Reply to Blackson

Brian Weatherson

2016

Thomas Blackson argues that interest-relative epistemologies cannot explain the irrationality of certain choices when the agent has three possible options. I argue that his examples only refute a subclass of interest-relative theories. In particular, they are good objections to theories that say that what an agent knows depends on the stakes involved in the gambles that she faces. But they are not good objections to theories that say that what an agent knows depends on the odds involved in the gambles that she faces. Indeed, the latter class of theories does a better job than interest-invariant epistemologies of explaining the phenomena he describes.

It isn't true that all interest-invariant epistemologies are alike, but it is certainly true that every interest-relative theory is interest-relative in its own idiosyncratic way. In fact, there are at least four dimensions along which a theory can be interest-relative.

I used to think (Weatherson 2005) that interest-relativity in knowledge was to be explained by interest-relativity in belief, but I came to think that's not true (Weatherson 2012). Some prominent defenders of interest-relativity in epistemology focus on practical interests – it's even there in the title of the book by Jason Stanley (2005) – but others of us think that theoretical interests matter too. At times Stanley writes as if interest-relativity means that there is an extra clause in the theory of knowledge for interests, but one need not think that. It could, for example, be that there is an interest-sensitive domain restriction on a quantifier in one of the clauses.

But for present purposes, the key divide among interest-relative epistemologists is between those who think that stakes are relevant, and those who think odds are relevant. I think, following Mark Schroeder (2012), that it is odds that matter. The key examples here are ones where there is little cost to gambling and getting it wrong, but even less to gain by gambling and getting it right.

So imagine that Ankita is walking to a restaurant she hasn't been to for a few months. She is stopped at the lights, reading baseball scores on her phone. She is almost, but not completely, certain that she should turn left at the next block, which indeed she should. If she had been wrong, she would have gone two blocks out of her way. She could avoid this risk by flipping from her baseball app to the map on her phone and checking the address, all of which she could do before the lights change. I say that in this circumstance, she doesn't know where the restaurant is. She should look up where it is, that's what maximises expected utility, but she needn't look up restaurants she knows the location of. So she doesn't know whether the restaurant is to her left or her

right. Ankita's case is not a high-stakes one. Even on a cold Michigan Fall evening, the downside of walking two extra blocks is not that high. But unless she really really cares about those baseball scores she's browsing through, deciding not to flip over to the map is a gamble on the correctness of her plans at incredibly long odds. That's not because the stakes are high, but because the gain from gambling is low.

This is a genuine case of interest-relativity though. The argument I just gave wouldn't go through if Ankita would have no disutility whatsoever from walking two blocks out of her way. (Maybe the exercise gain would completely outweigh the frustration.) In that case, perhaps she does know. But if the two block walk has many times the disutility of no longer browsing baseball scores, as it would in most realistic cases, her interests defeat her knowledge of the restaurant location.

The same thing I think is going on in the example Thomas Blackson (2016) gives. The agent has a three-way choice, between taking drug A, taking drug B and doing nothing. The upsides of both drugs are the same; they alleviate a minor medical condition. The downsides are the same too; they lead to death in rare cases. But this is much rarer still in the case of B than A. So, says Blackson, the person should take drug B. And I agree. But, says Blackson, this is a problem, because I'm committed to the following argument.

1. The agent knows that drug A and drug B won't kill them.
2. Any option an agent knows not to obtain can be left off a decision table.
3. So, from 1 and 2, the decision table the agent faces has only the upsides, and not the downsides.
4. So, from 3, the decision tables for the two drugs are the same.
5. So, from 4, the agent can be indifferent between the drugs.

Since 5 is false, one of the premises must be false. Blackson says that the false premise is 2. I say that it is 1; the patient doesn't know drug A won't kill them. Blackson anticipates this, and says that the move won't work because of a related case. If drug B didn't exist, the interest-relative theorist might say that the agent has enough evidence to know the drug won't kill them. Maybe that's true, but it isn't clear why it is relevant. The existence of drug B changes the gamble involved in taking drug A. It must, since taking drug A is irrational in the original case, but rational in the version where drug B doesn't exist. So if what the agent knows is sensitive to what gambles they face, then it isn't surprising that the presence of extra options changes what the agent knows.

And the interest-relative theory has a nice explanation of one feature of Blackson's case. Imagine the agent learns that drug A is only ever fatal with people with blood type A2B negative, and that's not their blood type. There's no similarly known marking for when drug B is fatal. Now the agent can say, "Now that I know drug A won't kill me, I'm going to take it, not drug B." That seems like just the right thing to say, but on Blackson's telling of the case, they can't say it. After all, they knew all along that drug A wouldn't kill them. Instead he has to say that the agent shouldn't take drug A before they learn this, because of the risk it would kill them, even though the agent knows the drug won't kill them. That doesn't sound at all right. Much better, I think, to say they

shouldn't take drug A before learning who it endangers, because it exposes them to a needless risk of death, but once they know it won't kill them, it is good to take it.

And that's the general strategy for defending interest-relative treatments of cases like Blackson's, and the others he describes. It's a strategy that I don't think is threatened by any example to date. There's some action that all parties agree would be irrational, or at least not rationally mandatory. But there's some evidence we can imagine the agent getting that would improve the rational status of the action. (E.g., change it from being rationally impermissible to rationally permissible, or from not rationally required to rationally required.) If we asked the rational agent why they changed their plans after getting that evidence, it seems to make sense for them to say that they now know the action would have a good outcome. That is, it makes sense for them to cite their knowledge as a reason for doing something different after the evidence-gathering event. And only the interest-relative epistemologist can explain why that is a sensible answer for them to give.

References

Blackson, Thomas. 2016. "Against Weatherson on How to Frame a Decision Problem." *Journal of Philosophical Research* 41: 69–72. doi: 10.5840/jpr201662868.

Schroeder, Mark. 2012. "Stakes, Withholding and Pragmatic Encroachment on Knowledge." *Philosophical Studies* 160 (2): 265–85. doi: 10.1007/s11098-011-9718-1.

Stanley, Jason. 2005. *Knowledge and Practical Interests*. Oxford University Press.

Weatherson, Brian. 2005. "Can We Do Without Pragmatic Encroachment?" *Philosophical Perspectives* 19 (1): 417–43. doi: 10.1111/j.1520-8583.2005.00068.x.

———. 2012. "Knowledge, Bets and Interests." In *Knowledge Ascriptions*, edited by Jessica Brown and Mikkel Gerken, 75–103. Oxford: Oxford University Press.