The Limits of Antiskeptical Infallibilism

Gregory Stoutenburg

gstoutenburg@ycp.edu

1 Fallibilism, Infallibilism, and a Reductio Argument

Infallibilism about knowledge—the view that a necessary condition of knowing that a proposition is true is that one’s basis for believing the proposition is sufficient for the proposition’s truth—has enjoyed a slight renaissance over the past two decades due (probably) to the influence of Williamson’s Knowledge and Its Limits (2000), which espoused an antiskeptical infallibilism.1 At least since Gettier, epistemologists have commonly accepted a version of fallibilism.2 The skeptical threat of infallibilism is easy to spot: if knowing requires having a basis for belief so secure that it is impossible that one’s belief is false given one’s basis for belief, and if furthermore, as is plausible, we rarely or never have such excellent grounds for belief, then there is little or nothing that we know. The appeal of fallibilism as the view to take on pain of widespread skepticism

---

1 The definition of “infallibilism” is itself a matter of dispute. It has been defined as the thesis that when one knows that p, one’s basis for belief ‘entails’ that p; that when one knows that p, Pr(P) = 1.0; and that when one knows that p, it is impossible that one’s belief is false relative to one’s basis for belief (Dodd 2011, Dutant 2016, Reed 2002). I prefer my gloss in terms of a basis for belief that is sufficient for truth because it avoids concerns about knowledge of necessary truths and does not presuppose that all evidence is propositional. The language of a basis for belief ‘guaranteeing’ the truth of the belief has similar advantages.

2 Dutant (2015) argues (to me, convincingly) that a kind of internalist infallibilist view about knowledge is historically the dominant view, and that only after Gettier (1963) that fallibilism became popular.
is obvious. But if a plausible non-skeptical version of infallibilism is tenable, then this advantage of fallibilism is illusory.

And that is a good thing, because fallibilism struggles mightily to accommodate several strong intuitions about knowledge, among them: that concessive knowledge attributions are self-contradictory (Dodd 2010, Dougherty and Rysiew 2009, 2011, Lewis 1996, Stanley 2005), that knowledge is closed under known entailment (Carter 2011, Feldman 1995, Sharon and Spectre 2012, Stine 1976), and that knowledge is incompatible with luck (Pritchard 2005, 2010, Stoutenburg 2015, 2018). But the argument that I will use to motivate infallibilism in this paper focuses on two ideas that I think are ultimately contradictory. The first is that when one knows that p it is impossible for one that not-p. The other is that one often knows that p when the connection between knowing that p and one’s support for p is insufficient for the truth of p. It is easy to show that those two claims are contradictory, which is bad news for fallibilism.

Here are those claims again:

1: One can come to know that p without one’s basis for belief guaranteeing that p.
2: If S knows that p, then not-p is epistemically impossible for S.

(1) and (2) jointly imply a contradiction. It might be easy for you to see the contradiction. (1) implies that one can know when it is possible that one’s belief is false. (2) says that whenever one knows, it is not possible that one’s belief is false. I think the contradiction is obvious, but it can be spelled out in more detail:

Assume for reductio that (1) S can come to know that p without S’s basis for belief guaranteeing that p and (2) that if S knows that p, then not-p is
epistemically impossible for S. Suppose that S’s true belief that p is very highly probable, for S, relative to S’s basis for belief. Perhaps S believes that p because S trusts a very reliable source that claims that p, or S calculated that the odds of p are extraordinarily high. Still, that does not eliminate the possibility that p is false, relative to S’s basis for belief. So, p is possibly false (for S, and relative to S’s basis for belief). Yet most epistemologists would grant that the bases for belief just indicated are regularly used to acquire knowledge. So, let us grant, in accordance with (1), that S knows that p. If so, then according to (2), it is not possible for S that p is false. But we said a moment ago that it is possible for S that p is false. Contradiction.

Fallibilists are clearly committed to (1). So, they must give up (2). But (2) is plausible, which tells against fallibilism.

Perhaps a particular example will help with finding the contradiction. Here is one:

I believe that I had eggs for breakfast, several hours ago. I believe this based on memory, as I am now at work and there are no eggs in sight. I may take myself, and ordinarily would, to know that I had eggs for breakfast: if I were asked, “Do you know what you had for breakfast?” I would reply, “Yes: eggs.” If later, I were offered either a donut or eggs, I might reason, “I can take the donut without guilt, because I already had some protein today when I ate eggs…” and so on. So, since I take myself to know that I had eggs, I take it that the possibility that I actually did not have eggs is no possibility at all.

Of course, I also know that memory is fallible. We misremember and confabulate. My belief that I had eggs for breakfast is based on memory. So, I am forced to conclude that it is possible that I merely seem to remember having eggs when I had something else for breakfast or no breakfast at all. Perhaps I had eggs yesterday and mistakenly regarded that memory as a memory of this morning instead of yesterday morning. In any case, I admit that it is possible that I did not have eggs for breakfast.
When I take myself to know something based on memory, I think of that belief as not possibly false. When I consider that my basis for belief is fallible, I think of that belief as possibly false. One of these must go. That is the point of the *reductio* argument.\(^3\)

2 Antiskeptical Infallibilism: E=K and Safety

Timothy Williamson is sometimes thought of as an infallibilist because of his commitment to the thesis that knowledge and evidence are identical (“E=K”) (2000). If my evidence includes p, then I know p; and if I know p, then my evidence includes p. Trivially, then, all known propositions have a probability of 1.0 on one’s evidence, because the probability of anything given itself is of course 1.0. This kind of infallibilist view is subject to the *reductio* argument that I just gave. Suppose that I form the belief that I ate eggs for breakfast based on my apparent memory, and further suppose that I really did have eggs and that there are no other oddities in the background (e.g. no Gettier stuff is going on). Then most would grant that I know that I had eggs for breakfast.

According to E=K, *that I had eggs for breakfast* is a part of my evidence. So, the probability that I had eggs for breakfast is 1.0. But I also recognize that it is possible that my memory has failed me, as it sometimes does. And if it is possible for me that p, then the probability for me that p must be above 0.\(^4\) So, the probability that I did *not* have eggs

---

\(^3\) I am not the first to notice that fallibilism is in tension with the idea that knowing implies that one’s belief is not possibly false (DeRose 1991, Reed 2013, Weatherson and Egan 2011). However, I offer this as an argument against fallibilism and as a constraint on plausible versions of infallibilism. An implicit assumption that I will not defend in this paper is that epistemic invariantism is true.

\(^4\) This is an attempt at glossing possibility-language as probability-language, but nothing very substantial rides on it. Dodd (2011) defines infallibilism probabilistically. The issue is tricky, though, since one might combine a safety theory of knowledge with an infallibilist view, and some safety theorists deny that probable alternatives are close alternatives. See Rabinowitz (2018).
for breakfast is greater than 0. So, the probability that \( p \) is 1.0 and the probability that \( \neg p \) is greater than 0. Contradiction.\(^5\)

I think the mistake here is to claim that one can assign probability 1.0 to propositions that one comes to believe inductively (cf. Littlejohn 2011). There are two possible implications of this. One is that fallibilism is true after all, so we should not assign a probability of 1.0 to all that many propositions, and that is fine because we still get to call those propositions knowledge. The alternative is that we do not know propositions like “I had eggs for breakfast”. Neither option is satisfying for one interested in non-skeptical infallibilism.

Is there a better option? Intuitively one has infallible knowledge of having had eggs for breakfast only if it is impossible that one did not eat eggs, relative to one’s basis for believing that breakfast included eggs. Whatever exactly eliminating a possibility amounts to, the no-eggs possibility is no possibility at all if one has infallible knowledge of having had eggs for breakfast. It is not plausible that inferentially-supported beliefs could meet that standard.

A more plausible candidate for an antiskeptical version of infallibilism is Julien Dutant’s (2016) safety-based view. His safety theory has a few components. The first is the safety condition on knowledge itself: “S knows that \( p \) only if S’s belief that \( p \) is sufficiently epistemically unlike any false belief” (150). A belief counts as knowledge

\(^5\) Similar arguments against specifically Williamson’s views about evidence and knowledge feature in Dodd (2007), Littlejohn (2011), and Dunn (2014).
only if (note that the claim as phrased is only a necessary condition) that belief is, epistemically speaking, not like any false belief. Safety, phrased in this way, implies infallibilism so long as being sufficiently unlike is absolute: one’s true belief that p is knowledge only if it is completely epistemically unlike any false belief.

The second important feature is the concept of ‘epistemic likeness’ that features in the definition of generic safety. Crucially, the concept of epistemic likeness is eventually replaced with the concept of epistemic peerhood, as the relation that interests Dutant must be asymmetric for reasons discussed below. More important than whether the relation is symmetrical or asymmetrical, though, is the character of the relation itself. Dutant recognizes that an account of epistemic peerhood given in terms of internal indistinguishability would allow that ordinary subjects and subjects deceived by a Cartesian demon have beliefs that are epistemically alike, and that, combined with safety, would imply that the ordinary believer lacks ordinary knowledge. So, some account of the epistemic peerhood relation is needed other than internal indistinguishability. In attempting motivate his account of the relation, Dutant offers some intuitions about knowledge. The first is that if one holds a single ticket in a fair lottery with 1,040,000 tickets and a single guaranteed winner, one does not know that one’s ticket will lose. Even if one formed the true belief that a particular ticket (de re) is the winner, one would believe similarly of any other ticket. Beliefs about individual tickets being the winning

---

6 As far as semantic concerns go, “peerhood” sounds like it expresses a symmetrical relation, just as “likeness” does.
7 In that case both subjects would lack knowledge, but only the deceived believer would lack knowledge on account of false beliefs.
ticket are epistemic peers in this way. One’s true belief that one’s ticket will lose is not knowledge, for even if the ticket does lose, the belief my ticket will lose has as its peers many false beliefs, namely every belief that one would have had about the other, losing tickets, had one held one of those tickets instead of this one. So, in keeping with intuition, one cannot know that one’s ticket will lose in a lottery with this sort of setup. (Presumably, the result will generalize to a fair lottery with a guaranteed winner and as many tickets as you like, since any true belief about a particular ticket being a winner will have many false peers.)

The other motivating intuition about knowledge that Dutant uses to flesh out an account of epistemic likeness is the intuition that one can know that throwing twenty fair coins will not result in all twenty coming up heads. There is some intuitive pull here, as we often say that we do know such things. Personally, I do not feel the intuition for that verdict very strongly, as the coin flip case no less than the lottery case plainly leaves open the possibility of false belief, but perhaps I am too much of a committed infallibilist. Nevertheless, Dutant points out that those who think knowledge in lottery cases is impossible but who also think we can know ordinary propositions about the world need some account of the distinct verdicts. As he argues, if we cannot know in the coin toss case, then we cannot know much about the world, because the occurrence of many common physical processes is about as improbable as twenty coins coming up heads in a row, but we think we know in such ordinary cases. Dutant uses the examples of the 1,040,000 ticket lottery and twenty fair coins all landing heads because the probabilities are almost the same. So, the question is: What makes lottery ticket beliefs peers and coin
toss beliefs non-peers, such that one lacks knowledge in the former case, but has it in the latter case?

Dutant’s answer is that while all beliefs about particular lottery tickets being losers are epistemic peers of each other, not all beliefs about coin tosses are each other’s peers. That is why mere similarity between true beliefs and false beliefs is not the belief-destroying property respected by the safety condition: true lottery beliefs are similar to false lottery beliefs, just as true coin toss beliefs are similar to false coin toss beliefs. But since Dutant thinks that we can have advance knowledge about coin flips but not about lottery outcomes, there must be some asymmetric relation that obtains in coin cases but not in lottery cases. Remember, a true belief is knowledge only if it is peerless.

The suggestion is that beliefs about the outcomes of coin flips are similar only to other beliefs about outcomes of coin flips that are formed in circumstances with similar actual frequencies of outcomes of the same kind. (Suppose that) in the history of coin flips, heads come up about half of the time.\(^8\) So, peers of my belief that not all will come up heads are only those beliefs that are formed in circumstances that are at least as normal as the circumstances that I am now in. Thus, I can know that not all twenty coins will come up heads because that possibility is highly dissimilar to the actual scenario, dissimilar enough that the beliefs that not all twenty come up heads and that half come up...

---

\(^8\) Externalists often make unjustified assumptions about the truth-values of contingent propositions. In assuming that we know that not all twenty coins will land heads, Dutant is assuming that heads and tails come up with roughly equal frequency. No one knows if that is the case. Similarly, externalists who think that we know truths about the external world assume that our belief-forming mechanisms typically result in true belief: but again, there is no additional reason offered as to why we should believe this.
heads are not peers. Dutant compares the role played by actual outcomes in his account with the role played by the environment in some accounts of knowledge. As he puts it, “the all-heads scenario belief differs from the half-heads scenario one like a fake barns scenario belief differs from an ordinary-barns scenario one” (159). In both pairs of cases, one’s knowledge depends upon the structure of the world that one is in.

Thus whether or not one can know that some run of tosses will result in a particular outcome has nothing to do with how many ways the results could turn out and everything to do with how close the predicted result is to how similar runs actually turn out. So, if the actual heads result of flipping twenty coins is 0.5, and the belief one had formed was that half of the coins would come up heads, then that belief’s peers should include the belief that eleven coins will come up heads and that nine come up heads, since 11/20 coming up heads is very similar to 10/20 coming up heads, and as 9/20 coming up heads is also similar to 10/20 coming up heads.

But that account of peerhood implies that we can have knowledge of all sorts of things that we intuitively cannot know. Assuming Dutant's account of belief-peerhood, it should be possible to know, prior to flipping twenty coins, that about half of the flips will come up heads. I stipulate “about half” to cover the range of outcomes that count as peers

---

9 In spelling this example out, Dutant says: “Let p be the proposition that not all coins will fall heads. Suppose in the actual outcome the frequency of heads is exactly heads. Then the belief in p I actually form is very similar to the beliefs in p I would form if the outcome had a frequency of heads of 1 above or below half. It is slightly less similar to the beliefs I would form if the outcome had a frequency of 2 above or below half…[T]he outcome plays a role analogous to that of the ‘circumstances’ or ‘environment’ in the Gettier literature…” (158-9).
of the belief that \textit{exactly half} are heads.\textsuperscript{10} Then there will be no false peers of the ‘about half’ belief because all beliefs that are outside of the ‘about half’ range will be normal only in circumstances that are sufficiently distant from the actual circumstances, wherein half of coin tosses come up heads. So, given Dutant’s account, one can know in advance of flipping twenty coins that about half of them will come up heads. But obviously it is \textit{possible} relative to everything the subject reasonably believes that the result will not be about half-heads, given the equal probability of any particular outcome. So, Dutant’s account does not provide a satisfactory way out of the \textit{reductio}.

One interested in defending this kind of infallibilist view should note further implausible consequences. It is not only beliefs about the results of future coin tosses that can be known. Once the principles used to generate that knowledge are understood, it turns out that we can have \textit{infallible} knowledge of all kinds of facts that intuitively we just cannot know. Using this account, we can know that there are zebras in the pen rather than cleverly-disguised mules, because in the actual world zoos put zebras, and not cleverly-disguised mules, in pens, so there are no false peers of the belief that \textit{there are zebras and not cleverly-disguised mules in the pen}. We can come to know that local circumstances of perception are normal, because there are no false peers of the belief that \textit{local circumstances of perception are normal}. We can come to know that we are not brains in vats because the belief \textit{that I am not a brain in a vat} has no false peers. A

\textsuperscript{10} In correspondence, Dutant notes (correctly, I think) that “about half” is vague: it could mean within 1 heads result of 0.5 distribution, or it could mean 2, or it could mean 3. The boundaries of “about half” are just not sharp. But I do not think that undermines my argument. Intuitively, we can have beliefs with contents that say ‘approximately such-and-such’ without the content of that belief being equivalent to any more precise number.
crucial motivation behind Dutant’s account of epistemic peerhood is resisting the skepticism that is normally thought to accompany an internalist version of infallibilism, but as we see here, this externalist infallibilist view makes “infallible” knowledge far too easy to have.

There are a few replies available to Dutant. One is to deny that beliefs that involve circumstances that are less normal than the actual circumstances are really peers. Let us stipulate for the sake of argument that by “about half” we mean “within 1 of half”. So, about half ranges from 9/20-11/20 heads. But, if epistemic peerhood only extends to beliefs that are at least as normal as the target belief, then my belief that within 1 of half of the flips come up heads will not have as a peer the belief that 8/20 or 12/20 come up heads. That proposal is ad hoc, as it is hard to see what would motivate putting the margin between knowledge and non-knowledge in exactly that spot aside from a desire to protect a favored theory.

Another reply is to claim that a belief’s peers include those that are not significantly less normal than the belief. This would allow us to avoid the consequence that the belief that 8/20 come up heads is not a peer of the belief that 9/20 come up heads. But with that revision, the account is now a kind of fallibilism: one can know that p when one’s belief that p does have some false peers: for example, I could know that about half of coins will come up heads even though I would believe similarly if 8/20 or 12/20 came up heads. To really call a view infallibilist requires at least, it seems to me, that one’s basis for belief is sufficient for the truth of the belief. With this reply, however, one gives up sufficiency, for it would be possible to be highly justified in believing that my coin
will land ‘about’ heads (between 9/20-11/20 heads) when it just might land 15/20 heads, even though the belief that the coin would land 15/20 heads is supposedly not to be considered a peer. Combining that with closure allows one to argue thus:

1. I know that my coin will land about half heads.
2. If (1), then I can know that my belief will not land 15/20 heads.
3. Therefore, I can know that my coin will not land 15/20 heads.

I went from a claim about my advance knowledge of coin flips to a conclusion that, intuitively, is false, for I lack any basis for believing that my coin will not delivery precisely that result.

Similarly, bracketing some technicalities just raised, I should be able to argue this way:

1. I know that my coin will land about half heads.
2. If (1), then I can know that coins in my environment usually land heads, or near enough to half-heads.
3. Therefore, I can know that coins in my environment usually land heads, or near enough to half-heads.

Is the threat of skepticism really so great that we should allow that one can infer facts about one’s environment based on one’s allegedly infallible knowledge about future outcomes? But the point remains: it is very hard to accept the idea that one can have infallible knowledge that \( p \) when a necessary condition of knowing that \( p \) is that certain truths hold about the circumstances of belief-formation, when one has no \textit{prior} knowledge that those truths do indeed hold: knowledge that, intuitively, is constitutive of one’s knowing that \( p \) (cf. Wright 2007).
In concluding this section I will reply to an objection that sometimes comes up when discussing this argument. The objection is that perhaps a relevant alternatives account of knowledge could navigate through the dilemma I posed in the *reductio* argument, by allowing that one can have ordinary knowledge by discarding ‘irrelevant’ possibilities of error *and* by claiming that when one knows that *p* in such a way then one’s belief that *p* is not possibly false. But a not-*p* possibility, even one that is allegedly irrelevant, remains a possibility, and so long as it does it is false to say that my belief that *p* is not possibly false when the ‘irrelevant’ not-*p* possibility remains uneliminated.\(^\text{11}\) To see this requires no more than repeating the argument I gave above. Suppose that I know that I have hands and that one irrelevant possibility is that I am a handless brain in a vat stimulated in such a way as to make it appear to me as though I have hands. Then I can reason thus:

1. I know that I have hands.
2. If (1), then I can know that I am not a handless brain in a vat.
3. Therefore, I can know that I am not a handless brain in a vat.

We can add an alternative conclusion by supplementing this argument with the further claim that all infallibilists should make about knowledge, namely that when one knows that *p* it is impossible for the subject that not-*p*. By doing so, we can arrive at this further conclusion:

4. It is impossible that I am a handless brain in a vat.

\(^{11}\) It is for this reason that a relevant alternatives view is a kind of fallibilism. Nevertheless, this objection surfaces from time to time.
That is clearly just too easy a refutation of skepticism.

3 Infallibilism Antiskeptical and Skeptical

What I hope to have shown thus far is that there are significant obstacles in the way of an antiskeptical version of infallibilism. Such a view is open to technical objections and unintuitive consequences concerning the range of propositions that can be known if the theory is true. To my mind, there is no better version of non-skeptical infallibilism out there. But what is the alternative? Well, there are no alternatives to be excited about, as any version of infallibilism that will avoid the objections I have given will likely turn out to imply a widespread form of skepticism.

We can define an internalist infallibilist view that avoids making infallible knowledge too easy to have and that avoids the technical problems given above, in this way:

Internalist Infallibilism: one knows that \( p \) if and only if one believes that \( p \), and one does so on a basis that guarantees that \( p \) is true, and one is aware of this connection between the truth of \( p \) and the basis for \( p \).

For a subject to be aware of the connection between the basis for belief that \( p \) and the truth of \( p \) will trivially entail that there are no uneliminated possibilities of error with respect to \( p \), and that there will be no instances wherein a subject knows that \( p \) and yet the epistemic status of the subject’s belief that \( p \) constitutively depends upon the subject knowing some other proposition that the subject could not possibly know. In short, the internalist infallibilist view just mentioned is a kind of Cartesianism about knowledge.

Where does that leave us? The initial motivation for an infallibilist view of knowledge has first and foremost to do with the implausibility of a fallibilist conception
of knowledge. We do not need to take seriously the idea that one can really know that a proposition is true when it is possible that the proposition is false, or that one can really know that a proposition is true by inferring it from even very excellent statistical grounds. That just does not count as knowledge the way we ordinarily think of it. Within the infallibilist camp, then, one must make choices: should we prefer a version of infallibilism that leads to skepticism, or a version of infallibilism that suffers from technical objections and makes ‘infallible’ knowledge very easy to have? While I think the choice is not terribly easy—especially if one holds out hope that a particularly sharp epistemologist will come along and address some of those technical objections (though I lack such a hope)—the force of the objection to a skeptical version of infallibilism could be largely absorbed by a principled account of how our knowledge-talk serves us well in ordinary discourse even if our knowledge claims are often false.

I will not argue in detail for such an account here. I have suggested in other places that the best account of our knowledge-attributing behavior ultimately supports the view that ordinary subjects accept very demanding standards for knowledge (Stoutenburg 2016, 2017a, b). Here I will very briefly sketch some key ideas that support an infallibilist, skeptical account of knowledge talk.

We must first observe that some of the terms that we use as predicates do not strictly apply in ordinary cases. Take “square”. We call an object—say, a building, or a parking lot, or a desk—“square” when it is in at least some respects close to being a square. But being close to being a square is not the same as being a square. Real squares have several properties: their sides are all (perfectly) straight and intersect at ninety-
degree angles (ninety: no more, no less, not even a little), and a line drawn across opposite angles traces the hypotenuse of two right triangles. The “square” in the parking lot has *none* of those properties. So, it is not a square. Nonetheless, it may be useful for ordinary purposes to call it a square. This is just one example. Similar remarks apply for nearly all quantity-words and geometric-property-words, among others.

The next main step is to provide a pragmatics of the appropriateness of false speech. This part is much more difficult: it is one thing to recognize that much of what we say is strictly-speaking false; it is quite another to show that that is perfectly all right. Indeed, error theorists about epistemic terms typically want to show that it is *not* perfectly all right to go on using epistemic vocabulary in familiar ways (Stoutenburg 2017b, Streumer 2017, Unger 1975). I think the view can be made to work in a Gricean framework. First, grant that speakers and hearers at least roughly accept infallibilist standards for knowledge. (This may sound shocking and implausible, but it should not: how else are we to make sense of our unwillingness to attribute lucky knowledge, to deny that one can know that p when it is possible that not-p, and so on? The very act of responding to thought experiments with intuitive judgments makes no sense if we do not *in some sense* accept some views about the matters at hand. *All* philosophers defending a view—at least, from the armchair—should claim that ordinary subjects accept the view being proposed.) Second, when a person claims that someone knows something, hearers should take that claim to violate the maxim of Quality, which says to assert only what is true: since per hypothesis knowledge claims are thought of as false, hearers will calculate to an implicature that satisfies all of the other maxims, too, including Relation. So, when
a person makes a knowledge-claim, hearers will look for a nearby claim that strikes them as true and that serves the relevant purposes at hand.

That is only a sketch of a view, but I think that when sufficiently filled out, it will show that a skeptical, Cartesian version of infallibilism might just be the best account of knowledge we can have.\(^\text{12}\)

References


\(^\text{12}\) The argument of this paper was greatly improved by feedback from the many UFSM philosophy faculty members and graduate students who attended my presentation at the conference. Thomas Grundmann and Matthew McGrath deserve special thanks for their incisive comments during the discussion period. Most of all, though, thanks go Julien Dutant for his very helpful correspondence about these issues.