

PHIL 640: Open-Mindedness

Brian Weatherson

November 29, 2021

I'm not going to prepare line by line notes on the papers for this week, instead the point of these notes will be to go over some things that I think are useful background for thinking about what's at issue in the papers, and generally about open-mindedness and closed-mindedness.

Logic

I want to start, perhaps a bit self-indulgently, with a bit of autobiographical background. As I've probably mentioned before, I did my studies (both undergrad and grad) in Australia in the 1990s. My thesis was on non-standard theories of uncertainty and decision. Or, at least, it was on theories that were then non-standard; one of them (imprecise probability) has since become pretty mainstream. And one of the big rolling debates in the background was over which (if any) is the correct logic. These two things overlapped a bit; the most technically innovative part of my thesis was exploring what probability theory looked like if you started with intuitionist logic rather than classical logic.

Now debates about logic, and to some extent debates about uncertainty, are distinctive in two ways that are relevant to this week's subject.

1. Some of the things we take for granted in other parts of philosophy are very much not taken for granted in logic. You can't simply assume that evidence against p is the same as evidence for $\neg p$. That doesn't work even in logics that are very close to classical logic. One big debate in Australia at the time was whether, as Graham Priest argued, there are true contradictions. And then you really can't say that evidence for $\neg p$ is evidence against p . Some (philosophical) arguments may provide evidence for both.
2. It really isn't clear at the start of play that there will be arguments for one or other side that are remotely convincing to the other. There are definitely things you can say back to Priest. You can argue (and many did) that his theory doesn't have the virtues that he claims for it, and in particular that it doesn't provide as unified an account of the paradoxes as he says. You can argue that there are alternative classical explanations of the phenomena that he wants to explain with true contradictions. But you really don't know you're going to succeed. Sometimes it seems like the best

thing to say is that if Priest's theory is true, then a contradiction follows, so Priest's theory is false. That's a rule of inference we literally teach to intro logic students, so we should use it here. But it's dogmatic, closed-minded, or whatever adjective you'd like to use. David Lewis's "Logic for Equivocators" is an important paper in this context. Apart from being slightly more rude than Lewis usually was (at least to his friends), it both gave substantive arguments against Priest (launching what became known as the fragmentation strategy), but also gave the straight-up denial that we should take this view seriously. The view implies contradictions. That's really bad!

The big picture here is that if you're coming from a background where logic itself is up for debate, you end up with a slightly different take on what are the permissible moves in a philosophical game. The idea that to have a justified belief that p you need to have arguments that would be even vaguely convincing to the other side seems ludicrous. You'd have to say that no intro student knows what the actually right answer is to any bit of logic homework, since none of them know how to reply to Priest. (Of course some of those students know what answers will get marked as correct, but they don't know whether the professors are grading the papers correctly.)

And it's really tempting to be a bit closed-minded here. You don't want to have to interrupt an argument about ethics every time you reason by disjunctive syllogism, or double negation elimination, just because there might be a non-classical logician lurking in the corridor. (Again, this is a more vivid concern when in fact in almost every philosophy department in the country there was in fact a non-classical logician somewhere in the corridor.) You just want to take logic for granted, and get on with doing philosophy. And while it wasn't quite as pressing, for a lot of purposes you also wanted to pretend that people like me - who quibble about the foundations of decision theory - aren't around either.

The upshot of this is when I see these general principles being states about dogmatism, or closed-mindedness, my first reaction is always to ask about how they would apply when debates about logic are going on around you. And it seems that for a lot of purposes it would be a very good thing to be somewhat closed-minded. You can't get anything done on non-foundational questions without being a little closed-minded.

But even within this, there are worries. It's fine to not worry about whether classical logic is correct when doing first-order ethics, let's say. But even if you're happy in principle with allowing this kind of closed-mindedness, what are its limits. Is it ok to assume orthodox Bayesian decision theory when doing ethics, for example, or should you worry a bit about non-standard decision theories. It is ok to assume classical logic when you are reasoning about how comparatives like 'taller' work? What about when you are reasoning about 'tall' instead of 'taller'? Weirdly, the standard practice is to take classical logic for granted when reasoning about comparatives, but not at all to take it for granted when reasoning about vague terms like 'tall'. We could probably do with tidying up our practices here.

One kind of foundational question I wasn't worried about in the 90s, but probably should have been, concerns reasons scepticism. In *The Philosophy of Philosophy*, one of Timothy

Williamson's arguments against certain kind of anti-dogmatist principles is that they would mean that we couldn't justifiably believe that the reasons sceptic is mistaken. But that means we can't justifiably believe that anything is a reason for anything. That doesn't quite mean that we get global scepticism - it might be possible to properly believe something for a reason without believing that there are such things as reasons - but it's a really bad result.

Part of the point of writing all this out is to kvetch about using things like Holocaust deniers as examples of people who we can properly be closed-minded towards. Deviant logicians, and deviant decision theorists, are much closer to home, and for many of us, it would make a much bigger difference to our everyday practices if we took them seriously.

Disagreement

I literally wrote a book that was about, among other things, the contemporary debate about disagreement. I fear that if I get started on this I'll spend all this week, all next week, maybe all next semester, starting on it. But for now I just wanted to stress one point that Fantl makes, and which I think gets really left out of the conversations about disagreement.

Consider a situation in which two people disagree. Neither of them have reason antecedently to think that they would be more likely than the other to get the correct answer. In that respect they are symmetric. But one of them actually has a reasonable belief and the other does not. In that respect they are asymmetric. The big dispute in work on disagreement in the 2000s-2010s was whether you took the symmetry or the asymmetry to be more important.

I'm on the team that says the asymmetry is more important. Or, at the very least, that the asymmetry matters.

The key point is that evidence matters. Or, if you don't want to assume evidentialism, reasons matter. If A has good reasons for their belief, and B does not have good reasons for their conflicting belief, then after the disagreement, A still has those initial good reasons. They may also have some reasons to move their belief somewhat in the direction of B, but they have the initial reasons. I've never seen a good argument for why those reasons shouldn't still matter. And that line of thinking supports being a teensy bit closed-minded. Or, if you'd prefer, it means that the virtue of open-mindedness can't be the kind of fickleness that we associate with conciliationism.

Intuitions, Evidence, and Models

I'm very sympathetic to Battaly's conclusion that some closed-mindedness can be useful to an inquirer, and to a community. But of the three following ways to argue for that, I suspect the first is the least convincing.

1. Appeals to **intuitions** about cases where it seems that closed-mindedness could be beneficial.
2. Appeals to **evidence** either about historical situations or examined situations, where it was good someone was closed-minded.
3. Appeals to **formal models** that show how closed-mindedness can be a good thing.

All three probably have a part. But if we have 2 and 3, then I'm not sure how much we should be relying on 1.

It would be nice to have some examples from the history of science that fall under 2. My impression is that the history of mRNA vaccines, for example, falls into category 2. A handful of people didn't agree with the conventional wisdom that mRNA vaccines weren't going to be viable to use in real-world settings. And here we are. (There is a short version of the history of these vaccines in Nature.) But I'm sure there are plenty of other cases. The scientist who bravely/foolishly stuck to his or her guns while everyone else believed that the project wouldn't work is almost a cliché. They should be the ones we cite as examples here.

There are some interesting recent attempts to formally model ways in which it might be good for a scientific community to include some people who are closed-minded. One class of models takes the members of the community to be procedurally rational in the sense that every scientist takes into account all the evidence they receive, they update by conditionalisation as they should, and they perform the experiments that will have the highest expected return given their knowledge. What the models do is vary not the rationality of the agents, but their information. But the same kind of model can be used for closed-minded agents. Instead of having uninformed ones, we can have ones who are informed, but who ignore that information. That is, we can have closed-minded agents. And sometimes, that leads to better outcomes. If you want to see some early modeling of this, start with these two papers by Kevin Zollman.

- Kevin Zollman (2007) "The Communication Structure of Epistemic Communities," *Philosophy of Science* 74 (5): 574–87.
- Kevin Zollman (2010) "The Epistemic Benefit of Transient Diversity," *Erkenntnis* 72 (1): 17–35.

But note that these models are not uncontroversial. It is arguable that the results where closed-mindedness is good are only a very small part of logical space. A version of this critique is in this paper.

- Sarita Rosenstock, Justin Bruner, and Cailin O'Connor (2017) "In Epistemic Networks, Is Less Really More," *Philosophy of Science* 84 (2): 234–252.

Closer to home, recent UM graduate Sara Aronowitz has argued that there can be long-run epistemic benefits to having beliefs that are over-confident.

- Sara Aronowitz (2021) "Exploring by Believing," *Philosophical Review* 130 (3):339–383.

The argument is complicated, relying both on claims about what kinds of things are only possible for believing agents, and about rational behavior in multi-armed bandit problems. The latter is really fascinating, in part because it messes up so many intuitions we have about decision theory.

Here's the intuition behind the theory. (There is a lot more algebra in the real version.) Imagine you move to a new town, and you want to find out the best pizza there, because you eat a lot of pizza. More precisely, you want to find the best pizza-sub-your-tastes. You find three (affordable) pizza places nearby, and you try the three of them. Or maybe you try them each twice, if you really want to be sure. And there is one of them you like best. What do you do next?

Well, you certainly should have the one you like best *most of the time*. But should you have it *all of the time*? Probably not. If one of the other pizza places got better, you'll never find out that way. You should occasionally see if the others get better. And you should do that even if (a) the one you use hasn't gotten worse on average, and (b) there is no evidence that the other ones have gotten better. It does matter that there is evidence that pizza places change in quality over time, but it doesn't matter whether you expect that change to be on average for better or worse.

Now in practice what you do is when you get a bad pizza from your favorite place, that prompts you to try one of the others. And as I understand it, that's actually a reasonable rule. Even if you know that your favorite place varies from time to time, and one bad pizza is probably just bad luck, not a sign in falling quality, it is (I gather - I'm not 100% sure of this) a reasonable strategy to use things like a bad pizza as a reason to check in on how the others are doing.

That's to say, sometimes there is a long-run advantage in doing something that does not, in the short-run, maximise expected utility. If you see how the other places are doing, you'll probably be reminded why you didn't eat there. But if you never try, you won't find out. So you should try.

What's nice is (a) there is a mathematical model that more or less backs up these intuitive behaviors, (b) it is a really hard mathematical problem to figure out how to optimise within this model - I gather it's something we still do more or less by trial and error, not by proof, and (c) just like you can show that it's best for people to not maximise expected utility on each occasion, you can argue that it's best for people to not perfectly conform their credences to the evidence. That's to say, it's useful to be closed-minded. Or, at least, that's what the math says given some quite substantial assumptions about the long-run effects of certain beliefs.

Now one thing you might note about Battaly's paper is that precisely none of this stuff turns up in it. In fact, there is remarkably little engagement with folks outside of the virtue epistemology community. It's a bit disconcerting that a literature about open-mindedness and closed-mindedness, and relatedly about epistemic bubbles is, well, a little closed-off.

There are all sorts of connections here to history of science, formal philosophy of science, philosophy of mind, decision theory, and so on. But very little of that turns up here, and that's a somewhat worrying feature of the virtue epistemology literature. One thing I hope you'll get from being at Michigan is the disposition to draw on a bunch of distinct fields, both inside and outside philosophy, when writing, so you don't end up with a relatively narrow range of stuff you engage with.

Reference Class

So I want to end this week's notes, and as it turns out all the notes, with an old philosophical problem. It feels at first blush like the kind of technical problem that you can be confident some clever graduate student will solve one day. But I think we see in Battaly's paper that it's got a bit more philosophical, and actually practical, relevance.

The original problem was an issue for **reliabilism**. Alvin Goldman (and others, but Goldman is the figure you most need to know in this context) argued that a belief was justified iff it was produced by a reliable process. This was a very clever idea that sidestepped a host of problems with traditional theories of justification.

For one thing, it is a theory where it is clear why justification matters without making justification require truth. And that combination seems both important, epistemic justification should have those features, but also somewhat hard to pull off. After all, if something doesn't imply truth, we need some story about why it is interesting to truth-seekers. The standard internalist response, that justification implies *probable* truth just shifts the problem back a step. But if justification means being *usually* true, then we can see why it would matter.

For another, it gives us a way around some of the classical problems about scepticism. Yes, it's true that if you were a brain in a vat, then the inference from *this look like this* to *things are like this* would fail. But as a matter of fact you're not a brain in a vat, and neither is anyone that you know. Perception is, at least in a lot of circumstances, reliable. So we don't lose justification by trusting it.

There are other advantages too, but let's move on to the disadvantages. If I fall for a visual illusion, and come to believe *p* when *p* is false, what should we say about the justification of my belief. Well, on the one hand, it was a belief produced by trusting visual perception. And visual perception is pretty reliable. On the other hand, it was a belief formed by trusting appearances when looking at an illusion. And, as the name suggests, that's not a very reliable method. So is the belief reliable or unreliable; justified or unjustified.

The problem arises because of a mismatch between what we want and what the theory gives us. Reliability is a property of a class of processes. What we want is the justification of the belief produced at the end of a token process.¹ To talk about the reliability of this

¹Actually even this is conceding too much to the reliabilist. A belief is the end of any number of overlapping processes, some of them which took microseconds, and some of which took decades. Which is the process the

belief, we need to associate a belief with a particular class of beliefs, and then ask about the reliability of that class. But a belief is a member of any number of classes of beliefs. (I guess 2^{n-1} classes, where n is the number of beliefs in the world.) Which is the special class that matters for reliability.

Some of the classes seem obviously inappropriate. Consider my belief that my cat is nearby. (She is; I just looked.) That's a member of a class of beliefs that consists of it plus thousands of token beliefs that covid vaccines don't work. That class consists mostly of false beliefs, so the beliefs in that class are, qua members of that class, unreliable. But my belief that the cat is nearby is justified. Is this a problem for reliabilism? No - that isn't the relevant class. Or - better - it isn't a relevant class.

But what are the relevant classes? Here's a way to make that question more concrete. Imagine that closed minded Claude doesn't listen to counter-evidence to his belief that p . But as a matter of fact, Claude knows that p . Which type of activity does this particular ignoring fall under?

1. Ignoring counter-evidence to one of his beliefs.
2. Ignoring counter-evidence to one of his true beliefs.
3. Ignoring counter-evidence to something that he knows.

Class 1 is presumably an unreliable kind of activity. But class 2 is definitely a good practice. You'll get more reliable that way. Still, it seems wrong to call it a class. We don't have the ability to just ignore counter-evidence to our true beliefs.

What should we say about class 3? Obviously it is also a good idea to ignore counter-evidence to your knowledge. It is misleading, and you should ignore misleading evidence. But can we ignore only counter-evidence to what we know? That's a tricky question.

On the one hand, we don't always know what we know. So if we try to implement the rule of ignoring counter-evidence to our knowledge, we are almost sure to misapply it at some time. On the other hand ... there is nothing interesting that we can identify with 100% accuracy. And rule will have conditions that we don't always identify. So any rule will have this problem. Is a rule that is sensitive to one's knowledge any worse off in this respect? I'm not sure - this feels like a good open question to end on.

belief is at the end of is a question the reliabilists have never satisfactorily answered, though they have helped themselves to the assumption that there is an answer.