Evidence and Fallibility

Joshua DiPaolo

11.24.15

Abstract: The "Evidentialist Dictum" is the thought that we must believe what our evidence supports, and the "Fallibility Imperative" is the thought that we must account for our fallibility when forming beliefs. Both of these thoughts are plausible, but there appears to be a tension between them. I illustrate this tension by articulating a puzzle that a proponent of the Evidentialist Dictum must address. I then consider the prospects of a novel Evidentialist position. In addition to having the ability to solve the puzzle, this position appears to have several other virtues. In particular, it appears to explain, in cases where other Evidentialist positions cannot, what's wrong with certain illegitimate dismissals of misleading evidence. However, this explanation opens it up to objections. I argue that this position is ultimately untenable, and that the tension between the Evidentialist Dictum and the Fallibility Imperative is genuine. Finally, I offer a brief sketch of how our epistemological theorizing can capture both principles, in spite of this tension, by distinguishing between two types of prescriptions, prescriptions of perfection and counsels of compensation. The Evidentialist Dictum is a prescription of perfection, whereas the Fallibility Imperative is a counsel of compensation.

A wise man…proportions his belief to the evidence.

–Hume, An Enquiry Concerning Human Understanding¹

In general, there is a degree of doubt, and caution, and modesty, which, in all kinds of scrutiny and decision, ought for ever to accompany a just reasoner.

–Hume, An Enquiry Concerning Human Understanding²

1. Introduction

Consider two plausible thoughts. The first thought – the Evidentialist Dictum – is that we must believe what our evidence supports. The second thought – the Fallibility Imperative – is that we must account for our fallibility when forming our beliefs. Little seems more obvious than the Evidentialist Dictum, and while the Fallibility Imperative may be less obvious, it's no less plausible.

Nevertheless, a tension between plausible elaborations of these two thoughts has been emerging in a growing literature.³ Consider the debate surrounding peer disagreement. Those who give priority to the Fallibility Imperative insist that when confronted with a disagreeing peer, you

¹ David Hume, (1999: 170). Locke expresses the same idea in his Essay Concerning Human Understanding 4.5.5.
should often reduce your confidence on the disputed issue. Those who give priority to the Evidentialist Dictum accuse these theorists of recommending that agents inappropriately ignore evidence. And even among conciliationists – those who think peer disagreement often requires a reduction of confidence – there are those who believe such epistemic modesty has its limits: although awareness of our fallibility should make us reduce our confidence on other disputed matters, no information about our fallibility can make it reasonable to reduce confidence in conciliationism itself. In other words, the evidence will always favor conciliationism, and no amount of disagreement can make it reasonable to ignore that evidence.

Is the tension between the Evidentialist Dictum and the Fallibility Imperative real or merely apparent? The answer is: it’s complicated. That’s where I’ll end up. I will begin, however, by investigating this apparent tension from the perspective of the proponent of the Evidentialist Dictum. In particular, I will articulate a puzzle based on considerations of fallibility that a proponent of the Evidentialist Dictum must address. I will then consider the prospects of a response to the puzzle – an “evidence-loss” response – that has the potential to dispel the tension. This sort of response has not been discussed in the literature, but both friends and foes of the Evidentialist Dictum have something to gain by considering it because the case against the Dictum depends on an evidence-loss account being untenable.

The tension between the Evidentialist Dictum and the Fallibility Imperative – hereafter, simply, the tension – may appear easily dissolvable: that we are fallible is itself just more evidence to be balanced against all the rest. Although this position was a compelling pretheoretical stance, several theorists have recently argued that evidence of fallibility differs importantly from more mundane evidence, and the puzzle described in the next section bolsters this claim. In the end, maybe these theorists are wrong, and maybe the puzzle can be easily solved. But the “just more evidence” view can no longer be taken for granted.

After articulating the puzzle in §2, and explaining and motivating the evidence-loss account in §3, I will argue in §4 that the evidence-loss account provides the wrong sort of response to the

---

5 Kelly (2010), Weatherson (ms.), and Schoenfield (forthcoming a).
7 The puzzle is inspired by Christensen (2010). See also Christensen (2011: 19-20).
8 Lasonen-Aarnio (2014: 317-8) nicely articulates this response, to which she is sympathetic. Feldman (2009) and Kelly (2010) are the main proponents of the “just more evidence” view of evidence of error.
9 See, for instance, Christensen (2010), Lasonen-Aarnio (2014), Schechtter (2013), and Schoenfield (forthcoming b).
puzzle and has an implausible consequence, viz. that it makes evidence possession “purpose-relative,” in a sense I will explain.\(^{10}\) In §5, I briefly reject other responses to the puzzle. Then I conclude in §6 by sketching what I take to be the metaepistemological upshot of the tension between the Dictum and the Imperative. In short, I believe that, like those working in other normative domains, epistemologists can engage with two types of theory, ideal and non-ideal, and that the Evidentialist Dictum may capture an important truth of ideal theory, whereas the truth behind the Fallibility Imperative belongs to non-ideal theory.

2. The Puzzle

We make all sorts of errors: we misremember, we think we see something we didn’t, we make fallacious inferences, we misapply good rules of inference, etc. These are epistemic errors because they often lead to false beliefs. Throughout, my reference to our fallibility signifies our proclivity toward epistemic error.

To begin to see the puzzle our fallibility gives rise to, let’s start with an example:

**Hypoxia:** Andy is piloting an airplane and after doing some calculations, she becomes highly confident that she has enough fuel to fly to Hawaii on the basis of the following evidence, \(E\):

- A full tank contains 20,000 miles worth of fuel.
- The tank is \(\frac{3}{4}\) full.
- Hawaii was 16,000 miles from Andy’s point of departure.
- Andy has flown 5,000 miles toward Hawaii.

Then air traffic control warns her that, as a result of a drop in the oxygen supply in her cabin, any reasoning she’s done in the last five minutes has probably been distorted without her knowing it.\(^{11}\)

In this example, Andy receives evidence that she has likely made an error in her calculations. Call this kind of evidence – evidence that one has made an epistemic error – higher-order evidence. The intuition that Andy should reduce her confidence is widespread,\(^{12}\) and it is typically encoded in the judgment that it would be epistemically irrational for her to maintain her confidence after she receives the higher-order evidence. I will suppose that this judgment is correct, and that it represents the fact that Andy, like the rest of us, must account for her fallibility in her doxastic practice.

To make the puzzle more precise, I will replace talk of the Evidentialist Dictum with talk of Evidentialism. There are several claims one might refer to with the label ‘Evidentialism,’ but on my

---

\(^{10}\) This is not the same as saying that evidence is “interest-relative,” as Stanley (2005: ch. 9) argues. I discuss this view in §4.

\(^{11}\) This case is widely discussed in the literature. It’s originally due to Elga (ms). I’ve added details about Andy’s evidence.

\(^{12}\) For discussion of this or similar cases, see, e.g., Christensen (2010, 2010a 2013), Elga (ms), Schechter (2013), Schoenfield (forthcoming a).
usage, *Evidentialism* is the view that believing a proposition is rational for someone just in case her evidence supports that proposition. On this view, rational belief supervenes on evidential support, rather than evidence alone: any two people for whom the same propositions are supported (to the same degree) by their respective evidence are justified in believing the exact same things. Evidentialism, then, is a broad view that includes both objective and subjective Bayesianisms, for instance.  

One additional claim is needed to generate the puzzle. It is a claim linking evidential support to entailment, so I’ll call it the *Entailment-Support* principle: If a body of evidence entails some contingent proposition, that body of evidence supports that proposition. This is the most controversial claim of the puzzle. Some will view it as downright implausible, while others will think (something like) this principle is the only objective evidential support principle. I’m not sure that it’s the only objective evidential support principle, but I’m confident that it’s not downright implausible. Indeed, I think it’s difficult to reject it, given a plausible conception of Evidentialism. In any case, I will offer reasons for endorsing it in §5. For now, I’ll make two notes: First, the principle does not say that if your body of evidence entails any proposition, your body of evidence supports it. In particular, it does not entail that every necessary truth is justified for everyone. Second, the puzzle can arguably be generated without appeal to this principle.

The gist of the puzzle is that Evidentialism cannot accommodate the judgment that Andy should revise her belief. This is because, by the Entailment-Support principle, Andy’s evidence supports the belief that she has enough fuel, since her evidence entails this proposition.

More carefully: Andy is in an *Entailment Case of higher-order defeat*—an *Entailment Case* for short. In an Entailment Case, at some time $t_1$, the subject’s evidence entails some proposition, $P$, and she believes $P$ on the basis of this evidence. Then, at a later time $t_2$, she acquires higher-order evidence

---

13 This leaves open whether permissivism is true. Notable papers containing arguments against permissivism include White (2005, 2010), Feldman (2007), Christensen (2007). Notable papers that defend permissivism include Titelbaum (2010), Meacham (2014), Schoenfield (2014), and Titelbaum & Kopec (ms.).

14 See Ballantyne & Coffman (2011) for discussion of different evidentialisms. Schoenfield (forthcoming a) treats Evidentialism as I do here. As stated, Evidentialism is something that might be accepted by self-proclaimed evidentialists like Conee and Feldman. However, anyone sympathetic with their (2008) developments of Evidentialism may not find the puzzle described in this section troubling. At the end of §5, I make brief remarks that must be addressed even by someone sympathetic to a Conee-Feldman-style Evidentialism.

15 This principle is best seen as a corollary of Evidentialism, since Evidentialists owe us an account of evidential support, and this principle provides a sufficient condition for that notion.

16 Subjective Bayesians are likely to adopt the latter stance. Titelbaum (2010) explicitly takes this stance.

17 I discuss a version of the puzzle that does not require the Entailment-Support principle at the end of §5.
that renders her belief in P irrational.\footnote{This is a case of higher-order defeat because receiving higher-order evidence that a belief likely resulted from a mismeasurement of the evidence renders that belief unjustified.} By Evidentialism and the Entailment-Support principle, Andy’s belief is justified at $t_1$, and by Evidentialism, Andy’s belief is not supported by Andy’s evidence at $t_2$. If it’s not supported by her evidence at $t_2$, then, by the Entailment-Support principle, it’s not entailed by her evidence at $t_2$. But entailment is monotonic: adding evidence to a body of evidence without removing any evidence leaves the entailment intact. Hence, her belief is entailed by her evidence at $t_2$, hence it is supported, hence it is justified. Thus, the judgment that Andy should revise her belief appears to be inconsistent with Evidentialism and the Entailment-Support principle. Since the latter principle is just a corollary of Evidentialism, this appears to be an instance of the tension: one cannot always simultaneously respect one’s evidence and respect one’s fallibility.\footnote{Notice this consequence isn’t an artifact of the Entailment-Support principle. The judgment that there is defeat in Andy’s case is crucial, and this judgment depends on the type of evidence she receives. If she had received only ordinary (non-higher-order evidence), we wouldn’t judge that defeat occurred.} Without a solution to this puzzle, Evidentialism will be unable to account for our judgments about how evidence of fallibility bears on rational belief.

3. Evidence-Loss

Abandoning Evidentialism at this point would be too hasty. There is a lacuna in the argument based on the puzzle, namely that Andy does not lose evidence between $t_1$ and $t_2$. Indeed, there is an argument from the puzzle to an evidence-loss account. Call it the Evidence-Loss Argument.

1. **Evidentialism**: Believing P is rational for S at t just in case S’s evidence supports P at t.
2. **Entailment-Support principle**: If S’s evidence entails (contingent) P at t, then S’s evidence supports P at t.
3. **Entailment Case**: At $t_1$, S’s evidence entails P, and S believes P on the basis of this evidence. At $t_2$, S acquires higher-order evidence that renders S’s belief irrational.
4. **Entailment-Loss principle**: S’s evidence entails P at $t_1$ and not at $t_2$ only if S loses evidence between $t_1$ and $t_2$.
5. S’s evidence doesn’t support P at $t_2$. (By (1) and (3))
6. S’s evidence doesn’t entail P at $t_2$. (By (2) and (5))
7. Therefore, S loses evidence between $t_1$ and $t_2$. (By (3), (4), and (6))

The Entailment-Loss principle is the only premise not discussed so far, but it follows from the monotonicity of entailment and a plausible understanding of a person’s evidence. Entailment relations between sets of propositions cannot change, but which set of propositions constitutes one’s evidence can change. And, according to this argument, at least sometimes when one acquires higher-order evidence, one’s body of evidence not only gains propositions, it loses propositions as well.
Philosophers have discussed versions of this puzzle before, and some do make reference to the claim that Andy (for instance) does not lose evidence when she receives her higher-order evidence. For instance, Thomas Kelly (2013: 46) writes, “If one’s original evidence entails that p, then it seems like one’s total evidence will always support the belief that p, no matter how much misleading testimonial evidence one subsequently acquires, so long as the original evidence remains part of the total set.” [Bold emphasis mine.] What Kelly says here is, of course, true. But he does not argue for the claim that, in these cases, the subject does not lose evidence. While this claim may appear to require no defense, in the face of the Evidence-Loss Argument more needs to be said to rule out such a response.

Indeed, I think something can be said to rule out this response, but before offering objections, it’s worth seeing what such a response looks like and what can be said in its favor.

The idea suggested by the Evidence-Loss Argument is that as a result of acquiring higher-order evidence that one has improperly assessed one’s original evidence, one no longer possesses some of that original evidence. Call this the Evidence-Loss Account. To avoid confusion arising from the ambiguity of the notion of “having less evidence,” we need to distinguish between evidence possession and support strength. If Bob has weaker evidence for a claim than Coco does, it’s natural to describe Bob as having less evidence than Coco. But Bob does not necessarily have less evidence in the sense relevant to the Evidence-Loss account; indeed, in that sense, Bob may possess more evidence than Coco. For example, suppose Bob and Coco both see someone who looks like Don standing on the corner. Suppose Bob knows, but Coco doesn’t, that Don is out of the country. Bob’s extra information about Don’s whereabouts reduces the support the shared evidence provides for Bob for the proposition that Don is standing on the corner. Thus, even though the (overall) strength of Bob’s evidence is less than Coco’s, Bob possesses more evidence than Coco. (Correspondingly, possessing less evidence can also increase the strength of one’s evidence.) The sense of “having evidence” relevant to the Evidence-Loss Account has implications regarding support strength, but it is the stronger, literal notion of possessing evidence that is at issue. On the conception of evidence I find most plausible (and will defend in §5), evidence is propositional and the evidence a person possesses is a set of propositions. According to the Evidence-Loss Account, when someone acquires higher-order

---

20 Christensen (2010: 195) briefly addresses this step as well. I should note that in the context in which this quote from Kelly appears, Kelly is addressing an issue arising from peer disagreement, whereas I’m focusing on fallibility more generally. However, Kelly only bolsters the point in the text that there seems to a tension between Evidentialism and fallibility, when he says (2013: 47): “Of course, it is difficult to make sense of the idea that someone who possesses entailing evidence should invest less than maximal credence in the entailed proposition; indeed, orthodox theories of evidential probability would seem to rule this out (at least in cases in which the person is certain of the entailing evidence itself).”
defeating evidence, a proposition that was previously in her evidence set no longer is. It is in this sense that she no longer possesses some of her original evidence.

Why might anyone believe the Evidence-Loss Account? There are at least four reasons, and we’ve already seen one: it solves the puzzle faced by the Evidentialist. As this Evidence-Loss Argument appeals to Evidentialism, it is an Evidentialist-based reason to believe the Evidence-Loss Account.

There is another Evidentialist-based reason that appears to favor the Evidence-Loss Account. A prior question to whether Evidentialism can explain higher-order defeat is what resources are available to Evidentialism to do so. The elements of Evidentialism – its basic resources – are nicely captured in the standard characterization of the distinction between undercutting and rebutting defeat: undercutting defeat “attacks the connection between the evidence and the conclusion, rather than attacking the conclusion itself [as rebutting defeat does].”21 This characterization appeals to three elements, and these plausibly exhaust the Evidentialist’s basic resources: evidence, conclusions, and the connection between the two. Several philosophers have argued that higher-order defeat is distinct from rebutting and undercutting defeat.22 If this is right, then given that the latter two elements are already spoken for – that each is already associated with a specific kind of defeat – it is a reasonable hypothesis that higher-order defeat is associated with the final element. That is, it seems reasonable to suppose that higher-order defeat attacks the evidence. And this is one way of articulating what the Evidence-Loss Account says. Call this the “elements of Evidentialism” argument.

The third reason in favor of the Evidence-Loss Account is that there is a precedent for it in the literature. Timothy Williamson (2000: 219) describes an example of evidence loss occasioned by acquiring evidence that one misassessed one’s original evidence:

I see one red and one black ball put into an otherwise empty bag, and am asked the probability that on the first ten thousand draws with replacement a red ball is drawn each time. I reply ‘1/2^{10,000}'. Part of my evidence is the proposition $e$ that a black ball was put into the bag; my calculation relies on it. Now suppose that on the first ten thousand draws a red ball is drawn each time, a contingency which my evidence does not rule out in advance, since its evidential probability is non-zero. But when I have seen it happen, I will rationally come to doubt $e$ …

Williamson concludes, “Thus $e$ will no longer form part of my evidence.” In this case, something very improbable occurs, given the subject’s assessment of what he takes to be his evidence. When he


observes this, he then loses evidence as a result of becoming reasonably worried that he misassessed his evidence. Thus, sometimes, at least, acquiring evidence that one misassessed one’s evidence can result in evidence loss.

Finally, there is a parallel between a sort of dogmatic response to higher-order evidence and the reasoning involved in the so-called Dogmatism Paradox for knowledge, and an explanation of the illegitimacy of the former reasoning that appeals to evidence loss parallels the standard explanation of the illegitimacy of the latter reasoning. Moreover, the two main hypotheses that compete with the Evidence-Loss Account on this score are unacceptable by the Evidentialist’s lights.

Several philosophers have cautioned against dogmatic responses to higher-order evidence. Suppose Andy were to reason as follows: “If I were suffering from hypoxia, I would very likely have made the wrong inference. But, in fact, I made the right inference: after all, I concluded that I have enough fuel to fly to Hawaii and that is just what my evidence supports—indeed, entails! So, I must not be suffering from hypoxia.” Learning that she’s probably suffering from hypoxia calls into question her ability to properly assess her evidence, so it would be illegitimate (question-begging or dogmatic) for her to appeal to her assessment of the evidence to conclude that she’s actually not suffering from hypoxia.

One may be tempted toward a similar form of reasoning in response to evidence against one’s knowledge. Consider Gilbert Harman’s (1973: 148) statement of this reasoning:

If I know that $b$ is true, I know that any evidence against $b$ is evidence against something that is true; so I know that such evidence is misleading. But I should disregard evidence that I know is misleading. So, once I know that $b$ is true, I am in a position to disregard any future evidence that seems to tell against $b$.

Both sorts of reasoning – regarding knowledge or one’s assessment of one’s original evidence – are illegitimate. What explains why these sorts of reasoning are illegitimate?

The standard explanation of what is wrong with the dogmatic reasoning about knowledge is that it incorrectly presupposes that once one has knowledge, one may always thereafter use it in one’s reasoning. It’s true that when one knows a proposition, one also knows that evidence against that proposition is misleading. But once one acquires such misleading evidence, one may no longer know the original proposition, and one may not reasonably dismiss this evidence as misleading. In effect, Harman’s explanation relies on the fact that acquiring new evidence can cause you to lose knowledge.

---


The Evidence-Loss Account’s explanation of why the dogmatic reasoning is illegitimate in the higher-order evidence case appeals to the fact that acquiring new evidence can cause you to lose evidence. Although Andy’s evidence previously entailed that she had enough fuel, this does not warrant her in disregarding further evidence, since getting that further evidence may change what her evidence entails. But it can only do this by causing her to lose evidence. And this is just what the Evidence-Loss Account implies: it is illegitimate for Andy to appeal to her original evidence because once she acquires the higher-order evidence, she no longer possesses that evidence. (Likewise for disagreeing peers, and people who learn they have been slipped reason-distorting drugs, etc.)

There are two alternative explanations of the problem with Andy’s reasoning in the literature, but the Evidentialist cannot accept either of them. First, there is David Christensen’s *Bracketing Explanation*. In accounting for her higher-order evidence, Andy must bracket her original reasons; she is barred from giving a certain part of her evidence its due. This account is a sort of “virtual” rather than “genuine” evidence-loss account: Andy does not actually lose evidence, but what she is justified in believing after she acquires the higher-order evidence depends only on a proper subset of her evidence, the evidence left over once she brackets her original evidence. The second explanation – what I’ll call the *Non-Monotonic Explanation* – is due to Thomas Kelly. Focusing on peer disagreement, Kelly explicitly discusses the parallel between the Dogmatism Paradox and the dogmatic reasoning involving higher-order evidence. Explaining why this reasoning is illegitimate, Kelly (2013: 45) writes:

> After I add the fact that you believe as you do to my stock of evidence, it will no longer be reasonable for me to believe that p, given what is my total evidence. And if it’s no longer reasonable for me to believe that p, then I lack any rational basis for inferring that your sincere testimony is misleading evidence.

The idea is that the support enjoyed by a subject’s belief before she was aware of the disagreement will simply be lost, in the standard way, once she learns of the disagreement. As a result, subsequently appealing to that belief to dismiss her peer’s belief will be irrational.

All three explanations have the same surface structure: the dogmatic reasoning is unreasonable because (some of) the premises of the dogmatic reasoning are unjustified. However, the Evidentialist can accept neither the Bracketing Explanation nor the Non-Monotonic Explanation. Remember, the problem cases are Entailment Cases: the original evidence entails the relevant beliefs, hence the Evidentialist must say that the original evidence supports the beliefs. No matter how much evidence one brackets, and no matter how much additional evidence gets added to one’s body of evidence, if one does not lose any evidence, then by the Evidentialist’s lights the

---

premises of the problematic reasoning remain justified. Indeed, Christensen would say that Andy’s evidence does still support her beliefs because, on his view, higher-order evidence defeats while leaving evidential support relations intact; his account is anti-Evidentialist. And while Kelly wants to provide an Evidentialist account, he admits that Entailment Cases give rise to a genuine puzzle. The Evidence-Loss Account is an Evidentialist account that can solve that puzzle.

To summarize: There appears to be a tension between respecting one’s evidence and respecting one’s fallibility. In particular, the Evidentialist faces a puzzle in accommodating higher-order defeat in Entailment Cases. The Evidence-Loss Account can solve this puzzle. And there seem to be three additional points in its favor. First, it coheres with the standard characterizations of defeat. Second, there is a precedent for it in the form of examples of evidence-loss occasioned by acquiring reason to believe one misassessed one’s evidence. Finally, it explains, in terms acceptable to the Evidentialist, what is wrong with dogmatically dismissing higher-order evidence, and this explanation mirrors the standard explanation of what’s wrong with the dogmatic reasoning found in the Dogmatism Paradox. Thus, the Evidence-Loss Account should strike Evidentialists and non-Evidentialists alike as a serious hypothesis.

4. Against the Evidence-Loss Account

In spite of its attractions, the Evidence-Loss Account is problematic. In this section, I consider three objections to this account. The first should not worry the Evidentialist much, I will argue. The second and third may not be decisive, but they strike me as strong objections.

The first objection can be put in terms of a question pertaining to the Hypoxia example: According to the Evidence-Loss Account, when Andy acquires the higher-order evidence that defeats the justification for her belief, she thereby loses some of the evidence that justified that belief. Question: Which evidence does Andy lose? Given the description of the case, there are four main options. But picking one rather than the other seems arbitrary. Without a principled answer to this question, the account seems to be in trouble.

This is a difficult question that should be addressed if the Evidence-Loss Account is to be defended further. However, not having a precise, general answer to this question is not a major embarrassment to the proponent of the Evidence-Loss Account, since the main alternative to this account faces a similar question, and also lacks a precise, general answer. This alternative is

28 Kelly (2013: 46).
29 The fact that the two explanations mirror each other is noteworthy because it allows what appear to be similar phenomena to be given a similar (perhaps, at a certain level, the same) explanation. This potential for explanatory unification counts in the Evidence-Loss Account’s favor.
Christensen’s Bracketing Account, according to which, accommodating higher-order evidence requires bracketing some of one’s original evidence. This account needs to address the question: Which evidence, or how much of it, must Andy bracket? It’s unlikely that any precise, general answer will be forthcoming. Thus, the Evidence-Loss Account can respond to this objection by adopting a companions-in-guilt strategy. This is not to say it’s off the hook, but only that this concern should be seen as a challenge rather than a strong objection.

According to the second objection, the Evidence-Loss Account gives the wrong sort of response to the puzzle. On the Evidence-Loss Account, Andy’s belief turns out to be irrational because she loses evidence. But it’s unclear that this is the right sort of explanation. A common concern that arises when we consider our fallibility is that we make “performance” errors while attempting to properly form beliefs. Hume (1978: 180), for instance, insists that although deductive rules of inference are certain and infallible, our application of them is not. Likewise, when Andy gets the hypoxia warning, this is worrisome because it makes it likely that she made a faulty inference. And, it seems, this sort of worry leaves the agent’s evidence untouched.

Compare Williamson’s ball example (§3) with the Hypoxia example. In Williamson’s case, it does seem that the subject should give up his belief in his evidence because what he observes is incredibly unlikely if his evidence is in fact true. Since incredibly unlikely things are probably false, what he took to be his evidence is probably false, too. This is why claiming that he loses evidence is plausible. But nothing like this occurs in Hypoxia. Andy takes her evidence to support a proposition, while having strong evidence that her take on the evidence is flawed. Unlike the subject in Williamson’s case, what Andy observes is not incredibly unlikely if her evidence is true. Putting the comparison slightly differently: the subject in Williamson’s example acquires reason to think he has misassessed what his evidence is, whereas Andy only acquires reason to think she has misassessed what her evidence supports. Whereas continuing to accept what she took her evidence to support when she gets the hypoxia warning is unreasonable, it seems unproblematic for her to rely on her judgments about what evidence she possesses. Thus, while it’s true that Williamson’s example sets a

---

30 This account underlies Christensen’s Bracketing Explanation. See §3.

31 Christensen (2010: 213; 2011). Schoenfield (2014) offers an account of bracketing. Evaluating it is beyond the scope of this paper, but if there is a principled account of bracketing, but no principled answer to the question facing the Evidence-Loss Account, so much the worse for the Evidence-Loss Account, and Evidentialism’s attempt at resolving the tension.

32 Schechter (2013: 448) makes a similar point.
precedent for the Evidence-Loss Account, this account doesn't properly explain all problematic Entailment Cases. 33

According to the third objection, the Evidence-Loss Account requires an implausible “purpose-relative” conception of evidence possession. Consider a case of peer disagreement. If you and your friend discover, while out to dinner, that you disagree about your shares of the bill, even if you thereby lose evidence, you can still legitimately appeal to this evidence for some purposes. Suppose it’s part of your original evidence that your bill, before tip, is $60. You can reasonably reject your friend’s claim that your share is $400, since the bill itself is only $60. 34 But, to avoid condoning dogmatism, the Evidence-Loss Account prohibits appealing to this evidence to dismiss a more plausible position (that, with tip, your shares are $35, rather than $36, say). Now suppose you’re out to dinner with two people (suppose that the third person is paying separately): one claims that your shares are $400 and the other claims that your shares are $35. Arguably, you can dismiss the $400 verdict, but not the $35 verdict. What explains this? The natural answer is that for the purposes of ruling out the ridiculous response you can reasonably appeal to your original evidence, but for the purposes of ruling out the sensible response you cannot. 35 The Evidence-Loss Account explains the former fact by positing evidential support, hence evidence possession, and it explains the latter fact by positing evidence-loss. Thus, the Evidence-Loss Account is committed to a purpose-relative account of evidence possession: you possess the evidence about the bill for the purpose of dismissing the ridiculous response, but not for the purpose of answering the sensible one.

While the claim that evidence possession is purpose-relative is not obviously incoherent, it does seem highly implausible. If a subject possesses some evidence for some purpose, then, in some

---

33 This is also a reason not to accept the Evidence-Loss Account on the basis of the “elements of Evidentialism” argument. Evidentialism has an open slot for higher-order defeat, but without a compelling explanation of why higher-order defeat should fill that slot, that fact is only a very weak reason to accept the Evidence-Loss Account.

34 Cf. Christensen (2007). Also, nearly all of the papers in Feldman & Warfield (2010) make a point like this.

35 Although it seems to me that Christensen (2007: 200-1) implicitly endorses appealing to the original evidence in a case like this as part of the explanation of why rejecting the ridiculous response can be rational, some of his theoretical commitments – e.g., his commitment to Independence – may prohibit this appeal. Things are more complicated in Christensen (2011) because although he doesn’t straightforwardly endorse this appeal in his discussion of similar cases, he admits that there are unresolved issues about the scope of Independence (see Christensen (2011: 18). In any case, those who share Christensen’s commitments may find my argument objectionable. But, the argument from the Evidence-Loss Account to purpose-relativity does not depend on this example. It only requires (i) that there is some purpose for which the agent who’s suffered higher-order defeat can use her original evidence, but (ii) that she could not use it for that purpose if she did not possess it. For instance, after providing her with the hypoxia warning, air traffic control might ask Andy how much fuel she has. It seems perfectly legitimate for her to appeal to her evidence that the tank is ¼ full. (Fill in the details however you like to focus on another piece of her evidence.) If that’s right, then in order to capture this point, the Evidence-Loss Account requires a purpose-relative conception of evidence possession. Thanks to Katia Vavova for discussion here.
intuitive sense of ‘available,’ the evidence is available to her. So why does she possess the evidence for one purpose but not another? The Evidence-Loss Account introduces relativity into evidence possession where there appears not to be any.

Of course, the proponent of the Evidence-Loss Account can insist that I’ve mischaracterized which evidence has been lost, and that, properly characterized, any trace of purpose-relativity will vanish. However, while this response may only depend on particular judgments about which evidence one loses in response to higher-order evidence, the more heavily the account relies on such judgments, the more urgent the need to answer the first objection begins to look. In other words, this response invites the question: Why does the subject lose that evidence rather than this evidence? Again, this is a question that we would do well to avoid having to answer.

There is, what might seem to be, a precedent for the purpose-relativity of evidence possession. Several philosophers have argued that knowledge and/or justification are ‘interest-relative,’ in the sense that whether a subject knows or is justified in believing something depends in part on her interests or what is at stake for her. If this view is correct, then at least some epistemic notions are interest-relative. And Jason Stanley suggests that all epistemic notions are interest-relative, including evidence. According to Stanley, this implies that one’s practical situation will affect the “evidential standing one has with respect to one’s belief that p.” Inspired by Stanley’s position – *Strong Interest-Relativism* – a proponent of the Evidence-Loss Account may dig in his heels and defend the purpose-relativity of evidence possession.

No doubt that by some people’s lights we need not take the Evidence-Loss Account seriously if it relies on Strong Interest-Relativism. But it’s important to note that even if Strong Interest-Relativism could be defended, the claim that evidence possession is purpose-relative is even more radical. First, recall that in §3 we distinguished between evidence possession and support strength, noting that the Evidence-Loss Account is primarily concerned with the former. Some of Stanley’s remarks seem to suggest that what is relative to one’s interests is evidential support strength, not evidence possession. But that sort of view will not justify the purpose-relativity of evidence possession. Stanley also discusses Williamson’s E=K account of evidence – according to which, one’s evidence is just what one knows. If that view is true and whether one knows something depends on one’s interests, then whether one possesses it as part of one’s evidence will also depend on one’s interests. Yet, this still doesn’t guarantee purpose-relativity. For the proponent of the Evidence-Loss Account must still claim that evidence-possession can vary while one’s interests are

---

36 See Hawthorne (2005), Stanley (2005), and Fantl & McGrath (2009).
38 Ibid. 180.
held fixed. There may be a lot or a little at stake for Andy in determining how much fuel she has or for you in figuring out what your share of the bill is, but once we hold fixed your stakes, Strong Interest-Relativism will have fixed what evidence you possess, whereas the Evidence-Loss Account will still leave this purpose-relative. So, in this respect, the Evidence-Loss Account is more radical than Strong Interest-Relativism.

Thus, although the Evidence-Loss Account may have an elegant explanation of what’s wrong with dismissing evidence of one’s fallibility, it purchases that elegance with a purpose-relative conception of evidence possession. Given this, and given the second objection, I think we should reject the account.

5. Defending the Puzzle

Evidentialism implies that higher-order evidence defeats justification by changing evidential support. Given the Entailment-Support principle, the Evidence-Loss Account seemed to be the Evidentialist’s only option for accommodating higher-order defeat. With reason to reject that account, we are apparently left with two options: reject Evidentialism or reject Entailment-Support. In this section, I briefly argue that if these are indeed our only options, then we should reject Evidentialism because on the most plausible version of Evidentialism, Entailment-Support will be true. The argument will be brief because it mostly depends on claims that others have already defended. At the end of the section, however, I suggest that even if this argument fails, Evidentialism may still have trouble accommodating higher-order defeat.

The argument goes as follows: Evidential support should be understood probabilistically, and evidence is propositional. Given these two claims, we should accept Entailment-Support.

Evidentialism analyzes justified belief in terms of evidential support. This is plausible because typically the greater epistemic justification one has for a belief, the more likely it is to be true and, likewise, typically the greater evidential support one has for a belief, the more likely it is to be true. When we ask about the strength of the evidence, i.e., how much the evidence tells for or against a hypothesis, we are asking how probable that hypothesis is given that evidence. This suggests that evidential support should be understood probabilistically.

There are several reasons why evidence should be understood propositionally. First, evidence is what inferences to the best explanations explain, and only propositions are the proper objects of explanation. Second, evidential support is a logical relation between something and hypotheses. Logical relations include entailment, perhaps the logical connectives, and relations of probability. As a logical

---

39 A reader who is convinced that Evidentialists cannot accommodate higher-order defeat may want to skip this section.
40 This section draws heavily from Williamson (2000: chs. 9, 10).
relation, then, that something must be a proposition. Third, thinking of evidential support probabilistically, two main types of quantities will interest us, the conditional probability of a hypothesis on evidence and the likelihood of the hypothesis—\( P(b|e) \) and \( P(e|b) \), respectively.\(^{42}\) Both rely on evidence being the sort of thing that can have a probability, and the sort of thing that can have a probability is a proposition. (It makes no sense, e.g., to discuss the probability of appearing redly to me or the appearance of red, etc.) Finally, our evidence can rule out hypotheses by being inconsistent with them, but only propositions can be inconsistent with hypotheses.\(^{43}\)

Now suppose that, indeed, evidential support should be understood probabilistically, and evidence is propositional. And suppose that \( E \) is S’s evidence and that \( E \) entails some contingent proposition \( H \). Does \( E \) support \( H \)? It’s hard to see how it could not. For one thing, the conditional probability of \( P(H|E)=1 \). But suppose entailment doesn’t suffice for support. The following seems plausible: if \( P(H_1|E)>P(H_2|E) \), then \( E \) provides more support for \( H_1 \) than for \( H_2 \). Now suppose that \( E \) entails \( H_1 \). If entailment doesn’t suffice for support, then we must endorse the following implausible triad: (1) \( E \) provides as much support for \( H_1 \) as it can for any hypothesis – indeed, \( H_1 \) can enjoy no greater support from any evidence; (2) \( E \) provides more support than it could for any hypothesis it doesn’t entail; (3) but it doesn’t necessarily support that hypothesis. And if evidential support doesn’t require the supported hypothesis to have probability 1 on the evidence, as it plausibly doesn’t, we’re forced to say that there are two hypotheses \( H_1 \) and \( H_2 \) such that (1) \( E \) supports \( H_2 \), but \( H_2 \) is not certain on \( E \), yet (2) \( H_1 \) is certain on \( E \) and \( E \) does not support it. This just seems implausible. Greater support of some hypothesis given some evidence seems tied to greater probability of that hypothesis given that evidence.\(^{44}\)

One might grant the Entailment-Support principle, but express a lingering worry as follows: the Evidentialist need not worry about cases like Hypoxia because although evidence can entail and thereby support propositions, the specification of Andy’s evidence is too broad. While it’s true that appearances and the like are not evidence, evidence is not directly about the external world. In Hypoxia, Andy’s evidence includes propositions like \( \text{the tank is ¾ full} \), but her evidence should be specified more narrowly, i.e., internally – e.g., \( \text{the fuel gage seems to read ¾ full} \). Otherwise, our evidence provides us with certainty about the external world, which is unobtainable, as Descartes taught us.

\(^{42}\) For a nice introduction to likelihoods see section 1.2 of Sober (2008).

\(^{43}\) See Williamson (2000: 196) for an articulation of the relevant sense in which evidence is inconsistent with hypotheses.

\(^{44}\) Fitelson (2007) argues for the claim that if \( E \) provides conclusive evidence (i.e., entailing evidence) for \( H_1 \), but non-conclusive evidence for \( H_2 \) (when these are all contingent claims), then \( E \) favors \( H_1 \) over \( H_2 \). His arguments, therefore, support an analogous, contrastive entailment-support principle.
This is a concern about evidence’s extent. Although this is a legitimate worry, it doesn’t undermine the puzzle. Insofar as this puzzle depends on the Entailment-Support principle, it depends on evidence being propositional. It doesn’t, however, depend on a particular account of evidence’s extent. Change the specification of Andy’s evidence however you think appropriate. That evidence will entail some proposition, and she can acquire evidence that she bungled her reasoning about that proposition. Given the judgment that maintaining her belief is irrational, Evidentialists will still face the same puzzle.

Finally, for those who remain unmoved by this defense of the Entailment-Support principle, it’s worth noting that Evidentialists may face a version of the puzzle even if that principle is false.45 Recall, several philosophers argue that higher-order defeat differs from both rebutting defeat and undercutting defeat. Higher-order defeat does not defeat by providing evidence that the defeated belief is false (like rebutting defeat does), by “attacking the connection” between one’s evidence and one’s belief (like undercutting defeat does), or by attacking the evidence (as in Williamson’s example). If higher-order defeat differs in these ways from these other sorts of defeat, how else can it undermine justification? This question points to the grain of truth in the “elements of Evidentialism” argument: Evidentialists can only accommodate defeat that attacks conclusions, evidence, or the connection between these, and higher-order defeat doesn’t necessarily do any of these.46 It’s tempting, then, to conclude that higher-order evidence defeats while leaving evidential support intact even in non-Entailment Cases. And this is just what Christensen (2010: 197) argues. Suppose my evidence strongly supports, but does not entail, some conclusion, and I adopt belief in that conclusion on the basis of an inference to the best explanation, only to discover that I was slipped a reason-distorting drug before I made the inference. Christensen argues that, in this sort of case, I cannot rationally maintain my belief even though my evidence still supports it. Why does the evidence still support the belief? Because the evidence itself is not in question (as in Williamson’s case). And the connection between that evidence and my belief remain incredibly strong, since these connections do not depend on any claims about me or my current reasoning abilities and the higher-

45 As Conee & Feldman (2008) express their evidentialist commitments, they are likely to remain unconvinced. This final argumentative maneuver may, therefore, be of special interest to those who find an evidentialist view like theirs compelling.

46 Another “element” in some versions of Evidentialism (for instance, Bayesianism) is the agent’s standards for evaluating evidence or her priors, but appealing to this element won’t solve the puzzle. For one thing, Bayesians should already be on board with the conclusion that Evidentialism cannot accommodate defeat in Entailment Cases because they are committed to the Entailment-Support Principle. But, in any case, insofar as I have a handle on the notion of one’s evidential standards being “attacked,” for the same reasons it seems wrong to say that higher-order evidence always defeats by attacking the evidence, it seems wrong to say it always defeats by attacking an agent’s standards. (It’s not as if Andy must adopt some wholly new take on the import of evidence in response to learning that she’s not properly assessing her evidence. She simply can’t fully trust her take on the evidence, regardless of how she tends to evaluate it.)
order evidence does not break those connections. Thus, if the proper moral of the “elements of Evidentialism” argument is that higher-order defeat cannot undermine justification on Evidentialism, or if Christensen is right that higher-order defeat leaves evidential support relations – even inductive ones – intact, then Evidentialists will face a version of the puzzle from §2, even if the Entailment-Support principle is false.

If our only options are to reject Evidentialism or the Entailment-Support principle, I think that, one way or another, we must reject Evidentialism because the most plausible version of Evidentialism includes the Entailment-Support principle. But even if Evidentialism doesn’t include the Entailment-Support principle, there’s reason to think that Evidentialism should still be rejected.

6. Ideal vs. Non-Ideal Epistemological Theory

We began with an apparent tension between the Evidentialist Dictum and the Fallibility Imperative. Evidentialism was one plausible elaboration of the Dictum, and higher-order defeat was taken to represent the Imperative. If what I’ve argued is correct, Evidentialism cannot accommodate higher-order defeat. The tension looks to be genuine. However, I don’t think we should reject the Evidentialist Dictum or the Fallibility Imperative. Since both clearly represent important standards of epistemic evaluation, our theorizing must make room for both. I will conclude with a brief sketch of how it can do this.

Begin with a story. Suppose that you get lost on the way to meeting a friend in a part of town you rarely visit. Your friend knows this part of town like the back of her hand, so when you call her for directions, she offers you two options: she can give you directions for the complicated, though much shorter route, or for the simple, though much longer route. Although you’d prefer to take the shorter route, since you’re likely to get lost again on a complicated route, you judge that you’d rather hear about the simpler route.

If some conditions of your situation were better, for instance, if your navigational abilities were stronger or more reliable in this situation, then you might have been able to pursue the more preferable option. However, given your limitations, it’s better for you to take the less preferable option. It’s a general fact that in your decision-making you must *compensate* for your limitations.

In this story, based on your situation, your friend is in a position to offer two different prescriptions, and which prescription she should provide depends on your abilities. Although this story was told in the second-person, it suggests a general question that might guide one’s third-person prescription-making practice: What should he do, given his situation and his abilities? More to the point, there is an epistemic analogue to this practical question: What should he *believe*, given his
situation and his abilities? This epistemic question provides the basis for the distinction between ideal and non-ideal epistemological theory.47

Ideal theory begins by articulating which abilities an ideal agent has. Perhaps, for instance, an ideal agent has perfect access to what she believes and what her evidence is, she never makes mistaken inferences, she is conceptually omniscient, and she knows all of this about herself. Ideal theory may also place constraints on the kinds of situations an ideal agent can encounter. One’s situation might include that one will violate certain optimal prescriptions, for instance, but an ideal agent would never be in \textit{that} sort of situation.48 Once the abilities have been fixed and the set of possible situations has been determined, prescriptions will be made. There may be a general, all-encompassing prescription or there may be more particular prescriptions relative to individual situations. This project has the potential to cleanly and simply systematize the rules of rational belief.

Non-ideal theory is likely to be considerably messier. Again, which prescription is appropriate will depend on an agent’s abilities and her situation. This project counts as ‘non-ideal’ because these abilities and situations will be \textit{limited} or \textit{imperfect} in various respects. There are two ways to proceed in non-ideal theory: either take the ideal theory as prior, and only make prescriptions relative to that theory (insofar as this is possible), or begin with the agent’s imperfections, and make prescriptions independently of the ideal theory. Although these procedures are distinct, one is not necessarily better than the other, and which procedure one adopts will partly be determined by one’s relative satisfaction with the current state of ideal theory. If you think there is no satisfactory ideal theory in the offing, then you may want to engage in the second sort of non-ideal theory; if you think the opposite, you may want to engage in the first sort.

Recognizing a distinction between ideal theory and non-ideal theory may help dissolve the tension between the Evidentialist Dictum and the Fallibility Imperative. On my proposal, the canons of reasonable belief include two sorts of prescriptions.49 First, there are \textit{prescriptions of perfection}, the

47 Sprinkled throughout David Lewis’s epistemological work are brief comments on the relationship between Bayesian ideals and how we fallible creatures should manage our beliefs. (See e.g., (1986a, 1986b, 1999).) These comments are expressed in similar language across these different contexts, suggesting that Lewis had some view in mind about the relationship between epistemic ideals and ideal and non-ideal agents. As far as I know, no one has developed these ideas in print, however, the picture they suggest seems plausible. The following sketch is inspired by Lewis’s comments and also borrows some of his language.

48 The story about the lost driver can illustrate this point. Your friend can offer you two different sets of directions. But the fact that \textit{those} are the available prescriptions depends on the fact that, in getting lost, you’ve already failed to conform to the ideal prescriptions. Since an ideal agent will never fail to conform to the ideal prescriptions, she will never be faced with such a situation.

49 My arguments support a view similar to the views of Lasonen-Aarnio (2014) and Schoenfield (forthcoming b). Lasonen-Aarnio argues that there is no non-paradoxical notion of justification/rationality that can capture both Evidentialist friendly
prescriptions ideal agents heed. Second, there are *counsels of compensation*, the prescriptions non-ideal agents, like us, ought to heed. If the Evidentialist Dictum captures an epistemic truth, it is a truth that belongs to ideal theory; it is a prescription of perfection. Not just *any* wise man proportions his belief to the evidence; only someone who can be absolutely confident what his evidence is, what his evidence supports, etc. should and will always believe what his evidence supports.\(^{50}\) We, who are not so wise, do well to avoid only believing what (we think) our evidence supports. As we are limited in various ways (e.g., we cannot be absolutely confident what our evidence is and what it supports), we need to remain open to the possibility that we have made mistakes in managing our beliefs and we need to heed evidence that this possibility is actualized when we acquire it. This is not what the ideal agent would do; it is, in some sense, imperfect.\(^{51}\) But it’s imperfection borne of imperfection. It’s compensation for our known fallibility. And it seems to be the reasonable way to proceed. The Fallibility Imperative is a counsel of compensation.

Finally, the theoretical upshot of fallibility, as I see it, is basically the one Hume already identified in the *Treatise*, and the one defended by many in the literature on higher-order evidence, namely, Calibrationism.\(^{52}\) Hume (1978: 180) claims, “all knowledge degenerates into probability; and this probability is greater or less, according to our experience of the veracity or deceitfulness of our understanding, and according to the simplicity or intricacy of the question.” Roughly, your confidence in a proposition should reflect your expected reliability regarding that proposition: if, independently of your first-order reasoning on a question, your expected degree of reliability concerning some proposition is some value \(r\), then \(r\) is the level of confidence you should adopt in that proposition.\(^{53}\) On this account, you attempt to take your evidence into account by judging which

---

\(^{50}\) Something similar should be said, I think, about the limits to modesty mentioned in the introduction. See fn. 6.

\(^{51}\) For what it’s worth, I disagree with Christensen’s (2007a) arguments that even an ideal agent would heed misleading higher-order evidence. Given the metacognitive capacities such an agent would have, she would view all of her inferences with as much confidence (probably more) as we view the simplest inferences we ever make. Any evidence suggesting she bungled her reasoning, no matter how strong, would be on a par with (probably weaker than) the evidence a mature adult gets when a child shakes her head after the adult proclaims that \(1 + 1 = 2\). We are so much more confident that we are right than that the child is, and this is how the ideal agent would be about all of her reasoning. Does this fly in the face of Williamson’s anti-luminosity argument? I don’t think so. That argument is best understood as applying not to all possible beings, but to beings with discriminative capacities relevantly similar to ours, as Srinivasan (2015) explains.

\(^{52}\) The name is due to Schoenfield (forthcoming a). But see fn. 4 for other references.

\(^{53}\) Cf. Schoenfield’s (forthcoming a) statement of Calibrationism. The framework I’ve sketched here allows the Calibrationist to answer an objection raised by Schoenfield. She argues that if this sort of Calibrationism is true, then it is the *only* rational requirement on belief, and that this is objectionable. But even if it were objectionable, the Calibrationist can
proposition your evidence supports, and then you match your degree of belief in that proposition with your expected degree of reliability. If you comply with this principle, you will sometimes fail to believe only what your evidence supports. However, you would have sometimes failed in this way if you hadn’t followed it. Complying with this principle is just the sort of compensation that our departure from ideal rationality requires.

References
Elga, A. (ms.) Lucky to be Rational.

reject that conclusion if she takes her principle to be merely a counsel of compensation while allowing that there are other rational requirements among the prescriptions of perfection.