A Dilemma for Calibrationism[1]

MIRIAM SCHOENFIELD
The University of Texas at Austin

The aim of this paper is to describe a problem for *calibrationism*: a view about higher order evidence according to which one's credences should be calibrated to one's expected degree of reliability. Calibrationism is attractive, in part, because it explains our intuitive judgments, and provides a strong motivation for certain theories about higher order evidence and peer disagreement. However, I will argue that calibrationism faces a dilemma: There are two versions of the view one might adopt. The first version, I argue, has the implausible consequence that, in a wide range of cases, calibrationism is the *only* constraint on rational belief. The second version, in addition to having some puzzling consequences, is unmotivated. At the end of the paper I sketch a possible solution.

1. Calibrationism

Calibrationists think that there is a very tight connection between how reliable you (reasonably) take yourself to be, and the credence that it is rational to adopt. To illustrate, consider the following case:[2]

> HYPOXIA: You are a pilot flying to Hawaii. You suddenly realize that you might not have enough fuel to get there. If you don't have enough fuel, you need to land right away. Fortunately, you have evidence that entails either that you have enough fuel to get to Hawaii (H) or that you don't (~H). You do some calculations and breathe a sigh of relief. It looks like you'll make it. You then get a message from ground control: "Dear pilot, you are flying at an altitude that puts you at great risk for hypoxia, a

---

[2]   This case is originally from Elga (ms.). It has since been discussed in a variety of papers including Christensen (2010ᵃ), Lasonen Aarnio (2014) and Elga (2013).

condition that impairs your ability to reason properly. People flying at your altitude, who do the sorts of calculations you just did, only reach the conclusions entailed by their evidence 50% of the time."

HYPOXIA is a case in which you have *higher order* evidence. I won't attempt to give a precise characterization of higher order evidence, but roughly, higher order evidence is evidence that bears on evidential relations, or evidence that bears on what is *rational*.[3] Here, and throughout, when I talk about the belief states that are *rational* for an agent to adopt I mean the belief states that the agent is *propositionally justified* in adopting. Furthermore, I will be assuming that the belief states that are propositionally justified for an agent are those that the agent's evidence supports.

I'll use the term "first order evidence" to refer to any evidence you have that is not higher order evidence. In HYPOXIA, your first order evidence includes, for example, the evidence provided by the gas gauge, whereas your higher order evidence includes the message you got from ground control concerning your potentially impaired reasoning.

It seems intuitive to many people that, in HYPOXIA, you should only have a 0.5 credence in H. But why 0.5? It's very tempting to think that the reason you should have a *0.5* credence in H has something to do with the fact that you learned that you are *50%* reliable in these circumstances. Indeed, if you think you should have a 0.5 credence in H in the case above, you probably also think that if, instead, you'd learned that you were *55%* reliable, your credence should be 0.55, if you learned that you were *60%* reliable, 0.6, and so on. So an attractive way of explaining the judgment in HYPOXIA appeals to a bridge principle connecting your expected degree of reliability to the credence that it is rational for you to adopt. In its most general form, it will look something like this:

> *Bridge*: If your expected degree of reliability concerning whether p is *r*, *r* is the credence that it is rational for you to adopt.

Your expected degree of reliability concerning whether p (at t) is the probability that it is rational for you assign to the proposition that your judgment concerning whether p (at t) is correct.

What exactly is a judgment? I will be using the term in a technical way. Your judgment, in my sense, is the proposition that you regard, or *would* regard, as most likely to be correct, on the basis of the first order evidence

---

3   See Christensen (2010a), Kelly (2005) and Feldman (2005) for a more detailed characterization of higher order evidence.

alone.[4] In HYPOXIA, your judgment is H, the proposition that you have enough fuel to land in Hawaii. It is important to realize that one can judge that p without believing (or assigning a high credence to) p. If you assign a 0.5 credence to H in HYPOXIA, you still judge that H since H is what you *would* believe if you only had the first order evidence.

Note, however, that *Bridge*, on its own, won't guarantee the result that you should reduce your confidence in HYPOXIA. To see why, imagine a cocky pilot who responds to the hypoxia warning as follows: "Look, I recognize that, in circumstances like this one, I'm generally only 50% likely to arrive at the correct conclusion. Indeed, had you given me this warning before I did my calculations, I would have only been 50% confident that I would judge correctly. However, I have gone through the reasoning. In fact, I *do* have enough fuel to land in Hawaii. Furthermore, I judged that I do. Therefore, this must be one of the occasions in which I *do* judge correctly and so I'm much more than 50% likely to be right in *this* instance." What the cocky pilot demonstrates is that, as many parties in the higher order evidence debate have observed[5], to get the desired result in HYPOXIA, we need to add to *Bridge* another principle:

> *Independence*: When calculating your expected degree of reliability, you need to do so in a way that is independent of the reasoning in question.

The thought behind *Independence* is that you can't appeal to your reasoning about the first order matter when evaluating the likelihood of this particular judgment's correctness. *Independence* raises all sorts of interesting questions: How much exactly must you set aside? What if the reasoning in question can't be neatly excised from your other beliefs? I won't delve into these issues here. Elga (2007) and Christensen (2011) discuss *Independence* at length. My aim is only to note that for a bridge principle connecting your credences to your expected degree of reliability to motivate the view that you should have a 0.5 credence in HYPOXIA, some independence principle along these lines is necessary. One thing that it's important to be clear about is that, however the details of what needs to be set aside get filled in, the fact that, in HYPOXIA, you have first order evidence that, *in fact*, supports your judgment doesn't bump up your expected degree of reliability. For suppose it did. Then we should expect your expected degree of reliability to be higher than 0.5. After all, given that your judgment is actually supported by your first order evidence, it's much more that 0.5 likely to be correct! The fact that your expected degree of reliability in HYPOXIA is exactly 0.5, even though your judgment is supported (indeed entailed!) by your first order evidence

---

[4]   I borrow this usage from Horowitz and Sliwa (ms.), and Weatherson (ms.).

[5]   For example, Elga (2007) and Christensen (2011).

demonstrates that your expected degree of reliability isn't improved by having a judgment that the first order evidence actually supports.

> The view that combines *Bridge* and *Independence* I will call "Calibrationism."

> *Calibrationism*: If, independently of the first order reasoning in question, your expected degree of reliability concerning whether p at time t is *r*, *r* is the credence that it is rational for you to adopt at t.[6]

Sometimes, for ease of expression, I will leave out the independence clause, and simply talk about the agent's "expected degree of reliability." This should always be understood to be calculated in a way that is independent of the reasoning in question.

Two notes about calibrationism:

First, conciliatory views about peer disagreement follow from calibrationism. When you disagree with someone who is your epistemic peer, it is true that, independently of the reasoning in question, your expected degree of reliability is 0.5. (Indeed, on at least some definitions of peerhood (such as Elga's (2007)), an epistemic peer *just is* someone such that, independently of the reasoning in question, your expected degree of reliability, in instances of disagreement with them, is 0.5.) Since calibrationism says that, when your expected degree of reliability is 0.5, your credence should be 0.5, it has the consequence that, in cases of peer disagreement, you should have a 0.5 credence. Note that calibrationism *isn't* just the view that you ought to reduce confidence *somewhat* when your expected degree of reliability is low. Views like Tom Kelly's (2010) "Total Evidence View" might recommend that you reduce your confidence in HYPOXIA, or in instances of peer disagreement, but not that you dive all the way down to 0.5. Such a view is not calibrationist. Calibrationists require that when your expected degree of reliability is 0.5, your credence be *exactly* 0.5.

Second, there are various bells and whistles one might want to add to the principle (for example, to avoid skeptical results), and I will discuss some of these later in the paper. However, for most the arguments I will be advancing, all that is relevant is the core commitment, which calibrationists share, that, in

---

[6]   Principles along these lines have been called by a variety of names ("Integration" (Christensen (2010a)), "Calibration Rule" (White (2009)), "Calibration" (Horowitz and Sliwa (ms.)), and "Higher Order Defeat" (Lasonen-Aarnio (2014))). The calibrationist position, as applied to cases of peer disagreement, has been called "conciliationism" (e.g. Christensen (2011) and "the equal weight view" (Elga (2007))). I chose "calibrationism" for this paper since the calibration theme seems to come up in many of the discussions, and having a name in the form of an "-ism" is useful for my purposes. Versions of calibrationism have been defended by Christensen (2007, 2010a), Elga (2007) and Horowitz and Sliwa (ms.). Calibrationist views about peer disagreement, in particular, are also defended by Feldman (2006) and Christensen (2011), among others.

a wide range of cases, one's credence should be calibrated to one's expected degree of reliability, where this is calculated independently of the reasoning in question. Thus, for simplicity, I will focus primarily on the principle above.

The purpose of this paper is to present a dilemma for calibrationism. There are two versions of the view one might adopt. The first version, I argue, has the implausible consequence that, in a wide range of cases, calibrationism is the *only* constraint on rational belief. The second version, in addition to having some puzzling consequences, is unmotivated. This means that the project of explaining and motivating our judgments in cases like HYPOXIA is much more difficult than one might have thought. At the end of the paper, I sketch a possible solution to the problem.

2. Setup for the Dilemma

The calibrationist thinks that your credence should equal your expected degree of reliability, r. The following question arises for the calibrationist: what proposition should you have credence r *in*? The most natural thought is that you should assign credence r to your judgment. This suggests:

> *J-calibrationism*: If p is your *judgment*, and r is your expected degree of reliability (where this is calculated independently of the reasoning in question), then the rational credence for you to assign to p is r.

However, there are problems with J-calibrationism, and these concerns have motivated some calibrationists to endorse:

> *E-calibrationism*: If p is *the proposition best supported by your first order evidence*, and r is your expected degree of reliability (where this is calculated independently of the reasoning in question), then the rational credence for you to assign to p is r.

As an illustration of the difference between these two views, let's consider a milder version of the hypoxia case. As before, you judge that H on the basis of your first order evidence. You then learn that, due to hypoxia's effects on reasoning, people in your circumstances only judge correctly *60%* of the time. As a matter of fact, your judgment that H was not supported by your first order evidence. In this case, J and E-calibrationism yield different results. J-calibrationism says that you should have a 0.6 credence in *H*, since H is what you judged. E-calibrationism says that you should have a 0.6 credence in *~H*, since ~H is what the first order evidence supports.

I will argue that both J-calibrationism and E-calibrationism should be rejected. In conversation, I have heard some people express the view that E-calibrationism is a non-starter. What could possibly motivate the thought

that, if you judge H, and your expected degree of reliability is 0.6, you should assign a 0.6 credence to ~*H*? I will take this up later in the paper, but, if this is your view, then just think of the "dilemma" for calibrationism as a "problem" for calibrationism, which, on this picture, will be equivalent to J-calibrationism, or one of its variants. My own view is that the problems with J-calibrationism motivate a serious consideration of E-calibrationism, which is why I devote a section of the paper to a discussion of this view. Those who are already convinced of the falsity of E-calibrationism may wish to skip that section.

It is a bit tricky to determine which version of calibrationism better aligns with the views of the various players in the higher order evidence debate since much of the discussion of higher order evidence has been focused on cases of peer disagreement. In the standard cases of peer disagreement, you and a peer disagree over whether p and it is stipulated that your expected degree of reliability is 0.5. Thus, in peer disagreement cases, J-calibrationists will recommend a 0.5 credence in your judgment and E-calibrationists will recommend a 0.5 credence in what the first order evidence supports. Your judgment and the proposition supported by the first order evidence may or may not coincide. However, each view will recommend a 0.5 credence in *a* member of {p, ~p}, and since your credences in these propositions must sum to one, it follows that they will each recommend a 0.5 credence in *both* members of {p, ~p}. Thus, in many of the standard disagreement cases, the two views yield the same verdicts. Nonetheless, Elga's (2007) suggests sympathies with J-calibrationism, and J-calibrationism is described (though not endorsed) by White (2009) and Weatherson (ms.). On the other hand, E-calibrationism is defended by Horowitz and Sliwa (ms.) and Christensen's (2011) view seems to best align with E-calibrationism.[7]

I am not claiming that J and E calibrationism are the *only* principles available to the calibrationist. As we proceed, I will discuss some

---

[7]   Christensen says:

> "It would, I think, be very damaging to conciliationist views [if they were committed] to insisting that one's original evidence was irrelevant to the rationality of the beliefs one ended up with after making one's conciliatory epistemic compromise. Nevertheless, I think that the most plausible version of the conciliationist position does not have this consequence…" (4)

J-calibrationism has exactly the consequence that Christensen says the most plausible view does not have. For, according to J-calibrationism, the rational credence, in the relevant cases, is determined completely by what you judged and the value of *r*. What the original evidence supported has no bearing on the credence it is rational for you to adopt after making your "conciliatory epistemic compromise." Christensen has confirmed, in personal correspondence, that he is opposed to J-calibrationism and that he is most sympathetic to a view *in the vicinity* of E-calibrationism (though he may wish to modify certain details, related to the bells and whistles I mentioned earlier, which need not concern us here).

alternative variants of these principles. However, I structure the argument around these two versions of the view because I think that the distinction between calibrating to one's judgment, and calibrating to what the first order evidence supports, is an important and illuminating one. I will argue that neither view is acceptable as it stands. I will then suggest an alternative way of capturing the thought that, I think, is motivating calibrationists.

3. The First Horn: J-Calibrationism

The problem with J-calibrationism is that it makes rationality too cheap.[8] My challenge to the J-calibrationist will be as follows: can the J-calibrationist account for cases in which someone has an irrational belief even though the person is doing what J-calibrationism recommends? If the answer is no, then J-calibrationism is in trouble, for it leads to the absurd conclusion that *all* that is required for rational belief is that we match our credences to our expected degree of reliability, where this is calculated in a way that is independent of the reasoning in question. Below, I will present an argument for the claim that J-calibrationism has this absurd consequence. However, there are ways to resist this argument and so my ultimate conclusion will not be that that the J-calibrationist is committed to the view that, *in all cases*, J-calibrationism is the only constraint on rational belief. Rather, I will show that the view is committed to the claim that, in *an unacceptably wide range of cases*, J-calibrationism is the only constraint on rational belief.

Such a conclusion may come as a surprise, for you might be thinking that calibrationism is a principle that applies to a special class of cases: *cases of higher order evidence*. If J-calibrationism is only about a very special kind of case, you might wonder how it could have *anything* to say about what constraints there are on rational belief more generally. Part of the lesson to be learned from the discussion that follows is that thinking of higher order evidence cases as special is a mistake. It is extremely rare that we are *not* in a higher order evidence case. Once we appreciate the breadth of cases which count as cases of "higher order evidence," the fact that J-calibrationism has very wide range implications will not be surprising.

To get the main idea on the table, let me begin by presenting the argument for the strong conclusion: that according to J-calibrationism, conforming to J-calibrationism is the only requirement on rational belief.

---

[8] My worry about J-calibrationism is similar in spirit to Tom Kelly's (2010) worries about conciliatory responses to disagreement.

(P1) According to J-calibrationism, if an agent who judges that p is highly confident that p, and conforms to J-calibrationism, her expected degree of reliability (calculated independently of the reasoning in question) must be high. (If it were low, being highly confident that p would violate J-calibrationism.)[9]

(P2) According to J-calibrationism, if the agent's expected degree of reliability (calculated independently of the reasoning in question) is high, then the rational credence for her to assign to her judgment (p) is high. (This follows directly from J-calibrationism.)

(C) According to J-calibrationism, if an agent who judges that p is highly confident that p, and conforms to J-calibrationism, the rational credence for her to assign to her judgment (p) is high. (From 1, 2 via hypothetical syllogism.)

In other words, according to J-calibrationism, any agent that judges that p, and believes (or is highly confident in) p, is rational, as long as she is conforming to J-calibrationism. There are no other constraints on rational belief. This is highly implausible. As I mentioned, I do not intend to defend (C) since I think that there are ways the argument above can be resisted. But in seeing how the argument can be resisted, we will see that the range of cases in which an agent conforming to J-calibrationism can form an irrational belief is extremely, and unacceptably, narrow.[10] I will consider two objections to P1 and two objections to P2.

---

[9]   Note that the "According to J-calibrationism" clause, at the beginning of P1, is not necessary. However, adding the clause makes the structure of the argument more clear.

[10]   It is worth noting that a version of argument above would go through even if the calibrationist rejected *Independence*, and was only committed to *Bridge*. Simply remove the parenthetical "independently of the reasoning in question" clauses and the result would be an argument from *Bridge* to (C). However, *Independence* actually has quite an important role to play. First, as we'll see, if someone endorsed *Bridge*, but was willing to deny *Independence*, there would be a way to modify the J-calibrationist principle to avoid the unattractive result. Second, (C) would not be nearly as problematic if we rejected *Independence*. For if turned out that what one's expected degree of reliability is depended heavily on what the first order evidence actually supported, one's expected degree of reliability could reflect various substantive rational requirements. To see this, imagine an extreme anti-*Independence* view according to which one's expected degree of reliability concerning p *just is* the credence in p that one's first order evidence supports. Then, the claim that calibrationism is the only constraint on rational belief would amount to the claim that the only constraint on rational belief is that one's credences be supported by one's first order evidence. This avoids making the notion of rationality trivial. However, as noted earlier, this option isn't available to the calibrationist who wants to use calibrationism to motivate judgments in cases like HYPOXIA. For if we abandon *Independence*, we won't get the desired result in HYPOXIA.

3.1. Objections to P1

> (P1) According to J-calibrationism, if an agent who judges that p is highly confident that p, and conforms to J-calibrationism, her expected degree of reliability (calculated independently of the reasoning in question) must be high.

The J-calibrationist might reject P1 by claiming that an agent can conform to J-calibrationism, *and* be highly confident that p, *without* having a high expected degree of reliability. This will happen in cases in which, setting aside the reasoning in question, the agent *has no expected degree of reliability at all*. When might this happen? I will consider two possibilities. First, the objector might claim that an agent lacks an expected degree of reliability when she fails to explicitly consider how reliable she is. The second proposal relies on the calibrationist's commitment to *Independence*. The objector might claim that an agent lacks an expected degree of reliability in cases in which so much needs to be set aside in order to evaluate her reliability, *independently of the reasoning in question*, that there's not enough left to ground an expected degree of reliability.

Objection 1: Explicit Consideration

The first proposal has two problems. First, on most accounts of what it is to have a credence, an agent can have credences at a given time in all sorts of propositions that she is not consciously entertaining. So the view that having an expected degree of reliability requires consciously entertaining a proposition about one's reliability is somewhat *ad hoc*.

Second, requiring that one consciously entertain ones expected degree of reliability results in a principle that is too weak to do the work calibrationists want it to do. For this view would have the consequence that, even in cases in which we have good evidence that our track records are quite bad, that many people disagree with us, and so on, we could still be rationally confident in our judgments, so long as we do not explicitly consider our expected degree of reliability.

Objection 2: Huge Propositions

The second proposal is inspired by a discussion in Elga (2007). Elga claims that the question of how reliable you are, setting aside a huge swath of your beliefs, may have no answer. He says: "These questions have no answers because the information to be set aside is enmeshed in too much of your reasoning to be cleanly factored off" (496). If he's right, then there may be no answer to the question: Setting aside my first order reasoning, what is my expected degree of reliability concerning my judgment about the

proposition that includes, for instance, *all of my beliefs about the external world,* or *all of my moral beliefs*.

However, even if Elga is right in thinking that I have no expected degree of reliability concerning propositions that encompass a huge swath of my web of beliefs (call these propositions "huge propositions"), this will not be much help. For if the problem with the argument above is that, in cases in which the belief in question is huge, the agent has no expected degree of reliability, a modified version of P1, which applied only to non-huge propositions, would yield a modified, but still unacceptable, conclusion: that the only rational constraint on *non-huge* beliefs is J-calibrationism.

3.2. Objections to P2

> (P2) According to J-calibrationism, if the agent's expected degree of reliability (calculated independently of the reasoning in question) is high, then the rational credence for her to assign to her judgment (p) is high.

Since P2 directly follows from J-calibrationism, both of the objections to P2 that I will consider go by way of suggesting that J-calibrationism should be modified in some way. I will set aside the question of whether or not the versions of J-calibrationism that I will consider are sufficiently well motivated so as to constitute an acceptable solution to the dilemma. For even if they *are* well motivated, there are other difficulties such views will face.

The first suggestion is that J-calibrationism be modified so as to apply only to cases in which agents have *positive reason* for their expected degree of reliability. The second suggestion is that we weaken the J-calibrationist principle so that it provides only a necessary, (rather than a necessary and sufficient), condition on the rationality of one's credences.

Objection 1: PRJ-Calibrationism

The first modification is inspired by Christensen (2011).[11] Christensen says that if your expected degree of reliability is low, but that is only because you *lack* reason to think that you are reliable, you ought not match your credence to your expected degree of reliability. For the principle to kick in, on this view, it must be that your evidence gives you *positive* reason to judge that your expected degree of reliability is as it is. (Christensen introduces this condition to avoid certain skeptical worries discussed in his paper).

---

[11]   Although the view under consideration was *inspired* by a comment from Christensen, this is not Christensen's view. As I mentioned earlier, Christensen is more sympathetic to E-calibrationism.

> *PRJ-calibrationism*: If p is your *judgment*, and r is your expected degree of reliability (where this is calculated independently of the reasoning in question), *and you have positive reasoning for assigning yourself expected degree of reliability r,* then the rational credence for you to assign to p is r.

PRJ-calibrationism doesn't yield the result that merely having a high expected degree of reliability is enough to rationalize a belief in one's judgment. For PRJ-calibrationism only says to match your credence to your expected degree of reliability, *when you have positive reason for having the expected degree of reliability that you have*. PRJ-calibrationism may, therefore, allow for the possibility that there are agents who conform to PRJ-calibrationism but nonetheless have irrational beliefs.[12]

Nonetheless, PRJ-calibrationism will not provide a satisfactory solution to our problem. For the view *does* have the consequence that, whenever an agent *does* have positive reason to assign a high expected degree of reliability, whatever conclusions she comes to, no matter how wacky, will be rational. You might think this is okay. For, perhaps, in most cases in which we're inclined to think of a belief as irrational, the agent lacks positive reason for her high expected degree of reliability. In fact, however, this is not so.

It is not at all unusual for us to have a great deal of positive evidence suggesting that we will judge correctly in some specific instance. We are constantly accumulating very large positive track records. They may not be as salient to us as bad track records *precisely because* they are so common. Furthermore, our positive track records are not just track records of being "generally reliable." They can be quite specific. I have lots of positive evidence that I will be good at doing simple inductive reasoning, involving coins, in the morning, after having my coffee. That doesn't mean that when I see a coin land heads 100 times, and I commit the gambler's fallacy, judging that the next flip will land tails, the resulting belief is rational. I have lots of positive evidence that I'm good at judging when and where colloquia will be, on the basis of announcements made in the department. If there is an announcement that all talks at 3 on Wednesdays will be epistemology talks, and I firmly believe, on the basis of that announcement alone, that the

---

12   Note that if PRJ-calibrationism is to help the calibrationist, it must be possible for an agent to have a *high* expected degree of reliability without any positive reason for having a high expected degree of reliability. *Can* an agent rationally have a high expected degree of reliability while lacking positive reason to think that she is reliable? It may depend on exactly what is required to have "positive reason" (and we'll get to this in a moment), but, on at least some views (Wright (2004) and White (2006)), we have a default warrant for the belief that we are reliable. If this is right, then, even lacking positive evidence, agents can reasonably assign themselves a high expected degree of reliability.

upcoming epistemology talk must be at 3 on a Wednesday (thereby doing the equivalent of affirming the consequent), my belief is unsupported by my evidence, despite my good track record. Given that we do usually have a great deal of positive evidence for our reliability, PRJ-calibrationism will not be of much help in explaining ordinary instances of irrational belief. PRJ-calibrationism, while consistent with the view that agents who conform to PRJ-calibrationism can form irrational beliefs, only allows for this in the rather special cases in which such agents lack positive reason for having a high expected degree of reliability.

One might try to get around this worry by imposing very strict requirements on what counts as "good reason" or "positive reason." For example, one might think that the reasons I described for thinking that I am reliable in the cases above are simply not sufficient. Alternatively, one might argue as follows: consider the case of my fallacious inference to the conclusion that the talk on Wednesday will be an epistemology talk. Perhaps the fact that I have the first order evidence that I do (which *doesn't* actually support my judgment) means that I *don't* have good reason to think that I am reliable in this instance.[13] The problem with both of these suggestions is that, if the calibrationist adopted either of them, the view would no longer do the work calibrationists want it to do.

To see this, let's consider a variant of the colloquium case. This time, when I reflect on my track record, I realize that I've made mistakes about the time and place of colloquia on the basis of departmental announcements a large percentage of the time. (Perhaps I even form some hypotheses about why this is: e.g., thinking about colloquia makes me excited and clouds my thinking.) Calibrationists will want to say that reflection on this *bad* track record would make it irrational for me to be confident in my judgment even though, let's say, as a matter of fact, in this particular instance, I judged correctly.

Now, if you thought that the evidence I got from my *positive* track record in the original case was insufficient to count as *positive reason to think I'm reliable* then it would be very hard to see why the evidence I got from my *negative* track record in the revised version of the case would be sufficient to count as *positive reason to think I'm unreliable*.[14]

Similarly, if one thought that having first order evidence which doesn't support my judgment means that I lack a good reason to think that I am *reliable* in the good-track-record case, it would be hard to see why, in

---

[13] Thanks to an anonymous referee for pointing out this potential response on behalf of the PRJ-calibrationist.

[14] Note that the same problem would arise if we tried to restrict the applicability of calibrationism to cases in which an agent is justified in having a *precise* expected degree of reliability. In many of the cases in which calibrationists want their principle to apply, like the bad track record case above, the agents' expected degree of reliability will be imprecise.

the bad-track-record case, having first order evidence that *does*, in fact, support my judgment wouldn't mean that I lacked a good reason to think I was *unreliable*. Since the PRJ-calibrationist needs to claim that having the first order evidence that I do in the bad-track-record case *doesn't* prevent me from having good reason to think I'm unreliable, it would be unacceptably *ad hoc* to say that having the first order evidence that I do in the good-track-record case, prevents me from having positive reason to think I'm reliable.

It is worth noting that the source of the problem facing the PRJ-calibrationist, here, stems from the *Independence* requirement, which says that an agent's expected degree of reliability must be calculated in a way that is independent of the reasoning in question. It also relies on this requirement functioning in the specific way I mentioned earlier: whether the reasoning you did was *in fact* good or bad, or whether you have first order evidence that *in fact* supports your judgment, or doesn't, can't bear on what your expected degree of reliability is. Nor can it bear on whether you have *good reason* to have the expected degree of reliability that you have. Why think the *Independence* requirement should function in this way? Because, as I argued above, if it didn't, the PRJ-calibrationist couldn't use it to motivate the kinds of claims she wants it to motivate in cases like HYPOXIA. So a version of PRJ-calibrationism with enough teeth to do the necessary work will indeed have the consequence that, in an unacceptably wide range of cases, it is the only constraint on rational belief.

In sum, like the huge-propositions proposal, PRJ-calibrationism only manages to carve off a very specialized class of cases in which agents who satisfy the principle can form irrational beliefs. In this case, it is the class of cases in which agents have no positive evidence for their reliability. I suggested that cases like this are quite rare. Thus, a modified argument would lead to the unacceptable conclusion that, in cases in which agents have positive evidence for their reliability, the only constraint on rational belief is PRJ-calibrationism.

Objection 2: WJ-Calibrationism

The second objection to P2 has us replace J-calibrationism with a weaker version of the principle.

> *Weak J-calibrationism (WJ-calibrationism)*: A necessary condition on the rationality of a belief state is that, for any p such that the agent judges that p, the agent's belief state assigns a credence to p that equals her expected degree of reliability (r) (calculated independently of the reasoning in question).[15]

---

[15]   Thanks to Gerard Vong for this suggestion.

A belief state is a function that assigns to every proposition (that the agent has some attitude towards) the doxastic attitude that the agent bears towards that proposition. A credence function is a paradigmatic example of a belief state, but on a tripartite model, a belief state will be a function that assigns propositions to one of {belief, disbelief, suspension of judgment}.

To see how WJ-calibrationism differs from the original principle, recall:

> *J-calibrationism*: If p is your judgment, and r is your expected degree of reliability (calculated independently of the reasoning in question), then the rational credence for you to assign to p is r.

J-calibrationism is stronger than WJ-calibrationism. For J-calibrationism says that, if p is your judgment, assigning credence r to p is a necessary *and* sufficient condition for having a rational credence in p. WJ-calibrationism, in contrast, is only committed to the claim that calibrating is a *necessary* condition on the rationality of a belief state. Thus, it is open to the WJ-calibrationist to claim that, if your initial judgment was not supported by the first order evidence (FOE), then, even if you calibrate, your credence will be irrational.

However, WJ-calibrationism does not provide a satisfactory solution to the problem. This is because WJ-calibrationism is committed to the existence of problematic epistemic dilemmas. Here is a quick version of the argument for the claim that WJ-calibrationism is committed to the existence of epistemic dilemmas. (I will argue that the dilemmas are problematic later in this section.) Consider any agent with expected degree of reliability r who judged inappropriately that p (by "inappropriate" I mean that the judgment is not in accord with the first order evidence). What belief state is it rational for this agent to adopt? If there is some belief state that is rational for her to adopt it either assigns credence r to p or it assigns a credence to p that is not r. The claim that the rational belief state assigns a credence to p that is *not* r is inconsistent with WJ-calibrationism. For WJ-calibrationism says that a belief state which includes a credence in your judgment that doesn't equal your expected degree of reliability isn't rational. Thus, if there is any belief state that it's rational for such an agent to adopt, it assigns r to p. So, assuming there is such a belief state, we can conclude that it's rational for an agent to assign credence r to p even if p was not supported by the FOE. But now WJ-calibrationism faces the same problems as J-calibrationism. For the view says that, even if the FOE doesn't support one's judgment, it's rational to assign credence r to one's judgment. So, if WJ-calibrationism is to avoid the problems with J-calibrationism, it must claim that, in at least some cases in which your judgment is inappropriate, there is no belief state that is rational for you to adopt. Hence, an epistemic dilemma.

Here is a more careful version of the argument: For simplicity, I will assume evidentialism: the claim that which belief state it is rational for an agent to adopt supervenes on her evidence.[16] (So, for example, the evidentialist will say that if the agent's evidence supports belief in p, it is rational for her to believe p.) Later I will talk about how this assumption can be relaxed. The argument begins by showing that the following four claims are inconsistent:

- *Inclusion*: If an agent judges that p, her evidence includes the fact that she judges that p.

- *WJ-calibrationism*

- *~J-calibrationism*: It's not always the case that: If p is your judgment, and r is your expected degree of reliability, the rational credence for you to assign to p is r.

- *No Dilemmas*: Every body of evidence rationalizes some belief state.

The argument is as follows: If the four claims above were true then,

(1) There is some case in which an agent judges that p, and yet, assigning credence r to p is not rational. (By *~J-calibrationism*)

Let's call an agent in such a case "Zoey" and suppose that Zoey judged that p.

(2) If Zoey assigns credence r to p, Zoey's belief state is not rational. (By (1))

(3) The fact that Zoey judged that p is part of Zoey's evidence. (By *Inclusion*)

(4) Some belief state is rationalized by Zoey's evidence. (By *No Dilemmas*)

(5) The belief state that's rational given Zoey's evidence, assigns credence r to p. (By (3), (4) & *WJ-calibrationism*)

The inference to (5) is justified as follows: Zoey's evidence includes the fact that she judged that p (by (3)). Thus, if there is any belief state that's

---

16   Recall that, as I am using the term, the belief state that is rational for an agent to adopt is just the belief state that she is propositionally justified in adopting.

rational given Zoey's evidence, it's one that assigns credence r to p (by WJ-calibrationism). (4) tells us that there *is* some such belief state. So we know that it assigns credence r to p. (2) and (5) are inconsistent. Thus, the four claims above are inconsistent.

Which of these four claims should the WJ-calibrationist reject? I will now argue that the WJ-calibrationist should reject *No Dilemmas*. I will then show why positing epistemic dilemmas in this context is problematic.

First, the WJ-calibrationist must accept *Inclusion*. For if the agent's judgment *isn't* part of her evidence, we will violate evidentialism (the claim that the rationality of an agent's belief state supervenes on her evidence). This is because WJ-calibrationism is a constraint on rational belief. A belief state that doesn't satisfy the requirements of WJ-calibrationism isn't rational. Whether an agent satisfies WJ-calibrationism depends on what her judgments are. But if facts about what she judged are not part of her evidence then the rationality of her belief state depends on facts that are not part of her evidence. Second, the WJ-calibrationist must, of course, accept WJ-calibrationism. Third, the WJ-calibrationist must accept ~J-calibrationism. Otherwise, the WJ-calibrationist will face all of the problems that the J-calibrationist faces. Thus, the WJ-calibrationist must reject *No Dilemmas*.

Note that a variety of non-evidentialist views are subject to the same argument. For example, suppose you think that what it's rational to believe doesn't supervene on the agent's evidence alone. Rather, it supervenes on a combination of the agent's evidence *and prior probabilities*, or the agent's evidence *and standards of reasoning*. This won't make a difference. Simply replace instances of "evidence" in the argument above with "evidence + priors," "evidence + standards," or whatever else you think might be relevant. Nonetheless, for simplicity, I will be assuming that evidentialism is true for the remainder of this section.

Before I argue for the claim that rejecting *No Dilemmas* in this context is problematic, let me address an objection. You might think that the WJ-calibrationist should reject the inference to (5) in the argument above. Recall,

(3)   The fact that Zoey judged that p is part of Zoey's evidence.

(4)   Some belief state is rationalized by Zoey's evidence.

(5)   The belief state that's rational given Zoey's evidence, assigns credence r to p. (By (3), (4) & *WJ-calibrationism*)

Perhaps the WJ-calibrationist can say the following: "The belief state that is rationalized by Zoey's evidence assigns credence r to ~p (and credence 1-r to p). Nonetheless, I'm not saying anything that is inconsistent with WJ-calibrationism. Because although Zoey *actually* judges that p, the belief

state that is rational for her to adopt is one in which she judges that ~*p* and assigns credence r to ~p. There is no problem with saying that the belief state that is rational for the agent to adopt includes a judgment that differs from what the agent currently judges."

Note that the distinction between J and WJ-calibrationism is a distinction between a narrow and wide scope version of calibrationism, with "rational" taking narrow scope in J-calibrationism, and wide scope in WJ-calibrationism. The line of thought that I am suggesting on behalf of the WJ-calibrationist is analogous to what one might say about other wide scope requirements. Consider, for example, *modes ponens*. Modes ponens might be thought of as a wide scope principle which says that a necessary condition on a rational belief state is that, if it includes the beliefs *A* and *If A then B*, it also includes the belief that *B*. However, it's not supposed to follow from this that, if an agent believes *A* and *If A then B*, it's rational for the agent to believe *B*. For it is possible that the agent's evidence actually supports believing ~*A* and ~*B*. Is the modes-ponens wide-scoper committed to epistemic dilemmas? Surely not. In a case in which an agent believes *A* and *If A then B*, despite the fact that ~*A* and ~*B* are supported by her evidence, the belief state that is rational for her to adopt may be one in which she believes ~*A*, *If A then B*, and ~*B*. Thus, there is no violation of the wide scope principle, no epistemic dilemmas, and we are still not committed to the narrow scope principle that says that any agent who believes *A* and *If A then B*, is rational in believing *B*. The thought on behalf of the WJ-calibrationist is that a similar approach might be taken in cases in which an agent judges inappropriately.

The proposed solution, however, will not work. The reason for this is that a judgment is not a belief, a credence, or anything that you will find in an agent's belief state. What an agent judges is a matter of what belief state she is disposed to be in given a body of evidence that may differ from her actual one. What the agent *actually* believes given her evidence doesn't entail anything about what she *would* believe if she had a body of evidence that differs from her current body of evidence. So when the objector above said: "the rational belief state given the evidence contains the judgment that ~p and assigns credence r to ~p", she was making a category mistake. There is no such thing as a belief state containing a judgment.

Note that it is very important for calibrationism that it be stated in terms of judgments rather than beliefs or credences. Calibrationists appeal to the notion of the agent's *expected degree of reliability*. But remember that an agent's expected degree of reliability is, roughly, the likelihood that her *judgments* are correct in the relevant circumstances. If we changed the notion of expected degree of reliability, such that what was relevant was the likelihood that an agent's *beliefs* were correct, we wouldn't get the right results in

cases like HYPOXIA. To see this, consider a version of the hypoxia case in which the order of events is reversed. First, the pilot gets evidence that people in her circumstances only perform the relevant calculations correctly 60% of the time. Then she gets her FOE, does her calculations, and judges that H (correctly, let us suppose). The calibrationist wants to say that such an agent should assign a 0.6 credence to H. But note that at no time did the agent ever *believe* H. So if her expected degree of reliability was defined as the likelihood that her *beliefs* would be correct, there would be no reason for her to assign a 0.6 credence to H. Since she has no belief about whether H, the reliability of her beliefs would be completely irrelevant. Finally, just to emphasize how un-belief-like judgments are, note that an agent can judge that p while being more confident in ~p than p. This will occur in cases in which one's higher order evidence suggests that one's judgments are *anti-reliable*. We will discuss such cases in more detail shortly.

You might think that the proposed solution could be modified as follows: rather than saying that the agent's *belief state* contains a judgment, we can say that the belief state that is rational assigns credence r to the proposition best supported by the FOE *and*, furthermore, one is rationally required to judge in accord with the FOE. But now the resulting view contains two components: a claim about which belief states are rational given the evidence, and a claim about the rationality of some other thing—a judgment. On this view, the claim about the rationality of a belief state is *E-calibrationism* and I will discuss the problems with E-calibrationism later in the paper.

There is, however, a sense in which I think a solution along the lines of the above proposal is on the right track. The reason we are running into trouble is that the object of evaluation is *a belief state*. If we changed the object of evaluation, there would be room to make the sorts of moves that this objector finds tempting. The solution that I will propose at the end of this paper is in the spirit of this suggestion. I will suggest that to make sense of calibrationism we need to move to a picture on which calibrationism isn't a principle about the rationality of belief states. Calibrationism is about another matter altogether. But this is getting ahead of ourselves. The important point for now is that, as long as we think of calibrationism as a principle about the rationality of *belief states*, the proposed solution doesn't work. Since judgments are not components of belief states, one can't say that the rational belief state is one which includes this or that judgment. Thus, the argument above is sound and the WJ-calibrationist is committed to the existence of epistemic dilemmas.

I will now argue that the kind of dilemmas that the WJ-calibrationist is committed to pose a problem. But let's begin by getting clear on what exactly the epistemic dilemma proposal amounts to. Recall that J-calibrationism faced problems because, according to J-calibrationism, regardless of

how an agent judged, it is rational for her to assign credence r to her judgment. I think that the most plausible version of WJ-calibrationism that avoids these problems is the following:

> *Dilemma WJ-calibrationism (DWJ-calibrationism)*: If an agent appropriately judges that p, (that is, p is best supported by her first order evidence), it is rational for her to assign credence r to p. If an agent inappropriately judges that p, (that is, p is not best supported by her first order evidence), there is no belief state that is rational for her to adopt.

Why might you accept DWJ-calibrationism? The thought would be that the theory of rationality is a theory of *ideal* rationality: it describes how ideally rational agents form beliefs. We shouldn't expect a theory of rationality to tell us what belief states are rational for agents who already displayed irrationality through inappropriate judgments. Such agents have put themselves in a bind for which there may be no way out: an epistemic dilemma.

Positing epistemic dilemmas, however, will pose a variety of problems, the first of which is that there may be general reasons for denying the existence of epistemic dilemmas. But there are also specific reasons why the calibrationist should avoid positing epistemic dilemmas in cases in which an agent judged inappropriately. I will describe three such reasons below.

First, consider:

> ORACLE TESTIMONY: I judge that tomorrow will be sunny on the basis of meteorological evidence (this is my FOE). The oracle then comes down from the heavens and tells me the following: "First, I'd like to let you know that your meteorological evidence does not support the proposition that it will be sunny. Quite the contrary, in fact. It supports the proposition that it will be rainy. In addition, I should alert you to the fact that your mistaken judgment is not an isolated incident. Whenever you get strong meteorological evidence that it's going to rain tomorrow, you become so depressed that you end up inventing very silly arguments for the conclusion that it will be sunny. On the other hand, if you get strong meteorological evidence that it will be sunny, your fear of disappointment leads to your judging that it will be rainy!"

Note that there is nothing paradoxical about learning that one's *judgments* are anti-reliable. There is, perhaps, something paradoxical about learning that one's *beliefs* will be anti-reliable, regardless of how one attempts to correct for their anti-reliability. But judgments, recall, are not beliefs.

According to DWJ-calibrationism, my evidence in ORACLE TESTIMONY doesn't rationalize any belief state since it includes an inappropriate judgment. However, to the extent that we have an independent grip on the notion of evidential support, it seems clear, that, in such a case, my total

evidence supports believing that it will rain. Think about it this way: in the case we are considering, my evidence includes first order evidence in support of rain, the trustworthy-as-one-can-get oracle who *tells* me that my first order evidence supports rain, and my judgment that it will be sunny.[17] Given the oracle's testimony about my anti-reliability, the fact that I judged that it will be sunny is *also* evidence that it will be rainy. So *all* of my evidence strongly favors rain. How could such a body of evidence *not* support believing that it will rain?

Second, DWJ-calibrationism will lead to a violation of single premise closure, the claim that if one knows p, and one knows that p entails q, then (if one believes q), one knows q. For suppose that, in ORACLE TESTIMONY, I know two further facts: that my FOE entails sunniness or entails raininess,[18] and that the proposition that the FOE supports is the proposition that the FOE entails. So I know:

(1) My FOE supports raininess (on the basis of the oracle's testimony).

(2) [My FOE supports raininess] ⊃ it will rain.

And yet I don't know that it will rain! I don't know that it will rain because according to DWJ-calibrationism my evidence doesn't support the belief that it will rain.[19] (I am relying on an assumption, widely held by evidentialists, that one can't know that p unless one's evidence supports belief in p.)[20] Thus, we have a case in which I don't know something that I know is entailed by something I know.

Finally, DWJ-calibrationism seems to license some awkward reasoning. To see this, consider a hypoxia case in which the pilot gets evidence that her expected degree of reliability is 0.6. Prior to getting that evidence, let us suppose, she judged, and believed H (that she has enough fuel to make it to

---

[17] Note that it is still true that I judge that it will be sunny even after I hear the oracle's testimony. This is because an agent's judgment is the proposition she *would* regard as most likely to be correct on the basis of the first order evidence alone. Even after hearing from the oracle that my first order evidence supports the proposition that it will be rainy, it's still true that I would regard sunniness as most likely if I only had the first order evidence to go on (which does not include the oracle's testimony).

[18] It's rare that we have evidence that entails facts about how tomorrow's weather will be, but it's not impossible. Nonetheless, if you prefer, you can substitute a case in which it is more plausible that one's evidence entails a proposition that it supports.

[19] This is because, according to DWJ-calibrationism, it's not rational for me to believe that it will rain, and since we are assuming evidentialism, this means that my evidence doesn't support the belief that it will rain.

[20] Conee and Feldman (2004, p.93), for example, say that knowledge requires that a belief be "well founded" and that well foundedness requires that one's belief be supported by the evidence.

Hawaii). And let us also assume that she knows that her FOE entails either H or ~H and that it supports what it entails. She is now contemplating whether she should reduce her confidence. The DWJ-calibrationist tells her that, if, in fact, she judged *appropriately* she should reduce her confidence to 0.6. However, if she judged *inappropriately*, there is no belief state that is rational for her to adopt. It seems like the pilot might respond as follows:

> AWKWARD REASONING: You're telling me that it's only rational for me to adopt a 0.6 credence if my initial judgment that H was supported by my FOE. Let's focus for a moment on the bearing that your view has on the rationality of my actions. I know that if I have a 0.6 credence in H, I'll return home right away. I wouldn't risk a 40% chance of death for a weekend vacation in Hawaii. On the other hand, if I maintain my high credence in H, I will keep on flying. Given a plausible theory about the connections between the rationality of beliefs and the rationality of actions, what you're essentially telling me is that *I should go home only if I have enough fuel to make it to Hawaii*. For your view licenses the following reasoning:
>
> 1. My judgment that H is appropriate if and only if H is true (for recall that the FOE, in this case, supports what it entails).
>
> 2. If my judgment that H is inappropriate, no belief state is rational for me to adopt. (By DWJ-calibrationism)
>
> 3. If I don't have enough fuel to land in Hawaii (~H), no belief state is rational for me to adopt. (1, 2)
>
> 4. If it's rational for me to adopt a lower credence in H than my initial high credence in H, then I have enough fuel to make it to Hawaii. (3)
>
> 5. If it's rational for me to return home this must be because it's rational for me to adopt a lower credence in H than my initial high credence (from principles connecting epistemic and practical rationality).
>
> 6. If it's rational to return home, then I have enough fuel to make it to Hawaii. (4, 5)
>
> Now, perhaps I'm shallow, but I'm not really interested in being rational for the sake of being rational. What I really care about is having an awesome weekend. If rationality can only require me to act in a way that will make my weekend worse than it would be if I stick to my irrational choices, I'm afraid I'm not much interested in being rational!

I don't mean to endorse this reasoning. Perhaps we *should* be interested in rationality for its own sake. However, it seems like even an interest in the

quality of one's weekends should lead to an interest in forming beliefs rationally. DWJ-calibrationism makes it mysterious why this is so.

I have raised three concerns for the DWJ-calibrationist. First, the view has the consequence that, even in a case in which all of one's evidence seems to favor p, the total evidence doesn't support believing p. Second, the view is inconsistent with single premise closure. And third, the view makes it mysterious why an agent's practical interests should align with an interest in having rational credences. The DWJ-calibrationist might be willing to bite some bullets here. But given the odd consequences of the view, I think it is worthwhile to explore some alternative options.

Let's sum up. J-calibrationism is problematic because it makes rationality too cheap. It (and the PRJ-calibrationism variant) has the consequence that, in a wide range of cases, J-calibrationism is the *only* constraint on rational belief. WJ-calibrationism seems to avoid this unattractive consequence. However, WJ-calibrationism is committed to the existence of problematic epistemic dilemmas. In the next section I turn to the second horn of the dilemma—E-calibrationism.

4. Second Horn: E-Calibrationism

Recognizing the problems with J-calibrationism might make E-calibrationism seem like a natural solution. One shouldn't have credence r in *whatever it is that one judged*! Rather, one should have credence r in *whatever the first order evidence best supports*. On this view, regardless of how you judged, there *is* a credence that is rational to adopt (credence r in the proposition that the FOE best supports), so there are no problematic epistemic dilemmas. But E-calibrationism doesn't make rationality too cheap either, for it doesn't entail that one is rational in assigning credence r to one's judgment, regardless of what one judged.

Unfortunately, E-calibrationism doesn't possess quite the same charm as J-calibrationism. How, you might wonder, could it could possibly be rational to assign a 0.6 credence to ~H, when you judged that *H*? Ultimately, I think that it cannot be. But note that the E-calibrationist isn't committed to thinking that an *agent* who judges that H, and assigns a 0.6 credence to ~H is rational. The E-calibrationist can say that any agent who judges inappropriately that H will not be *doxastically* justified in assigning a 0.6 credence to ~H. Nonetheless, the E-calibrationist can maintain that she will have *propositional* justification for assigning a 0.6 credence to ~H.[21]

E-calibrationism faces two problems. First, it leads to some strange results, and second, it is theoretically unmotivated.

---

[21]   This is similar to a suggestion that Christensen makes in his (2011), p.4.

4.1. Some Strange Results

Let's return to ORACLE TESTIMONY. In this case, I learn that my initial judgments are *anti*-reliable. Whenever I get good meteorological evidence that it's going to rain tomorrow, I become so depressed that I end up inventing very silly arguments for the conclusion that it will be sunny. On the other hand, if I get strong meteorological evidence that it will be sunny, my fear of disappointment leads to my judging that it will be rainy.

Suppose that I judge that it will be sunny. Then, I can know that

(a) The meteorological evidence supports raininess.

I also know that

(b) My expected degree of reliability is very low (the likelihood of my judgments being correct will be close to zero, since my judgments are anti-reliable).

Since I know that E-calibrationism recommends assigning credence r to the proposition that my first order evidence best supports, I can conclude that

(c) E-calibrationism recommends assigning a very low credence to raininess. (From (a) and (b))

So...

(d) E-calibrationism will recommend that I assign a very high credence to sunniness. (From (c))

Thus, E-calibrationism recommends that I stick to my judgment, and believe that it will be sunny, while also believing that the meteorological evidence supports raininess! But if I believe that the meteorological evidence supports raininess, why would I believe that it will be sunny? Such a belief seems patently irrational.[22]

More general principles, in the spirit of E-calibrationism, exhibit similar problems. For example, suppose you have a 0.8 credence that your child will win the dance competition. You then find out that you have always been more confident than you should be in your children's talents. For ease, let's just stipulate that you find out that your credence has always been 0.2 too high. A principle in the E-calibrationist spirit would recommend that

---

[22] The E-calibrationist might contest that your judgment can change in response to the higher order evidence that you get. But this is impossible, since your judgment is *what you would be most confident in given the first order evidence alone*. It is important that judgments be understood this way for the calibrationists to get the desired verdicts in cases like HYPOXIA.

you reduce your confidence in this case, and it would license the following reasoning: "My first order evidence must support 0.6, since the credence I arrived at on the basis of the first order evidence was 0.8. The credence that is rational for me to adopt is 0.2 less than *whatever the first order evidence supports*. Since the first order evidence supports 0.6, I should reduce my confidence to 0.4!" But that surely is the wrong result. Once you've established that the first order evidence supports 0.6, it would be absurd for you to further reduce your confidence to 0.4.

4.2. Motivating Calibrationism

The second problem with E-calibrationism is that it is theoretically unmotivated. To begin, it will be helpful to think about how calibrationism, in general, might be motivated, by focusing on cases in which the E and J calibrationists agree. In HYPOXIA, for example, when the agent is considering whether H, and the agent is 50% reliable, E and J calibrationism both yield the verdict that the agent should have a 0.5 credence in H and a 0.5 credence in ~H. The question is, why 0.5? The views which I will be calling "anti-calibrationist" views are views according to which if, in fact, the first order evidence supports H, the agent should be more confident in H than in ~H.[23]

What motivates anti-calibrationist views is the following sort of thought: to determine what the agent's total evidence supports in HYPOXIA we need to consider both her first order and higher order evidence. Let's suppose that the first order evidence supports H. The higher order evidence about the agent's reliability may warrant some dampening in confidence but it doesn't, in and of itself, favor H over ~H or ~H over H. So there is an asymmetry in the total evidence. The total evidence favors H more than it favors ~H. What is interesting about the calibrationist's position is that this asymmetry in evidence isn't reflected in the credences recommended. Although the evidence asymmetrically favors H, the calibrationist recommends equal confidence in H and ~H. Why is this?

Here is Christensen's (2010a) explanation:

> "In accounting for the [higher order evidence] about the drug, I must in some sense, and to at least some extent, *put aside* or *bracket* my original reasoning for my answer. In a sense, I am barred from giving a certain part of my evidence its due. After all, if I could give all my evidence its due, it would be rational for me to be extremely confident of my answer, even knowing that I'd been drugged…yet it seems intuitively that it would be highly irrational for me to be confident in this case…Thus it seems to me that although I have conclusive evidence for the correctness of my answer,

I must (at least to some extent) bracket the reasons this evidence provides." (195-6)

Indeed, "bracketing" is what is needed to get the intuitive results in the cases Christensen describes. But bracketing is odd. That we need to suspend judgment even when we have conclusive evidence for the correctness of our judgment is something that needs to be explained. The problem with E-calibrationism is that it doesn't have the resources to provide such an explanation.

The claim that E-calibrationism is in special trouble when it comes to theoretically motivating calibrationism might come as a surprise. You might have thought that a nice feature of E-calibrationism is that, although it requires one to go through the bracketing rigmarole to come up with r, the first order evidence does get to play *some* role: it determines which proposition the credence, r, should be invested *in*. In fact, many criticisms of calibrationism along the lines of "you're ignoring your first order evidence!"[24] have been responded to by showing that the first order evidence *does* have a role to play, even for the calibrationist.[25] However, I will argue that this feature of E-calibrationism actually makes the view more, rather than less, difficult to motivate. There are two reasons for this. First, the best explanations for bracketing are not available to the E-calibrationist. And second, E-calibrationism makes it puzzling why one's expected degree of reliability is relevant at all.

(a) Unavailable Motivations

The Thermometer Motivation

One way to motivate calibrationism and explain the bracketing phenomenon is to appeal to the thought that we should treat ourselves in the same way that we treat other information processing devices like, for example, thermometers.[26] Suppose you have a thermometer that gives a reading of 53 degrees but you know that the thermometer is only 50% reliable. Surely, you should only be 50% confident that the temperature is 53 degrees.[27] Similarly, if you take your judgments to be only 50% reliable, you should only be 50% confident in your judgments. Now, it would

---

24 See especially Kelly (2010).

25 See Horowitz and Sliwa (ms.) and Christensen (2011).

26 See Enoch (2010) and White (2009) for discussion of the thermometer analogy.

27 Crucially, the reading of the thermometer needs to be the only evidence you have. This makes the analogy tricky since, plausibly, in higher order evidence cases, your judgment *isn't* all of the evidence you have, but let's just grant the calibrationist that this worry can be resolved.

clearly be absurd, in the thermometer case, to argue as follows: "Look, the temperature is 53 degrees [psst...it says so on the thermometer], and the thermometer *says* it's 53 degrees! So there's clearly more than a 50% chance that the thermometer is right on this occasion." Similarly, the calibrationist might say, it would be absurd to think: "Look, the answer to the math problem is 53 [psst...I figured this out on the basis of my reasoning], and my reasoning led me to the conclusion that the answer is 53! So there's clearly more than a 50% chance that I judged correctly on this occasion." The moral of the story: we can't take for granted the results of an instrument in determining how reliable the instrument is. This is why we need to bracket.

It's not just thermometers that we treat this way. There are other cognitive faculties whose outputs we treat very much like thermometers. Suppose that the eye doctor tells you that your color vision is only 50% reliable. You look at a wall and it looks red. How confident should you be that it is red? Surely only 50% confident. Once again, it would be absurd to reason as follows: "The wall is red, and I judge it to be red. I must have gotten lucky this time!" This is unacceptable. When figuring out how likely it is that your vision will yield correct results on a given occasion, you need to do so in a way that doesn't rely on your vision.

The thermometer and color vision examples are meant to illustrate that bracketing isn't odd at all. It's the sort of thing we do *all the time*, and calibrationism is motivated by the very same motivations that motivate our judgments in the thermometer case and the vision case.

Whatever the merits of this motivation, it is unavailable to the E-calibrationist. Here is the problem: with thermometers and with color vision, we match their expected degree of reliability *to their "judgment"*. If the thermometer *reads* 53 degrees, we ought to be 50% confident *in the reading of the thermometer*. If our color vision is 50% reliable, we should be 50% confident that the color that the wall *appears to us to be* is the color that it is. If we were to treat our rational faculties in the same way that we treat thermometers and visual faculties, then we ought to have credence r in *whatever the results of our reasoning are*. But this is *J*-calibrationism. E-calibrationists think that you should have credence r in the proposition that the first order evidence *actually* supports, even if that differs from your judgment.

Note that nobody thinks that we should be 50% confident in what the temperature it *actually is*, or 50% confident that the wall is the color it *actually is*. Now, there is no problem, per se, with treating rational processes differently from thermometers and visual processes. But the E-calibrationist cannot appeal to the idea that we should treat ourselves like thermometers if the view is that our credence in *what is in fact supported*

by the first order evidence should be r. That is a very unthermometer-like treatment.

The Guidance Giving Motivation

A different kind of motivation for calibrationism comes from the thought that the principles of rationality should provide agents with helpful guidance. One might think that a problem with anti-calibrationist views is that they fail to provide guidance in the relevant respect. For simplicity, let's consider an anti-calibrationist who thinks that one's credence in a proposition should equal the credence warranted by the agent's first order evidence (this is Kelly's (2005) "Right Reasons" view). "Look," a calibrationist might say. "It's all well and good to say that your credence should equal the credence supported by the first order evidence. In an ideal world, this would be just dandy. The problem is that we frequently get evidence that suggests that we're *not* in an ideal world and that we're not in a position to evaluate our first order evidence properly. The Right Reasons view fails to give us useful advice in precisely the circumstances in which it is meant to apply because it requires that we be in a position to evaluate our first order evidence. However, in the relevant circumstances, that is the very thing which we (believe) we *aren't* in a position to do."

Determining what conditions a principle must meet in order to count as "guidance giving" is notoriously tricky business and one might worry about the argument I just described (does it appeal to a problematic luminosity assumption?). But delving into those depths is a topic for another day. All I aim to show here is that, whatever merit these considerations have, they are not ones that the E-calibrationist can appeal to in motivating her position. This is because E-calibrationism doesn't fare much better than Right Reasons in the guidance giving domain. Indeed, the very same complaint made against Right Reasons can be made against E-calibrationism. "Look," someone might say. "It's all well and good to say that you should have credence r in the proposition *best supported by the first order evidence*. In an ideal world, this would be just dandy. The problem is that we frequently get evidence that suggests that we're *not* in an ideal world and that we're not in a position to determine which proposition is best supported by our first evidence. E-calibrationism fails to give us useful advice in precisely the circumstances in which it is meant to apply because it requires that we be in a position to properly evaluate our first order evidence. And, in the relevant circumstances, that is the very thing which we (believe) we *aren't* in a position to do." So if the complaint against Right Reasons is that its recommendation relies on our ability to determine what the first order evidence supports, even when that ability is believed to be impaired, the very same complaint can be made against E-calibrationism.

Brian Weatherson (ms.) has claimed that what is motivating calibrationists is the thought that "judgments screen evidence." What he means by this is that once one has arrived at one's judgment, the original evidence no longer bears on the rational credence to have in some proposition H. (In Bayesian terms Pr(H|J) = Pr(H|J&E) where J is a proposition describing what you judged and E is the first order evidence.) Whatever the merits of this motivation for calibrationism, it is not one that the E-calibrationist can make use of. This is because, for the E-calibrationist, judgments *don't* screen evidence. Although it's true that the first order evidence needs to be bracketed when coming up with *r*, it still has a very large role to play. (You might think that this, in itself, is an odd feature of E-calibrationism. If the first order evidence is available for use, why not use it to its full extent? Why can't it be relevant to determining the likelihood that one judged correctly in this instance? White (2009) writes: "But now if the evidence *e* is to play *some* role in determining my credence, why shouldn't it just do the whole job? And if the evidence has to divide up its belief-determining duties with my expected reliability, why is the evidence relegated to this task?" (240).)

(b) Why is r Relevant at All?

The last point I want to make about the motivational problem is that, on the E-calibrationist view, it's quite unclear why r should play any role whatsoever.[28] For what does r represent? r is the likelihood that p is true *given that you judge that p*. In other words, r *just is* the reliability of the agent's *judgment*. It is not how likely p is given that the first order evidence supports p. Given that r is just a number that represents judgment reliability, it would be odd for a principle to recommend that r be assigned, not to the judgment, but to the proposition that the first order evidence best supports.

5. A Possible Way Out

I don't think that there is an easy way out of the calibrationist's dilemma. Of course, one could bite various bullets. But since I am friendly to calibrationism, and I find the bullets difficult to bite, I would like to end by offering a very general strategy that might be helpful. Details of the view I will be suggesting require further development. The aim of this section is just to give a sketch.

The basic idea is that we should not think of calibrationism as a principle of epistemic rationality, where such principles are understood as principles

---

[28]   Thanks to Brian Cutter for this point.

about the relationship between bodies of evidence and belief states. Rather, we should think of calibrationism as a principle of *reasoning*. I am understanding principles of reasoning as principles about which transitions of thought, or reasoning processes, should occur in the process of deliberation. Examples of principles of reasoning might include things like: "if you believe p and if p then q, believe q," or "update your credences by conditionalization." Unlike principles of rationality, which take as input total bodies of evidence, and output a belief state, principles of reasoning can take as inputs beliefs, seemings, judgments, hunches, credences, and so forth and can output similar sorts of things.[29] If principles of reasoning are supposed to be helpful, and guide deliberation, J-calibrationism looks like a promising candidate. It tells us to take the output of whatever reasoning processes we just employed and then calibrate that judgment to our expected degree of reliability. Furthermore, if we think of calibrationism as a principle of reasoning there is no worry that J-calibrationism makes rational belief too cheap. On this reading, J-calibrationism says nothing about rational belief. J-calibrationism is about an entirely different subject matter: transitions of thought.

You might have the following worry: although we can distinguish the evaluation of belief states, from the evaluation of transitions of thought, or reasoning processes, what are we to do when the two theories yield conflicting recommendations? For example, if a calibrationist principle of reasoning recommends transitioning from my judgment to credence r in my judgment, but the theory of rationality recommends a high degree of confidence in what the first order evidence supports, what am I to do? A full discussion of the ways in which principles of reasoning and principles of rationality interact is beyond the scope of this paper. I will say only two things: First, on my view, the principles of *reasoning* are going to be the principles that play the role of guiding deliberation from the first personal perspective.[30] Evaluation of *belief states* will be of primary interest from a third personal perspective. Since these two theories play very distinct roles, the fact that they deliver conflicting recommendations will not pose a problem. Second, a substantive fact about the way that these principles interact is that if an agent knows what *belief state* is rational, a principle of reasoning will recommend that she transition to that belief state. So an agent will never find herself in a situation where she knows that it's rational for her to be in belief state 1, but the principles of reasoning recommend that she transition to belief state 2. If the agent knows that it's rational for her to be in belief

---

[29] See Harman (1986) and Broome (2013) for a thorough discussion of principles of reasoning.

[30] See Schoenfield (ms.) for a more detailed discussion of this point.

state 1, the principles of reasoning will recommend that she transition to that belief state.[31]

Many questions remain concerning how reasoning and rationality interact. The aim of this section was just to demonstrate that there is hope for a principle in the calibrationist spirit, as long as the principle is not understood as a principle that governs relations of evidential support between bodies of evidence and belief states.

6. Conclusion

Calibrationism is a principle that tells us to match our credences to our expected degree of reliability, r. The question then arises, credence r in *what*? In the proposition that we judged most likely to be true on the basis of the first order evidence alone (J-calibrationism), or in the proposition that is best supported by the first order evidence (E-calibrationism)? This poses a dilemma for calibrationists. J-calibrationism makes rationality too cheap and E-calibrationism is unmotivated. I suggested that we may be able to salvage calibrationism by thinking of it, not as a principle about relations of evidential support, but as a principle of reasoning. Recognizing that there is more to rational evaluation than evaluation of *belief states* opens the door to alternative ways of conceiving of calibrationist principles.

References

Broome, J. (2013). *Rationality through Reasoning*. Wiley-Blackwell.

Christensen, D. (2007). "Does Murphy's Law Apply in Epistemology? Self-Doubt and Rational Ideals." *Oxford Studies in Epistemology* 2: 3–31.

—— (2010a). "Higher Order Evidence." *Philosophy and Phenomenological Research* 81(1):185–215.

—— (2010b). "Rational Reflection." *Philosophical Perspectives* 24(1):121–140.

—— (2011). "Disagreement, Question Begging, and Epistemic Self Criticism." *Philosophers' Imprint* 11(6).

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

---

[31]   The view that I am sketching bears some similarity to a view that has recently been defended by Han van Wietmarschen (2013). van Wietmarschen argues that conciliatory views about peer disagreement shouldn't be thought of as principles that describe relations of evidential support. In fact, on his view, if you conciliate, your doxastic state will *not* be supported by the evidence. Rather, he says, conciliating is a necessary condition for your doxastic state being *well grounded*. (He suggests that we might understand well groundedness as something like *doxastic justification*.)

Elga, A. (2007). "Reflection and Disagreement." *Noûs* 41(3):478–502.

—— (2013). "The Puzzle of the Unmarked Clock and the New Rational Reflection Principle." *Philosophical Studies* 164(1):127–139.

—— (ms.). "Lucky to be Rational."

Enoch, D. (2010). "Not Just a Truthometer: Taking Oneself Seriously (but Not Too Seriously) in Cases of Peer Disagreement." *Mind* 119(476):953–997.

Feldman, R. (2005). "Respecting the Evidence." *Philosophical Perspectives* 19:95–119.

—— (2006). "Epistemological Puzzles about Disagreement." In Hetherington S. (ed.) *Epistemology Futures*. Oxford University Press.

Horowitz, S. and Sliwa, P. (ms.). "Respecting *All* The Evidence."

Kelly, T. (2005). "The Epistemic Significance of Disagreement." In John Hawthorne and Tamar Gendler (eds.), *Oxford Studies in Epistemology, Volume 1*. Oxford University Press.

—— (2010). "Peer Disagreement and Higher Order Evidence." In Alvin Goldman and Dennis Whitcomb (eds.) *Social Epistemology: Essential Readings*. Oxford University Press. 183–217.

Lasonen-Aarnio, M. (2010). "Unreasonable Knowledge." *Philosophical Perspectives* 24(1):1–21.

—— (2014). "The Limits of Defeat." *Philosophy and Phenomenological Research* 88(2):314–345.

Schoenfield, M. (ms.). "Bridging Rationality and Accuracy."

van Wietmarschen, H. (2013). "Peer Disagreement, Evidence, and Well-Groundedness." *Philosophical Review* 122(3):395–425.

Weatherson, B. (ms.). "Do Judgments Screen Evidence?"

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

—— (2009). "On Treating Oneself and Others as a Thermometer." *Episteme* 6(3):233–250.

Wright, C. (2004). "Warrant for Nothing (and Foundations for Free)?" *Proceedings of the Aristotelian Society Supplementary Volume* 78:167–212.