Responses to Referee comments

I didn't have much to do in response to referee 1's comments, and I've basically adopted all of the suggestions. The only thing I've done that referee 1 might not like is drop the example of Con, which referee 2 didn't like. I needed to cut something to make enough space to respond to other comments, and since a referee didn't like the example, it is now gone.

There was more to be done in response to referee 2's comments, so I'll go through point by point the replies.

> One problem, though, as I see it, is that there are really two papers here.

I basically agree. Indeed, I've written something like the other paper the reviewer as well. The opening sections weren't designed to really defend, or even fully lay out, a version of IRI, so much as they were designed to say enough to make what follows make some sense. I've clarified a little what's going on, and included a citation to the other paper. (It's currently blinded, but it will be filled in if the paper is accepted for publication.)

> As I said, I think this case could be used to argue for a novel kind of IRI. . . . No effort is made to show that Stanley, Hawthorne, et al support the author's view—that it is in some sense "what really matters" to those trying to defend versions of IRI.

I think the exegetical issues here are pretty delicate, and they are certainly less clear than I suggested in the original version. So I'm glad the reviewers pushed me on this.

Just whether it's best to interpret Hawthorne, Stanley et al as having a different view to the one I'm putting forward, or a view that is neutral on some of the questions I'm taking a stance on, or of having something like the same view, could be the subject of a long (and boring!) paper analysing the nuances of what they'd written. The answer would probably be different for Hawthorne and Stanley and Fantl and McGrath. Rather than take sides on this, I wanted to simply not claim credit that might not be due to me. So I've added a footnote that is much more neutral on the exegetical issues than was the original paper, noted that I'm happy to accept any judgments of originality that come my way, but made clear that I think the primary value in what I'm doing is in the *defence* of IRI, not the *articulation* of it.

> The author repeatedly presents a dilemma meant to motivate IRI. One horn has the agent believing that what is in fact a bad choice is really a good choice. (In Stanley's bank case this would presumably be Hannah's decision to wait until Saturday to go to the bank when the financial consequences of its being then closed would be dire.) The author says without defense that this is "irrational." Is it really irrational? After all, by hypothesis, the bank will be open on Saturday (otherwise this would itself be enough to rule out knowledge), and, again by hypothesis, Hannah has good evidence that it will then be open. (She has the kind of evidence that would support knowledge were less riding on the truth of what's believed.) So more needs to be said to defend the charge of irrationality.

I think there is more that can be said here. Basically, waiting until Saturday doesn't, I think, maximise expected utility, and the situation calls for maximising expected utility. But both conjuncts of that last sentence need further defence, and the paper was getting long. So the paper restricts itself to a more cautious claim: holding fixed that the agent is an expected utility maximiser, the change in interests changes whether she knows. That's enough to show that knowledge is interest-relative, which is all the aim is here. Defending the stronger claims about how much interest-relativity there is is, I think, a job best left for a paper setting out a particular version of IRI.

> I suggest eliminating section 3 altogether. The author provides no evidence that Stanley, Hawthorne, Fantl and McGrath, and other IRI theorists regard themselves as offering an "existential" theory. The author can hold off on saying that he favors this (weaker) version of IRI until section 9 where he uses it to respond to Jessica Brown's counterexample.

Basically done. I kept a little bit of section 3 at the start, though nothing like a full section, then moved the rest to the discussion of Brown.

> The author needs to give a precise characterization of the version of IRI from which the verdict that there is knowledge in Low Stakes Bridge but no knowledge in High Stakes bridge follows from conformity of the cases to the equation on lines 47-53 of p. 7.

I've been much more explicit in the setup about how odds relate to belief and hence to knowledge. So I've made explicit the principle **Relevant Practical Coherence** that does a lot of the work. Given that principle, and given that the agent's preferences follow expected utilities, the equation in question follows with some fairly basic algebra. I've omitted the algebra in question, though if need be I can include that too. But I think the request was for the theory from which this followed, not for the algebraic workings out, and that is now in the initial section.

> Moreover, it is unclear to me that this part of the defense should be published, as the criticism to which the author is responding has not yet published.

The paper I'm responding to has now appeared in print.

> The variants on Stanley's bank case that Russell and Doris use to argue against IRI are not presented with sufficient detail and clarity for the reader to gauge the effectiveness of the author's response to them.

At referee's 2 request, I've spelled out the cases Stanley uses, and Russell and Doris's variants on them somewhat. I still haven't filled in **all** the details though for a couple of reasons. First, if a reader really wants more details, Russell and Doris's paper isn't that hard to find - it is in this journal! Second, I wanted my descriptions to be somewhat abstract. I think that much of the debate about IRI has gotten caught up in inessential features of examples, and it helps to have descriptions of the cases that only focus on the core features. There is a cost to this, I acknowledge; it means intuitions are not as immediate or sharp. But I think the way it helps clarify what is at stake in the examples more than makes up for this cost.

> The author discusses a fake barn kind of case (Harman's dictator) where a subject reads a true report of a dictator's demise unaware that there are lots of misleading reports of his survival—reports she is in no position to rebut. The author says that if the subject doesn't care at all about politics (and is therefore presumably in little danger of reading the misleading reports) the misleading reports do not undermine her knowledge of the dictator's death. I find this claim unintuitive. At the very least, it requires argumentative defense.

I've added an example that is meant to respond to this concern, the example about global warming.

> The case of George and Ringo should not be introduced until after Blome-Tillmann's argument is introduced. The current presentation is confusing.

Done.

> Again, take out "con" and everything from l. 46 on p. 14 right up to the end of the section. The point's been made. Blome-Tillmann's criticism has been shown to be ill considered. No need to elaborate further.

Done.

> I did not find the response to Neta's criticism at all convincing. The subject can have the antecedent belief that street signs are accurate and still look for further evidence as to whether she is on Main. There is nothing at all incoherent about this.

I've tidied the response to Neta up to clarify what the relevant background belief of Kate's is. I think the original draft suggested that what mattered was whether she had a belief that the generic *street signs around here are accurate* is true. But what I really meant for her to be focussing on was whether the very street signs she can see are correct, i.e., whether everything she can see on a street sign is true. If she has that belief, then her observational evidence entails she is on Main, so there is no need for further investigation. Of course, if she just has the generic belief, it is possible that the sign that says Main is one of the outliers, so there's no entailment that she's on Main. I'm glad the reviewer pushed me to clean this up.