinct perception, one could not but form the belief that *p* is true, (2) since *p* is a metaphysical truth, it is impossible that *p* is false. From (1) and (2) it follows that (3) whenever one forms a belief about whether *p* on the basis of clear and distinct perception that *p*, one's belief is true. So one satisfies the modal infallibility condition with respect to *p*. What the psychological condition (clear and distinct perception) does is just to ensure that one could not have come up with the belief that not-*p*. By contrast, if one's belief that *p* was based on a guess, one could have made the opposite guess, and one would then have been wrong.[57]

Finally, modal infallibilism explains (d). Why does Descartes say that his experience does not allow him not know that he is not dreaming or deceived by an evil demon? It is metaphysically possible that Descartes has the experience he has and that

---

56 That is true only on the assumption that all beliefs have some basis. But the assumption is unproblematic. If there are cases where a belief just pops into existence, we will say that that belief is its own basis.

57 On the present view, Descartes would rely on something like (MI'') to characterize the fallibility of some beliefs in necessary truths. See section (4.2).

he is dreaming. Thus, if modal infallibilism holds, Descartes does not know that he is not dreaming on the basis of his experience.

One might object two things to that account of Descartes' position. First, Descartes does not think it metaphysically possible that he is deceived by an evil demon. He precisely argues that God's benevolence makes it impossible (*Med. First Phil.*, IV). Second, if the foregoing account of knowledge of metaphysical principles was right, Descartes would never refuse to count clear and distinct mathematical beliefs as knowledge – but he does (*Med. First Phil.*, I) .

The second objection does not goes through, because Descartes does not think it metaphysically impossible that true mathematical propositions are false. He refrains from saying that God could not have created different mathematical facts.[58]

The first objection can be answered in several ways. One is to move the whole discussion one order up: Descartes' position at the beginning of the *Meditations* would not be that he does not know that he is not deceived, but that he does not know *that he knows* that he is not deceived. Similarly, Descartes claims that an atheist mathematician has "*cognitio*" but not "*scientia*" of mathematical truths, where *scientia* seems to imply that one knows that one knows.[59] Another reply is to rely on the psychological mechanism of doubt: before he bases his belief that he is not deceived on the proof of the existence of God, Descartes's is liable to doubt whether he is deceived. Because of that, it was just as well possible that he formed the belief that he *is* deceived. And that belief would be wrong, assuming such deception is metaphysically impossible. So the experience-based belief that one is not dreaming is metaphysically fallible, even though it is metaphysically impossible that one is deceived.[60]

The second reason to think that modal infallibilism captures the core infallibilist intuition is that it explains why some infallibilists would endorse evidential infallibilism. The explanation is simply this: if the possibility in the modal infallibility condition is understood as logical possibility, then evidential infallibilism follow. It is pretty straightforward to see that. Suppose one's basis for q is one's knowledge that p, but that p does not logically entails q. Then typically it will be logically possible that one has that basis and q is false.[61] So the logical version of modal infal-

---

58 See *Med. First Phil.*, 5th and 6th replies.

59 See DeRose (1992).

60 A third objection would be that Descartes' metaphysical infallibilism would license a more permissive closure principle than deductive closure. And thus Descartes' alleged acceptance of metaphysical infallibilism would fail to explain why Descartes thinks that only deductive inferences yield knowledge. But as I said earlier (note 40), it is not clear whether Descartes' notion of "deduction" is restricted to logically valid inferences. And even if we assume that it is, it is plausible that Descartes does not clearly distinguish logical from metaphysical possibility. After all, the distinction was unclear until Kripke's work, and as far as I know Leibniz was the first to point it out. So it would not be a surprise if Descartes relied on metaphysical possibilities in his account of intuition, and logical ones in his account of deduction.

61 There are exceptions. For instance, let q be "S knows that p". Then if S's knowing that p is the basis for S's belief that S knows that p, it is logically impossible that S has that basis while S does not know that p. That kind of case cannot be blocked by saying that the basis should be a belief, and not knowledge, because (since by assumption one knows that p) the belief will have an infallible basis, and it will typically be the case

libilism implies that ones does know *p*. Conversely, one knows *p* on the basis of evidence *e* only if *e* entails *p*. Thus evidential infallibilism follows from modal infallibilism and the logical conception of possibility.[62]

To sum up, Descartes' view can be seen as an instance of modal infallibilism where possibility is understood as metaphysical possibility, and evidential infallibilism can be seen as an instance of modal infallibilism where possibility is understood as logical possibility. So modal infallibilism turns out to be the most general account of infallibilism, and arguably captures the core infallibilist idea.

4.4 Two arguments for modal infallibilism

Now why would one subscribe to the core infallibilist idea? There are two main arguments for it: first, it is needed to solve the Gettier problem, and second, it provides an account of lottery cases.

The first argument is the following: if our account of knowledge rejects modal infallibilism, then a Gettier case can be built against that account. Conversly, if no Gettier case can be built, then our account includes modal infallibilism. The argument does not show that modal infallibilism is *sufficient* to prevent Gettier cases, but it shows that it is necessary, and that is enough to show that if one thinks that Gettier subjects lack knowledge, one has to accept modal infallibilism.[63]

As is often noted, Gettier cases consist in a double luck situation.[64] S has a justification, or basis, that makes it very likely but not necessary that *p* is true. But the situation is just one in which one's basis would normally yield a false belief – that is the "bad luck" part. However, *p* turns out to be true in a manner unconnected to one's justification – that is the "good luck" part. For instance, one has a fairly reliable capacity to recognize sheeps, but one happens to be looking at a sheep-shaped rock in the distance – that is the "bad luck" part; in that situation one would normally acquire the *false* belief that there is a sheep there. However, there happens to be a sheep hidden behind the rock, and thus one's belief that there is a sheep there is true – that is the "good luck" part. It is thus a matter of luck that one's belief is true, and one does not know.

---

that one's having that basis is not only incompatible with *p* being false, but also with *one not knowing that p*. So the infallible basis for one's first order belief that *p* will typically be also an infallible basis for one's second-order belief that one knows that *p*. However, such cases are limited and can be left aside here. (Note in passing that the foregoing argument shows that modal infallibilism explains the intuitive appeal for the KK principle: if the basis of one's second-order belief that one knows *p* is the same as that of one's belief that *p*, then it will typically (but not always) be the case that that basis makes it necessary that one knows that *p*, and consequently the second-order belief will satisfy the infallibility condition as well.)

62 Reed (2002 : 145) says that (what is here called) modal infallibilism is equivalent to (what is here called) evidential infallibilism. But crucially, he fails to notice this is so only if possibility is understood as *logical* possibility.

63 By contrast, if one thinks that true belief is knowledge (Sartwell 1992) or that justified true belief is knowledge, then one does not have to accept modal infallibilism.

64 Zagzebski (1994, 1999), Pritchard (2005 : 149).

Now it is easily seen why rejecting modal infallibility exposes one to Gettier cases. Suppose it is argued that S has knowledge but that the basis of S's belief is fallible. Thus there are possible situations in which S has the same basis but her belief is false. Take one such situation, and make it so that the belief is true in a manner unconnectedd to S's basis. The resulting situation is one in which S satisfies all the conditions of the original account, and yet S does not have knowledge because it is just a matter of luck that her belief is true. The sheep case illustrates how the construction goes.

It might be objected that it is not guaranteed that in each case it will be possible to make it so that the belief is true in a manner unrelated to S's basis. But typically, a candidate fallible condition (for instance that one's belief has a subjective probability of .9 or more given what one knows) will provide a vast number of cases in which one has a false belief whose basis satisfy the fallibility condition, and it would be surprising if none of these is such that it can be made so that the belief is true in a manner unconnected to S's basis. At any rate no such analysis has been provided so far.[65]

The bottom line is this: if one argues that in some cases S knows *p* even though it was possible that S believed that *p* on that basis while *p* was false, then one's analysis of knowledge will classify some Gettier cases as knowledge.[66]

The second argument for modal infallibilism is that it accounts for lottery cases. As Kyburg first noted, we have the intuition that one cannot know that one will loose a lottery, no matter high the odds are and even though one will in fact loose.[67] If there is an objective chance that one's ticket will be drawn, then intuitively one cannot know that one will loose, however slight the chance is. As Hawthorne notes, lottery cases are instances of a more general principle:[68]

(18) If there is an objective chance that *p*, then one does not know that not-*p*.

Conversely:

(19) If one knows that not-*p*, there is no objective chance that *p*.

---

65 In order to block the above recipe for Gettier cases, a candidate fallible basis would have to have the following feature: whenever the situation is such that the fallible basis would normally yield false beliefs, it *necessarily* yields false beliefs. Such a fallible basis would thus have the surprising feature of counter-infallibility: when it is in default, it *infallibly* yields false beliefs. It would be most surprising to find a plausible fallible condition for knowledge with that feature.

66 Howard-Synder, Howard-Snyder and Feit (2003) argue that something slightly less that infallibility is required to solve the Gettier problem : in substance, that one's belief is such that *if true*, it is infallibly true. While their condition allows them to classify as warranted some false beliefs, it makes no difference in cases of knowledge, since in the latter one's belief is true. So their view entails modal infallibilism.

67 Kyburg (1961). More precisely, Kyburg noted that if a subjective probability of less than 1 was sufficient for rational acceptance, then one would be lead to rationally accept a contradiction, namely that nobody will win the lottery (since for each player the probability that he looses is close to 1), and that somebody will (since there is a winning ticket). See Hawthorne (2004 : 1) for further references on the lottery paradox.

68 Hawthorne (2004 : 93). More precisely, lotteries where the draw *has not taken place yet* are instances of that principle – assuming the drawing process is genuinely indeterministic. I will return to that point shortly.

Principle (19) is very intuitive.[69] Modal infallibilism provides a straightforward explanation of it: according to infallibilism, if one knows p, then one has a basis for one's belief that p such that one's having that basis makes it necessary that p. Hence the objective chance of not-p being the case is zero.

Given that arguably, only future events (if any) are objectively chancy, the argument provides support only for modal infallibilism concerning knowledge of future events. However, there is a *prima facie* reason to treat knowledge of the past in a parallel fashion. Suppose that the lottery *has* been drawn, but that one has not been told its outcome. Then the chance that one wins is 0 or 1, depending on the outcome. However, the objective conditional probability that one won given that one had ticket k is $1/n$, where n is the total number of tickets. Intuitively, we want to treat that case just as if the draw had not taken place: that is, we would want to say that one does not know because one's belief that one lost could be wrong. That is precisely what modal infallibilism gives us: if the objective conditional probability that one won given one's basis for belief is above 0, then it was possible that one had that basis and one lost, and thus, by modal infallibilism, one does not know. That suggests the following generalisation of (18):

The Chance-Knowledge Principle. One knows that p on basis b only if the objective conditional probability that one's belief is true conditional on one's having basis b is 1.

Assuming that the conditional probability of p given q is zero iff it was not possible that p and q, the Chance-Knowledge Principle is equivalent to modal infallibilism. Thus one who is willing to explain lottery ignorance by the Chance-Knowledge principle is thereby committed to modal infallibilism.[70]

A further reason to accept modal infallibilism is that it provides a justification of epistemic closure. Suppose that one knows that p and that p implies q. By modal infallibilism, one's bases for these two beliefs are such that it is not possible that one has them without q being true. Consequently, if they also are the basis for one's belief that q, then one's belief that q satisfies the modal infallibility condition too. If, as seems plausible, one's having an modally infallible basis for q is sufficient for one's being in position to know q, that gives us an explanation of why one is then in position to know that q.

4.4 Does modal infallibilism lead to scepticism?

The main reason to reject the modal infallibility condition is that it seems to lead to scepticism. In fact, some sceptical arguments can be built on the condition. But as I will argue, whether they go through depends on one's account of possibility.

---

69 Hawthorne (2004 : 91-93) shows that (19) can be derived from two intuitive principles : Lewis' principle according to which if one knows that the objective chance of p is k, then one should assign an epistemic probability of k to p, and what has been called here epistemic infallibilism.

70 Hawthorne and Lasonen (forthcoming) provide further arguments in favour of the Objective Chance-Knowledge principle by showing that attempts to avoid the principle have problematic consequences.

As we have seen (section 2.4), closure-based sceptical arguments lead to a standoff between Mooreans and sceptics, unless sceptics provide an independent reason to think that ones does not know that (for instance) one is not a brain in vat. Now modal infallibilism allows one to build such an argument. Let *b* be one's basis for one's belief that one is not a brain in a vat:

(20) It was possible that one had been a handless brain in vat.

(21) If one had been a handless brain in vat, one could have had *b* and yet one's belief that one is not one would have been false.

(22) By (20) and (21), one could have had a false belief on basis *b*.

(23) By modal infallibilism, one does not know that one is not a handless brain in a vat.[71]

The argument is valid. One can avoid the conclusion only by rejecting at least modal infallibilism or one of (20)-(21). But whether (20) and (21) hold crucially depends on two things: how possibility is understood and how bases for beliefs are individuated.

On some understandings of possibility and some individuations of bases for beliefs, (20) and (21) come out true. For instance, if one's basis for one belief is *internal* in the sense that one could have it however the rest of the world is, then (21) will come out true for any proposition about the external world. And if possibility is logical possibility, (20) comes out true. (It will plausibly do so with metaphysical possibility as well.) That seems to be the motivation behind many sceptical arguments and intuitions about our knowledge of the external world, notably Cartesian scepticism.

Externalists about mental content and disjunctivists will typically object to premisse (21). Externalists about mental content argue that beliefs are dependent on one's causal relations with the world, such that one could not have one's actual belief if one was, for instance, a brain in a vat.[72] Disjunctivists argue that perceptual experiences are similarly world-involving, so that for instance one could not have had the experience one has in seeing a tree had that tree not existed.[73] I cannot do justice to their arguments here. Let me just point out that using externally-individuated mental states as bases for belief in the infallibility condition leads to wrong predictions in simple cases. Suppose two coins A and B are perceptually indistinguishable. Coin A has been shown to Harry, then it was put in a box with coin B, and then a coin was randomly drawn out of the box and shown to Harry. Suppose coin A was in fact drawn. Then Harry's having an A-involving perceptual experience makes it necessary that the coin Harry now sees is coin A. So the infallibility condition would be satisfied. Yet intuitively, Harry could be wrong about the coin being coin A, and he cannot know that it is coin A. However, that does not mean that modal infallibilists

71 Brueckner (2005) offers essentially the same reconstruction of infallibilism-based sceptical arguments. He argues that sceptics rely on a "entailment principle" which is closely related to what is here called "modal infallibilism". However, he claims that the "entailment principle" is not warranted, while we have seen that there are powerful reasons to accept modal infallibilism. But, as I argue in the present section, it remains to be seen whether modal infallibilism warrants the strong infallibility condition that is equivalent to the entailement principle (in which case scepticism follows) or some weaker infallibility condition (it which case it does not).

72 See Putnam (1981).

73 See Snowdon (1980).

should reject disjunctivism. They can say that Harry's basis is the acquiring an (object-involving) experience of the coin, whichever it is. But then that revised basis is fallible with respect to coin A, and the analogue of (21) comes out true. So it is doubtful that externalism or disjunctivism warrant the rejection of (21).

Another type of objection to (21) comes from those who individuate bases for beliefs externally without thereby individuating mental states externally. For instance, one's basis for one's belief that it is 2pm could be described as *one's reading 2pm on a well-functioning watch*. Then even though one could have been misled by a broken watch, that possibility does not impugn one's knowledge because one's basis for one's belief would then be different. Thus the analogue of (21) would be false. In the brain-in-vat case, it could similarly be said that one's basis for belief is one's having the perceptual faculties of a normal human being. However, these kind of solutions typically run into trouble with so-called fake barn cases. Suppose Ian is driving through an area where unbeknownst to him barn facades have been erected for a movie set. Ian points to one of the buildings and says to his son that it is a barn – and in fact it is one. Now if we are allowed to describe Ian's basis as *perceiving a normal barn*, we will wrongly classify his belief as infallible. It is unclear whether one can deal with such cases without giving up the external individuation of bases.

Let me turn to (20). Whatever one says about (21), it is crucial to the sceptical argument that one accepts (20), namely that it could in fact have been the case that some sceptical scenario (such as one's being a brain in vat) obtained. But it is unclear whether it is so. It might be in fact physically impossible that there be brains in vats. For instance, there might no possible history of our universe that leads to a scientist stimulating a brain in a vat in a way that is remotely similar to the one in which a real brain is. If that is so, will one insist that it is possible, because it is metaphysically or logically possible? Moreover, there are even more restrained notions of alethic possibility. For instance, it is intuitively true that one cannot move one thousand kilometres away from where one is in ten minutes. This is true because given the present state of human technology, that is impossible, even though that is physically possible. Given that, it is impossible that one's belief that one will not be one thousand kilometres away in ten minutes cannot be wrong. Furthermore, it is not even clear that everything that is physically possible given the present conditions should be treated as possible. Contemporary physics ascribes extremely small but non-null chances to such events as a falling apple stopping in mid-air. We might want to say that such virtually impossible events have no "real" chance of taking place. And if that notion is the relevant one, massive physical indeterminism combined with modal infallibilism will not be as threatening for our knowledge of the future as it seems at first sight. Inversely, the relevant notion of possibility might also be more liberal that physical possibility: for instance, even if the world turned out to be deterministic, we might still want to say that a flipped coin that landed on heads could easily have landed on tails, because a slight difference in the initial conditions would have made it so.[74]

To sum up: if sceptics are right to claim that we could have been brain in vats, and if that is the notion of possibility relevant to the modal infallibility condition, then we should concede that scepticism is true. But my point here is that that depends on

74 See Williamson (2000 : 123). More generally, one could rely on the conception of chance used in statistical mechanics, which is compatible with determinism (see Albert, 2000).

substantial questions about the metaphysics of possibility and the semantics of modals that require more careful examination. Several situations might arise. A) sceptics are right to claim that we could have been in a sceptical scenario, and scepticism is true; B) sceptical scenarios are in fact impossible, and we know that they do not obtain; C) it is impossible for us to know whether sceptical scenarios are possible, and thus if B) is true then we know but cannot know that we do; D) statements of possibility might be context-sensitive, so that it is true in certain contexts to say that sceptical scenarios are possible, but false in others – and consequently ascriptions of knowledge will be context-sensitive as well.

The same questions arise for more mundane sceptical hypotheses. Take the painted mule case. Chris is looking at a zebra in a pen. Could it have been the case that that zoo had a painted mule here that Chris would not distinguish from a zebra? It is not obvious to me that the answer is "yes". The zoo authorities might be such that they could never had accepted to do that. It might have been impossible to get a mule, or to put a mule there without it being quickly noticed, and so on. Or it might be that the affirmative answer is true in some conversational contexts, but false in others. Before those issues are settled, it is not clear that modal infallibilists will have to deny that Chris knows that there is a zebra in the pen, or that that ascription in false in ordinary contexts.

The moral of the infallibility-based sceptical argument is thus the following: while Descartes is right to assume that knowledge requires that one's belief could not have been wrong, the sceptical consequences he initially draws might go wrong in two ways: first, he might be wrong to assume that the bases our beliefs are internally individuated phenomenal states, and second, he might be wrong to assume that every metaphysical possibility is a genuine possibility.

5 Conclusion

I have distinguished three notions of infallibilism that can be found in the literature. The first, epistemic infallibilism, is just a consequence of the highly intuitive principle of epistemic closure, and it should not properly be called "infallibilism", since it is compatible with such fallibilist accounts of knowledge as the one according to which knowledge is just true belief. The second, evidential infallibilism, is the requirement that one's inferential knowledge be deductively based on known evidence. I agree with most contemporary epistemologists that it should be rejected. However, it is an unsatisfactory account of the infallibilist intuition. The third, modal infallibilism, is the idea that knows only if one could not have been wrong. Modal infallibilism captures the core infallibilist intuition, allowing us to derive Descartes' infallibilism and evidential infallibilism under some (controversial) assumptions. I have argued that modal infallibilism should be accepted in order to solve the Gettier problem and to account for our ignorance in lottery cases. Unfortunately, modal infallibilism is also a source of sceptical arguments. However, I have argued that whether these arguments are sound depends on substantial questions about the metaphysics of possibility and semantics of alethic modals. But whatever the upshot of these latter issues turn out to be, modal infallibilism is a constraint with which we will have to live.

References

Albert, D. (2000) *Time and Chance*, Harvard, Harvard University Press.

Alston, W. (1992), "Infallibility" in J. Dancy and E. Sosa, *A Companion to Epistemology*, Oxford, Blackwell.

Audi, R. (1998), *Epistemology*, London, Routledge.

Ayer, A.J. (1956), *The Problem of Knowledge*, London, Penguin.

BonJour, L. (1985), *The Structure of Empirical Knowledge*, Cambridge, MA, Harvard University Press.

BonJour, L. (1998), *In Defense of Pure Reason*, Cambridge, Cambridge University Press.

Brueckner, A. (2005), "Fallibilism, Underdetermination, and Skepticism", *Philosophy and Phenomenological Research*, 72:2, pp. 384-391.

Carrier, L.S. (1993), "How to Define a Nonskeptical Fallibilism", *Philosophia* 22: 3-4, pp. 361-372.

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

Cohen, S. (1988), "How to Be a Fallibilist", *Philosophical Perspectives*, 2, pp. 581-605.

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

Davidson, D. (1986) "A Coherence Theory of Truth and Knowledge", *in* E. Lepore (ed.), *Truth and Interpretation*, Oxford, Blackwell.

DeRose, K. (1992), "Descartes, Epistemic Principles, Epistemic Circularity, and *Scientia*", *Pacific Philosophical Quarterly*, 73, 220-238.

Dretske, F. (1970), "Epistemic Operators", *Journal of Philosophy*, 67, pp. 1007-1023.

Dretske, F. (1981), "The Pragmatic Dimension of Knowledge", *Philosophical Studies*, 40, pp. 363-378.

Dougherty, T. and Rysiew, P. (forthcoming), "Fallibilism, Epistemic Possibility and Concessive Knowledge Attributions", *Philosophy and Phenomenological Research*.

Feldman, R. (1981), "Fallibilism and Knowing That One Knows", *Philosophical Review*, 90, pp. 266-282.

Fogelin, R. (1994), *Pyrrhonian Reflections on Knowledge and Justification*, Oxford: Oxford University Press.

Gettier (1963), "Is Justified True Belief Knowledge?", *Analysis* 23, pp. 121-123.

Goldman, A. (1976), "Discrimination and Perceptual Knowledge", *Journal of Philosophy*, 73, 771-791.

Goldman, A. (1979), "What is Justified Belief?", *in* G.S. Pappas, *Justification and Knowledge*, Dodrecht, Reidel.

Hawthorne, J. (2004), *Knowledge and lotteries*, Oxford, Oxford University Press.

Hawthorne, J. and Lasonen, M. (forthcoming) "Knowledge and Chance", in D. Pritchard and P. Greenough (eds.), *Williamson on Knowledge*, Oxford, Oxford University Press.

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

Howard-Synder, D., Howard-Synder F., and Feit, N. (2003), "Infallibilism and Gettier's legacy", *Philosophy and Phenomenological Research*, 63, pp. 304-327.

Hyman, J. (1999), "How Knowledge Works", *Philosophical Quarterly*, 49: 433-451.

Jeshion, R. (2000), "On the Obvious", *Philosophy and Phenomenological Research*, 60:2, pp. 333-355.

Klein, P. (1998), "Foundationalism and the Infinite Regress", *Philosophy and Phenomenological Research*, 58, 919-926.

Kyburg, H. (1961), *Probability and the Logic of Rational Belief*, Middletown, Conn., Wesleyan University Press.

Lehrer, K. (1974), *Knowledge*, Oxford, Clarendon Press.

Lehrer, K. (1990), *Theory of Knowledge*, Boulder, Westview Press.

Lewis, D.K. (1996), "Elusive Knowledge", *Australasian Journal of Philosophy*, 74, pp. 549-567.

Millar, A. (1991), *Reasons and Experience*. Oxford: Clarendon Press.

Moore. G.E. (1959), "Certainty", in his *Philosophical papers*, London: Allen & Unwin.

Moser, P. (1989), *Knowledge and Evidence*. Cambridge, Cambridge University Press.

Nozick, R. (1981), *Philosophical Explanations*, Oxford, Oxford University Press.

Prichard, H. (1950), *Knowledge and Perception*, Oxford, Clarendon Press.

Pritchard, D. (2005), *Epistemic Luck*, Oxford, Oxford University Press.

Putnam, H. (1981), *Reason, Truth and History*, Cambridge, Cambridge University Press.

Reed, B. (2002), "How to think about fallibilism", *Philosophical Studies*, 107, pp. 143-157.

Sartwell, C. (1992). 'Why Knowledge is Merely True Belief', Journal of Philosophy, 89, pp. 167-80.

Schaffer, J. (2005). "Contrastive Knowledge" in T. Gendler and J. Hawthorne, (eds), *Oxford Studies in Epistemology*, I, Oxford, Oxford University Press.

Sextus Empiricus (1933-49), *Sextus Empiricus with an English Translation*, 4 vols., tr. R.G. Bury, London: Heinemann.

Siegel, H. (1997), *Rationality Redeemed?* London, Routledge.

Snowdon, P. (1980), "Perception, Vision and Causation", *Proceedings of the Aristotelian Society* 81.

Stanley, J. (2005). "Fallibilism and Concessive Knowledge Attributions", *Analysis*, 65:2, pp. 126-131.

Unger, P. (1971) "A Defence of Skepticism", *Philosophical Review*, 80, pp. 198-219.

Unger, P. (1975), *Ignorance: A Case for Skepticism*, Oxford, Clarendon Press.

Vogel, J. (1990), "Cartesian Skepticism and Inference to the Best Explanation", *Journal of Philosophy*, 87: 658-666.

Williams, M. (1991) *Unnatural Doubts: Epistemological Realism and the Basis of Scepticism*, Oxford, Blackwell.

Williams, M. (2001) *Problems of Knowledge: A Critical Introduction to Epistemology*, Oxford, Oxford University Press.

Williamson, T. (2000), *Knowledge and its limits*, Oxford, Oxford University Press.

Zagzebksi, L. (1994), "The Inescapability of Gettier Problems", *Philosophical Quarterly* 44, pp. 65-73.

Zagzebski, L. (1999), "What is Knowledge?", in J. Greco and E. Sosa, *Epistemology*, Oxford, Blackwell.