HOW DO INTERESTS MATTER?

Jane Friedman, New York University [*] [†]

In *Knowledge: A Human Interest Story*, Brian Weatherson fills in and refines a story about the interest-relativity of knowledge that he has been building for well over a decade. Someone who thinks knowledge is interest-relative thinks that some of the factors that determine whether S's belief p is knowledge are S's interests. Traditionally, these interests have been practical: the phenomenon according to which interests play a role in determining what's known has been dubbed 'pragmatic encroachment'. In *Knowledge*, Weatherson extends the view in a couple of interesting ways: First, it's not just knowledge that is interest-relative, but belief, evidence, and more. And second, it's not just practical interests that matter.

In this brief discussion, I want to stay focused on knowledge, but raise some questions about exactly how interests are meant to matter for Weatherson, both given his discussion of theoretical interests, and given some of his remarks about the interaction between knowledge and inquiry. These discussions add new layers of richness and complexity to Weatherson's view, but they also serve to make some of the mechanics of his 'Interest-Relative Theory' (IRT) harder to get a grip on. So these comments are a request for a bit more detail.

There is a standard sort of case that makes plausible the idea that what one knows at a time is in part a function of one's practical interests at that time. This isn't quite the case that Weatherson focuses on, but I think it's still helpful to have on the table. First, we imagine a subject (call her Lois) on the way home from work, deciding whether or not to stop at the bank to deposit a cheque. The bank is busy, and the queue looks long. It's Friday, and so Lois wonders whether the bank will be open tomorrow. She remembers that she went there on Saturday a few weeks ago and thinks about how banks are usually open on the weekends these days. Lois decides to come back tomorrow to deposit the cheque. Anyway, if the bank is closed, it's not a big deal; she can always deposit it on Monday. We tend to think that Lois knows the bank will be open tomorrow, and that her decision to drive home and go to the bank tomorrow is a rational one.

But then we imagine a second subject, Hiram. Hiram is facing the same decision as Lois, and almost all other features of his case are the same as hers

[*] jane.friedman@nyu.edu

[†] DRAFT: 2026-02-24. This paper is part of a symposium on Brian Weatherson's book, *Knowledge: A Human Interest Story*, to appear in *Philosophical Studies*. Comments welcome! Please ask before citing.

(it's Friday, same evidence about the bank, etc.). The main difference between Hiram's situation and Lois' is that it is a big deal if Hiram doesn't get the cheque in the bank before the bank opens on Monday morning. If he doesn't deposit the cheque before that, he'll default on his mortgage and will lose his house. In this case, we tend to think that Hiram — despite having the very same evidence as Lois — does not know that the bank will be open tomorrow and that it would not be rational for him to forgo stopping at the bank now.

The conclusion many have drawn from cases like these is that one's practical interests can make a difference to whether or not one knows something. But which interests mattered in the story I just told? And how did those interests matter? Neither question has an obvious answer.

If what makes it that Hiram doesn't know while Lois does is something about their interests, then Lois and Hiram have to have different interests (of the relevant sort). But there are many potentially relevant interests that Lois and Hiram share: they are both interested in the question of whether the bank will be open tomorrow (and are both actively investigating that question). And we can imagine Lois and Hiram as having many of the same interests in a more general sense — they can have many of the same preferences, goals, desires, etc.: neither wants to default on their mortgage, neither wants to waste time in a queue, and so on. Lois and Hiram are in different practical situations though: Hiram has to deposit a cheque or lose his house, and that is not the situation Lois is in. So while they both have an interest in not losing their respective houses, only one of them is facing a decision that could have that result. It's this combination of interests and potential practical consequences that seems to wreck Hiram's knowledge.

One helpful way to think about the relevant difference between the two cases is in decision-theoretic terms. Lois and Hiram face a decision: whether to stop at the bank on the way home. For Lois, the expected utility of stopping at the bank is lower than that of not stopping — stopping means queuing, but not stopping isn't expectedly worse than that. But for Hiram, the expected utility of stopping is much higher than that of not stopping, given how catastrophically bad it would be if he didn't stop and the bank was closed for the weekend. This is part of what makes it rational for Lois to skip the bank on Friday but irrational for Hiram to do the same. Even though Lois isn't absolutely sure that the bank will be open tomorrow, she can act as if it is; Hiram cannot do the same. And this is a good part of the explanation of why we think Hiram doesn't know. If he knew the bank was open tomorrow, he could act on that information and skip the queue today.

Weatherson's main example in *Knowledge* is different from the classic bank case in a few ways. They are worth going through since they help to clear away some possible confusions about exactly how the relevant sorts of interests do

their work. The central case is one Weatherson calls the 'Red-Blue game' (pp.25-6). Here are the rules: Two sentences are written on the board, one in red, one in blue. Then, a player has to pick one of the sentences and say whether it is true or false. If they are right, they get $50, and if they are wrong, they get nothing.

Anisa is our player. Last night she read some medieval history, and read that the Battle of Agincourt took place on October 25, 1415.[1] Today, these two sentences are written on the board, the first in red, the second in blue.

- Two plus two equals four (r)

- The Battle of Agincourt took place in 1415 (b)

Anisa decides to play 'Blue-True': she decides to pick the blue sentence and declare it true (after all, she read the reliable history book last night, which told her that the sentence was true). Weatherson thinks this is an irrational play on Anisa's part; she should have played Red-True.

Weatherson doesn't offer an intuitive verdict about whether Anisa knows b when she plays the game. He says he has no clear intuition either way. The intuition he does have is that Anisa's decision to play Blue-True is irrational. He then goes on to argue that the best explanation of Anisa's irrationality is that she does not know b. When she read the book last night, she did know b, but once she starts playing the Red-Blue game, she loses this knowledge.

Anisa's case is different from the classic bank case in a couple of key ways. First, Weatherson claims that the only intuitive verdict is about what it's rational for Anisa to do. That she also doesn't know b is an inference to the best explanation. Second, Anisa's case is not meant to be a 'high stakes' case — the potential gains are not large, and in the worst case, she'll end up no worse off than she started. Both of these strike me as important insights. By removing the potential noise of stakes, and being clear about what lesson we should draw about the case, Weatherson makes the mechanism of knowledge loss by way of practical interests fairly clear.

He draws this all out nicely as follows (pp.46-7). Take a case in which a subject is facing a practical decision. There are two courses of action, one safe (go to the bank now/play Red-True), one more risky (drive home and go to the bank tomorrow/play Blue-True). The safe option has a guaranteed return of S. The risky option's return varies depending on how the world turns out: if things turn out well (bank is open tomorrow/b is true), the subject does at least as well: $S + G$. But if they turn out badly (bank is closed tomorrow/b is false), they do worse: $S - L$. So whether the risky choice leaves the subject in

[1] This is true, and she has no reason to doubt it; she knows her history book is reliable; etc.

their good case or bad case, depends upon whether some proposition is true (bank is open/b).

What is generally true in the encroachment cases at issue is that subjects should maximize expected utility. And Weatherson says that the risky choice will maximize EU only if the following equation is true:

$$Pr(p)/(1 - Pr(p)) > L/G \qquad (\star)$$

When ($\star$) doesn't come out true, and the risk isn't worth taking, we should say that the subject doesn't know p. In the cases at issue, the subject's credence in p is high (typically significantly higher than their credence in $1 - Pr(p)$), but the ratio of potential losses to potential gains is lopsided as well: G is significantly smaller than L. Given this, ($\star$) doesn't come out true. In the high-stakes cases like the bank case, L is giant, but part of what Weatherson is bringing out here is that this isn't essential — it just needs to be big enough in proportion to the potential gains. If the odds of p are (say) 9 to 1, then if L is (say) ten times greater than G, the risk won't be worth taking, regardless of how large or small (in more absolute terms) L is.

The claim that a subject's 'interests matter' to whether they know can be interpreted in a number of ways, some trivially true, others obviously false. One challenge for the interest-relativist is to pin down the sense that is central to their view. The details we just ran through seem to do this pinning down in Weatherson's IRT. But Weatherson quickly complicates this story in a few ways. For me, these complications were significant enough to leave me wondering how, in the end, interests do their work in IRT.

The complications come in two spots. First, Weatherson argues that it's not just practical interests that matter, but theoretical ones. And second, he argues that it's often rationally permissible to inquire into what you know. Both of these raise a number of questions for me. I'll discuss each in turn.

First, Weatherson argues that it's not just our practical interests (or our practical situations more generally) that can undermine our knowledge, but our interests more generally, and in particular our 'theoretical' interests. If one were in a particular kind of epistemic decision-theoretic frame of mind, one might imagine Weatherson's thought here unfolding parallel to the practical case: We have particular kinds of theoretical/epistemic interests — we want true beliefs, more accurate credences, to avoid falsehood and inconsistency, and

so on. We are faced with decisions about how to update our doxastic states, and in these decisions we are confronted with potential (theoretical) losses and gains given our theoretical interests. And then, just as in the practical case, we can get into situations where, given that we are facing a particular sort of (doxastic, epistemic) decision, something we previously knew will no longer count as knowledge given those potential losses and gains. Epistemic decisions can be risky too.

But this is not how Weatherson's theoretical story unfolds. How do theoretical interests matter in IRT? It's worth looking at the examples he gives. The first (pp.49, 104) is a case just like our original case of Anisa and the Red-Blue game, except that we're to imagine that instead of actually playing the game, Anisa is just imagining what she would do were she to play the game. When she's going through this hypothetical scenario, she shouldn't be indifferent between the two hypothetical bets; she should still prefer Red-True. But in this case she's not facing any practical decision at all. Some interests of Anisa's seem to be undermining her knowledge that b, but those don't look like standard practical interests.[2]

I think this is an interesting case, but I don't know exactly what we can take it to show. While it's true that Anisa isn't facing an actual decision about how to bet, she still has to reason from something like practical interests (even if these are hypothetical ones). Moreover, I don't have a strong intuition about whether Anisa retains her actual knowledge when she performs this thought experiment, nor am I sure that Weatherson's IBE-style argument gets us to that conclusion. I like this case, but I'm just not sure what it can tell us about how our theoretical interests impact our knowledge.

Weatherson really only gives us one other sort of case to bring out the way in which he thinks that our theoretical interests can undermine our knowledge — cases that involve some sort of (illicit) bootstrapping. Here is one version of this (p.20): Say S has a body of evidence E that supports p in the right sort of way to put S in the position to know p. S reasons from E and comes to know p. But then S wants to figure out more precisely the extent to which E supports p, e.g. S wants to know the evidential probability of p. It seems S shouldn't reason from p in this case and conclude that the evidential probability of p is 1.[3] A related, and more concrete case, is as follows (p.104): Say it's the day after Anisa has read her history book and learned that the Battle of Agincourt took place in 1415 (b). Instead of playing any sort of betting game, she starts wondering how likely it is that her history book is right about b ($\mathtt{B}$).[4] Again,

[2] Weatherson here is drawing on an insight from Fantl and McGrath (2009) — what they call the 'Unity Thesis'.

[3] Confession. This needn't be so bad?

[4] I'll use these typewriter letters for questions.

it seems as though Anisa shouldn't take b as a premise in her reasoning about $\mathtt{B}$ here. That is, it seems she shouldn't reason: the book says b and b is true, so the book is right about b.

There is a lot to say about this case too, but what I want to bring out is just how different this sort of case is from the main Anisa Red-Blue case (and all the standard cases of pragmatic encroachment). Here, the idea seems to be that Anisa's interest in the question $\mathtt{B}$ undermines her knowledge that b.[5] But the mere fact that Anisa is interested in or even actively investigating the question of how she should play the Red-Blue game does not, on its own, undermine her knowledge that b. Whether that knowledge is undermined also depends on some of her other practical interests, e.g. getting money. But in the sorts of theoretical bootstrapping cases at issue now, there are no analogous (theoretical) interests. In that case, just thinking about the question seems to undermine knowledge. But (as we saw) just thinking about whether the bank will be open tomorrow or whether one should stop at the bank now, does not, on its own, undermine one's knowledge of the bank's opening hours — both Lois and Hiram do this.

In general, in the standard encroachment cases, whether one's knowledge is undermined depends not (or not only) on the fact that one is investigating some question, but on a distinct variable: one's practical interests. But in the main theoretical cases Weatherson discusses, it's the mere fact of investigating that undermines some knowledge, regardless of any other interests.

So we have what seems like two quite different phenomena here, two quite different ways in which 'interests matter' to what we know. In the theoretical cases, no ($\star$)-involving explanation is doing any work. Do we have a principled explanation of what is happening in the relevant sort of theoretical cases? Weatherson says no. He says, "individual inquiries have their own logic, their own rules about what can and can't be taken for granted" (p.104).

That said, this sort of talk of 'taking things for granted' in inquiry is a way that Weatherson unites these two different ways in which interests matter in IRT. Neither our original Red-Blue game Anisa nor our potentially bootstrapping Anisa is allowed to take b for granted in their inquiries. And that, for Weatherson, means (or at least shows) that neither Anisa knows. He says, "Knowledge goes away if the thinker starts conducting an inquiry where the purported knowledge is an inappropriate starting point" (p.2), and "only

[5] In fact, I'm not sure this is the best description of what's going on with Anisa. It does not seem to be her mere interest in $\mathtt{B}$ that is doing the work. She plausibly had a standing interest in $\mathtt{B}$ when she read the book and did come to know b. The crucial thing in this case is more plausibly Anisa's starting an actual inquiry into $\mathtt{B}$ (which might involve a kind of interest in $\mathtt{B}$, but also more than that).

knowledge is appropriate as a starting point for inquiry" (p.14).

The idea that only knowledge is an appropriate starting point for inquiry needs to be hemmed in to be plausible. We start inquiries on hunches, suspicions, hypotheses, and all manner of vibes. These are all perfectly appropriate starting points in some sense of 'starting point'. As I was reading I had in mind that Weathersonian starting points were the propositions an inquirer used as premises in some inquiry-relevant (deductive) reasoning. Weatherson seems to have something like this in mind as well. He says, "the inquirer knows that p only if they can properly use p as a reason in "thought directed at the question" they are considering. That is, they can use p as a step in this reasoning" (p.103). When Anisa is trying to figure out $\mathtt{B}$, she can't use b as a step in her $\mathtt{B}$-relevant reasoning, so she doesn't know.[6]

So maybe the big picture is something like this: when we look at the practical cases — the standard pragmatic encroachment cases — we find an instance of a more general phenomenon. When we are investigating some question $\mathtt{Q}$, there is typically a body of information we can permissibly reason from in that investigation — these are the inquiry's 'appropriate starting points'. Weatherson is arguing that something is an appropriate starting point in this sense, just in case it is known. And part of what he shows is that this means that starting an inquiry into a question can change what you know, either alone or in concert with some of the inquirer's other interests.

A bit of a wrench gets thrown into this in Chapter 5 though. This chapter is framed as a response to an objection — what Weatherson calls the 'objection from double checking'. The (putative) problem is that it seems as though there are cases in which subjects who know p reasonably double-check whether p is in fact true, e.g. I know I locked the door, but I double-check (seems fine). Weatherson worries that this is a problem for IRT. I believe his thought is roughly this. Say p is the complete answer to $\mathtt{Q}$ and S knows p at t_1. Then S starts checking $\mathtt{Q}$ at t_2. Either this triggers a loss of knowledge in p or not. If it does, then is S really double-checking (given that they aren't really checking at t on something they know at t)? But if S doesn't lose knowledge, then is it really reasonable to investigate further (e.g. by going back and fiddling with the door)? Can't they just reason: p, therefore p? Weatherson doesn't like either result.

What Weatherson argues instead is that, "It is possible to properly conduct an inquiry into p, even when one knows that p, and even when knowledge provides appropriate starting points for inquiry. That's because it is often

[6]This aligns with a way others have characterized pragmatic encroachment, e.g. Hawthorne (2004). Also see Willard-Kyle (2023) for a different way in which we have to start inquiry with knowledge.

appropriate to deliberately restrict oneself in inquiry, and use fewer resources than are otherwise available" (p.125). The idea here is that our knowledge is kind of like evidence in a court of law: sometimes known facts cannot be used as evidence in a trial. Our knowledge, according to Weatherson, is like that: sometimes we know things that we shouldn't 'use' in a given inquiry. I think this means that sometimes we know p, and p is an appropriate starting point for some inquiry into $\mathtt{Q}$, but we cannot use p as a premise in our reasoning in that inquiry into $\mathtt{Q}$. But this puts us at an impasse, given what I said a few paragraphs ago about Weathersonian starting points.

Weatherson manoeuvres around this impasse by distinguishing kinds of appropriateness. For Weatherson, to say that some proposition p is an appropriate starting point into some inquiry is not to say that it's appropriate in every sense, but just in one particular alethic sense. These kinds of appropriate starting points are ones the use of which in reasoning leaves the inquirer/reasoner immune from a particular kind of criticism — criticism that p might not be true (p.110).[7] In some double-checking cases it can be 'alethically appropriate' to reason from some known proposition, but not appropriate in some other sense. So were I to reason, the door is locked, so the door is locked, I would be immune from criticism that my premise might not be true, but still might be doing something inappropriate (e.g. if my aim were to get a new piece of evidence on the matter).

I have some questions about the sort of criticism at issue here and the normative work it is doing. But I want to shelve those and instead think about the picture that results now. In particular, the space opened up now allows for cases structurally like the ones we've been discussing — ones in which a subject's interests seem to undermine their knowledge — but in which knowledge remains intact but can't be appropriately used. It is not obvious how to distinguish these cases (which Weatherson discusses in Chapter 5) from our original ones. The overall effect for me, given this (plus some of the details I already discussed), was that it is hard, in the end, to pin down just when or how our interests can undermine our knowledge according to IRT.

The central case in Chapter 5 is Jessica Brown's well-known surgeon case (see Brown (2008)). This case centres around a surgeon who has to perform a serious operation on a patient — remove a diseased left kidney. We're to imagine that the surgeon has done all the relevant preparations for the surgery, and when she's en route to the hospital for the surgery, she knows it's the left kidney to be removed (k). But as the surgeon is heading into the operating

[7] Alternative description of the immunity that comes with alethic appropriateness: subjects are "immune to criticism on the grounds that what they are doing is epistemically risky" (p.137).

room, she has the option of looking at the patient's chart again and seeing what it says about which kidney is to be removed. The surgeon does look at the chart, and (it seems) should look at the chart. But, Brown argues — and Weatherson agrees — the surgeon nonetheless does know k. Brown and Weatherson take this to be a case in which the surgeon knows k but is allowed (or even required) to check the chart and see whether k is indeed true.

But how, for Weatherson, is this case importantly different from the Anisa Red-Blue game case or the original banks cases? Brown's surgeon case has all the main markings of the original cases (wasn't that part of the force of the case qua counterexample?), but in this case we're to conclude that the surgeon knows, while in the earlier cases we were to conclude that the relevant subjects don't know.[8] And going back to the Weatherson's claim about alethic appropriateness: in the surgeon case we're to say that were the surgeon to reason from k, she would be immune from the relevant sort of alethic criticism, but this is not the case for Anisa or Hiram.

This isn't to say that Weatherson doesn't offer more of an explanation of what he takes to be going on with our surgeon. And what he says there — about 'sensitivity chasing' — is interesting and plausible. My worry is just that if we're looking to understand how IRT works, and in particular how it is that practical interests can encroach on what we know, it seems as though our best guide — the one that went by way of ($\star$) — is no longer offering guidance. ($\star$) is at the core of a fantastic potential explanation of why the surgeon should check the chart. But that explanation also seems to tell us that the surgeon doesn't know k (at least not at the crucial moment).[9]

My grip on how theoretical interests can undermine knowledge is loosened by the cases in Chapter 5 as well. In the theoretical case we already had less guidance on the mechanism of interest-driven knowledge loss, but we had something of a clear case. Anisa shouldn't bootstrap her way to answering the question, 'Is the book right about the date of this battle?', by reasoning from what she learned about the date of the battle in the book. Weatherson uses this thought to argue that when Anisa is investigating the reliability of the book on this matter, she no longer knows b. But a case very much like this one also comes up in Chapter 5 as well. In that case though, I take it that Weatherson thinks that the relevant subject does know.

[8] Perhaps the closest case to Brown's surgeon case in *Knowledge* is the Darja case (pp.47-8). There, Darja is faced with a choice of acting on p or checking Google as to whether p is true and then acting. Here, Weatherson argues that Darja should check and therefore doesn't know p.

[9] If we think the surgeon is certain that k is true, then perhaps ($\star$) gets us the result that Weatherson is after here. He doesn't say that we should think of the case that way, but he does think that checking the chart doesn't maximize EU for the surgeon, which points in a certainty-ish direction.

In this case (p.129) we imagine Florian, who owns the best coffee scale around, and one he's used reliably many times before. But Florian nonetheless starts wondering whether the scale is reliable. Like the case of Anisa wondering about the reliability of her history book, it seems as though Florian shouldn't just reason from the premises that the scale said 24g and that the weight of the coffee is 24g to the conclusion that the scale is fine. But Weatherson characterizes the Florian case as a type of case in which one can (reasonably) inquire into whether some proposition is true while retaining one's knowledge of that proposition.[10] While there are some differences between this case and the Anisa anti-bootstrapping case, again, it's hard to see why those should make a difference. If Anisa loses her knowledge, why doesn't Florian? If Anisa could be criticized for using a premise that might be false, why couldn't Florian be so criticized as well?

In general, the cases in Chapter 5 open up a particular kind of conceptual space. They are — very much like the canonical encroachment cases — cases in which subjects are not permitted reason from some proposition in their inquiry and not permitted to take some propositions for granted in those inquiries. In these new Chapter 5 cases though, the reason those propositions cannot be used is not that they are not known. Instead, those inquirers have other reasons not to use those propositions. But couldn't we say the same about Red-Blue Anisa, potentially bootstrapping Anisa, and Darja? Why do they lose their knowledge while the new inquirers don't?[11]

REFERENCES

Brown, J. 2008. Subject-Sensitive Invariantism and the Knowledge Norm for Practical Reasoning. *Noûs*, 42(2):167–189. [8]

Fantl, J. and McGrath, M. 2009. *Knowledge in an Uncertain World*. New York: Oxford University Press. [5]

Hawthorne, J. 2004. *Knowledge and Lotteries*. Oxford University Press. [7]

Willard-Kyle, C. 2023. The Knowledge Norm for Inquiry. *Journal of Philosophy*, 120(11):615–640. [7]

[10] Weatherson doesn't explicitly say that this is how we should read the Florian case, viz. as one in which Florian knows p but is 'sensitivity chasing' with respect to p, but it looks as though that's exactly the work the case is meant to be doing.

[11] Special thanks to the Haribo Crew for helpful discussion: Tez Clark, Eugene Ho, Ali Rezaei, Richard Roth, Ethan Russo, and Chris Sun.