“Planning for Failure”

Dissertation Abstracts for Robert Steel

Committee: Karl Shafer (chair), Kieran Setiya, James Shaw, Mike Caie, Anil Gupta

Short Abstract:

> Sometimes I not only judge that P, but I also have some further “higher-order” evidence about my reliability when it comes to this sort of judgment. Perhaps I make an arithmetical judgment, but at the same time I have evidence that I am drunk and so arithmetically unreliable. I argue that the rational response in such a situation is to calibrate one’s confidence in P to the level suggested by one’s higher-order evidence—in this particular case, however confident I would be in P just given the description “I judged that P, drunkenly.” As a special case, this “calibrationist” doctrine entails a conciliatory view of peer disagreement. Some have found conciliatory views to be disturbing, taking them to collapse into psychologism and, at the limit, skepticism. I develop a conciliatory view which need be neither of these things, which is fortunate, since I also argue it is correct.

Long Abstract:

Later tonight I am going out to dinner. I think: “at some point, I will have to figure out my share of the check. If I drink, then my chances of botching the division go up substantially. Given what I know of my own reliability in these matters, I anticipate two drinks will reduce my chances of getting it right on the first try to a mere 50-50.” Fast forward to dinner, where I have two drinks and then I go on to try to divide the check. How confident should I be that I’ve gotten it right on the first try? I say: I should take it to be 50-50. I should match my current confidence to my earlier estimate.

My view is thus “calibrationist” in character: it requires calibrating one’s judgements to one’s antecedent estimate of their reliability. This advice furnishes us with a picture of both how to understand “higher-order evidence” and how to respond to it. What is higher-order evidence? Higher-order evidence is evidence on the basis of which one can estimate the reliability of one’s judgements. How ought one respond to higher-order evidence? One ought match one’s confidences to the estimates formed on its basis.

I hold that following this advice is *all* one needs to do to acquire an *ultima facie* rational belief. An immediate consequence: even if one’s initial judgment failed to match the first-order evidence, the belief arrived at by correctly calibrating it ends up rational all the same. As such, on my view there is a sense in which the weight of the first-order evidence turns out to be irrelevant to what one should believe at the end of the day. Namely, we can determine what one should believe just by citing one’s initial judgment and one’s expectation of its reliability, and in so doing we need never say a word about what the first-order evidence actually supported.

Notice, after all, that this is just what I did with the opening check-splitting case. The story I told contains no information about what the correct calculations support; it contains no first-order considerations at all. Yet, nonetheless, I was ready to pronounce on what my ultimate confidence should be. I could do this this because on my view we don't need to know what the first-order considerations support to know what my ultimate confidence should be.

We can represent this picture somewhat sloganistically with the claim that *rational belief is directly sensitive only to higher-order evidence (and not first-order evidence)*. So sloganized, it's easy to see the natural alternatives. So, for instance, there's the mirror image: "right reason" views hold that rational belief is directly sensitive only to first-order evidence (and not higher-order evidence). And then there's the pluralistic choice: "interactionist" views hold that rational belief is directly sensitive to *both* first- *and* higher-order evidence. Both these alternative styles of view are well-represented in the literature, and so my first project is cataloguing their weaknesses.

Against right reason views, I argue they are extensionally inadequate. They yield remarkably incautious advice, most especially when we consider contexts where our beliefs are of practical significance. Defenders have tried a variety of strategies to try to explain away such counterexamples, but I argue all are hopeless. Against interactionist views, I argue that they are theoretically undermotivated. I show that correctly following interactionist views sometimes requires you to adopt an attitude you take to be irrational; I diagnose this as revealing that interactionist views cannot specify a coherent perspective from which to pursue their aims.

Having done some work to clear the way, I then turn to arguing more directly for calibration. In my central positive argument, I show that calibration has the following lovely feature: trying to follow it maximizes your expected epistemic returns, and, as such, calibration is the best view *to be committed to.*

It may seem suspect, though, to take the desirability of being committed to calibration to be a reason to think that it's actually correct; doesn't this confuse the attitude that it's rational *to have* with the attitude that it's rational *to want to have*? In order to dispel such worries, I turn to meta-epistemology. I start by examining the nature of the epistemic 'should'; what is the point of making claims about what a person should believe, and what do such claims do for us? I develop a view—inspired by recent expressivist work in meta-ethics—where the function of such beliefs includes the regulation one's future epistemic behavior, and I show how plausible versions of such a theory ground the requirement that the correct attitudes maximize expected performance. So, I hold, meta-epistemic reflection shows that in this case that the desirability of being committed to calibration *is* actually a good reason to think it correct. In the end, then, what we get is not just an extended defense of my favored view on higher-order evidence, but also a novel exploration of the meta-epistemic conditions that make it so.