Knowledge and Interests

Brian Weatherson

---

1 The Variable Threshold Model of Belief

2 Weaknesses and Refinements

3 Fantl and McGrath's Objections

Fantl and McGrath's primary complaint against the view in Weatherson (2005) is that it is not strong enough to entail principles such as (JJ).

(JJ) If you are justified in believing that p, then p is warranted enough to justify you in ϕ-ing, for any ϕ. (Fantl and McGrath, 2009, 99)

It's true that the variable threshold view cannot be used to derive (JJ), at least on its intended reading. But that's because on the intended reading, it is false, and the variable threshold view is true. So the fact that (JJ) can't be derived is a feature, not a bug.

The problem arises because of cases like the case we just discussed, namely cases where the agent has irrational beliefs elsewhere in her web of belief. Let's lay out the case carefully, because it will have a large role in the rest of this section. Here's what we're going to stipulate about S.

- She knows that p and q are independent, so her credence in any conjunction where one conjunct is a member of $\{p, \neg p\}$ and the other is a member of $\{q, \neg q\}$ will be the product of her credences in the conjuncts.
- Her credence in p is 0.99, just as the evidence supports.
- Her credence in q is also 0.99. This is unfortunate, since the rational credence in q given her evidence is 0.01.
- She has a choice between taking and declining a bet with the following payoff structure.

    - If $p \wedge q$, she wins $100.
    - If $p \wedge \neg q$, she wins $1.
    - If $\neg p$, she loses $1000.

- The marginal utility of money is close enough to constant that expected dollar returns correlate more or less precisely with expected utility returns.

As can be easily computed, the expected utility of taking the bet given her credences is positive, it is just over $89. Our agent S takes the bet. She doesn't compute the expected utility, but she is sensitive to it. That is, had the expected utility given her credences been close to 0, she would have not acted until she made a computation. But from her perspective this looks like basically a free $100, so she takes it. Happily, this all turns out well enough, since p is true. But it was a dumb thing to do. The

expected utility of taking the bet given her evidence is negative, it is a little under -$8. So she isn't warranted, given her evidence, in taking the bet.

I also claim the following three things are true of her.

1. p is not justified enough to warrant her in taking the bet.
2. She believes p.
3. This belief is rational.

The argument for 1 is straightforward. She isn't warranted in taking the bet, so p isn't sufficiently warranted to justify it. This is despite the fact that p is obviously relevant. Indeed, given p, taking the bet strictly dominates declining it. But still, p doesn't warrant taking this bet. Had the rational credence in p been higher, then the bet would have been reasonable. Had the reasonable credence in p been, say, 0.9999, then she would have been reasonable in taking the bet, and using p as a reason to do so. So there's a good sense in which p simply isn't warranted enough to justify taking the bet.[1]

The argument for 2 is that she has a very high credence in p, this credence is grounded in the evidence in the right way, and it leads her to act as if p is true, e.g. by taking the bet. It's true that her credence in p is not 1, and if you think credence 1 is needed for belief, then you won't like this example. But if you think that, you won't think there's much connection between (JJ) and pragmatic conditions in epistemology either. So that's hardly a position a defender of Fantl and McGrath's position can hold.

The argument for 3 is that her attitude towards p tracks the evidence perfectly. She is making no mistakes with respect to p. She is making a mistake with respect to q, but not with respect to p. So her attitude towards p, i.e. belief, is rational.

I don't think the argument here strictly needs the assumption I'm about to make, but I think it's helpful to see one very clear way to support the argument of the last paragraph. The working assumption of this paper is that talking about beliefs and talking about credences are simply two ways of modelling the very same things, namely minds. If the agent both has a credence 0.99 in p, and believes that p, these are not two different states. Rather, there is one state of the agent, and two different ways of modelling it.[2] So it is implausible, if not incoherent, to apply different valuations to the state depending on which modelling tools we choose to use. That is, it's implausible to say that while we're modelling the agent with credences, the state is rational, but when we change tools, and start using beliefs, the state is irrational. Given this outlook on beliefs and credences, premise 3 seems to follow immediately from the setup of the example.

So that's the argument that (JJ) is false. And if it's false, the fact that the variable threshold view doesn't entail it is a feature, not a bug. But there are a number of

---

[1] I think this is exactly the sense in which 'is warranted enough' is being used in (JJ), though I'm not entirely sure about this. For present purposes, I plan to simply interpret (JJ) that way, and not return to exegetical issues.

[2] We'll return to this point, at some length, in the discussion of Sturgeon below.

possible objections to that position. I'll spend the rest of this section going over them.[3]

Objection: The following argument shows that *S* is not in fact justified in believing that *p*.

1. *p* entails that *S* should take the bet, and *S* knows this.
2. If *p* entails something, and *S* knows this, and she justifiably believes *p*, she is in a position to justifiably believe the thing entailed.
3. *S* is not in a position to justifiably believe that she should take the bet.
C. So, *S* does not justifiably believe that *p*

Reply: The problem here is that premise 1 is false. What's true is that *p* entails that *S* will be better off taking the bet than declining it. But it doesn't follow that she should take the bet. Indeed, it isn't actually true that she should take the bet. This case might be a little controversial, but the general point here should not be controversial. If I'm standing by a roulette wheel, the thing that will make me best off is betting heavily on the number than will actually come up. But I shouldn't do that; I shouldn't place any bets at all, since all the bets have a highly negative expected return. Only a very crude kind of consequentialism would identify what I should do with what will have the best returns, and that crude consequentialism isn't true.

Objection: Even though *p* doesn't *entail* that *S* should take the bet, it does provide inductive support for her taking the bet. So if she could justifiably believe *p*, she could justifiably (but non-deductively) infer that she should take the bet. Since she can't justifiably infer that, she isn't justified in taking the bet.

Reply: The inductive inference here looks weak. One way to make the inductive inference work would be to deduce from *p* that taking the bet will have the best outcomes, and infer from that that the bet should be taken. But the last step doesn't even look like a reliable ampliative inference. The usual situation is that the best outcome comes from taking an *ex ante* unjustifiable risk.

It may seem better to use *p* combined with the fact that conditional on *p*, taking the bet has the highest *expected* utility. But actually that's still not much of a reason to take the bet. Think again about cases, completely normal cases, where the action with the best outcome is an *ex ante* unjustifiable risk. Call that action ϕ, and let $B\phi$ be the proposition that ϕ has the best outcome. Then $B\phi$ is true, and conditional on $B\phi$, ϕ has an excellent expected return. But doing ϕ is still running a dumb risk. Since these kinds of cases are normal, it seems it will very often be the case that this form of inference leads from truth to falsity. So it's not a good inductive inference.

More generally, we should worry quite a lot about *S*'s ability to draw inductive inferences about the propriety of the bet here. Unlike deductive inferences, inductive

---

[3]Thanks here to a long blog comments thread with Jeremy Fantl and Matthew McGrath for making me formulate these points much more carefully. The original thread is at http://tar.weatherson.org/2010/03/31/do-justified-beliefs-justify-action/.

inferences can be defeated by a whole host of factors. If I've seen a lot of swans, in a lot of circumstances, and they've all been blue, that's a good reason to think the next swan I see will be blue. But it ceases to be a reason if I am told by a clearly reliable testifier that there are green swans in the river outside my apartment. And that's true even if I dismiss the testifier because I think he has a funny name, and I don't trust people with funny names. Now although S has evidence for p, she also has a lot of evidence against q, evidence that she is presumably ignoring since her credence in q is so high. Any story about how S can reason from p to the claim that she should have to take the bet will have to explain how her irrational attraction to q doesn't serve as a defeater, and I suspect that will be a high bar to cross.

Objection: In the example, S isn't just in a position to justifiably believe p, she is in a position to *know* that she justifiably believes it. And from the fact that she justifiably believes p, and the fact that if p, then taking the bet has the best option, she can infer that she should take the bet.

Reply: It's possible at this point that we get to a dialectical impasse. I think this inference is non-deductive, because I think the example we're discussing here is one where the premises are true and the conclusion false. Presumably someone who doesn't like the example will think that it is a good deductive inference.

Unlike the inductive inference mentioned in the previous objection, I think this is generally a good inductive inference. Whenever you justifiably believe p, and the best outcome given p is gained by doing ϕ, then *usually* you should ϕ. But what's usually true isn't always true, and inferences from things to common correlates can be defeated if you have related irrational beliefs. (That's what the swan example above shows.) So if S tried to infer this way that she should take the bet, her irrational confidence in q would defeat the inference.

Unlike the previous replies, perhaps, I don't think this reply will really convince the person who presents the objection. But since one form of the objection, the form where the objector thinks that S can reason *deductively* that she should take the bet, presupposes the view I'm arguing against, there's not much I can do about that.[4]

Objection: If S were ideal, then she wouldn't believe p. That's because if she were ideal, she would have a lower credence in q, and if that were the case, her credence in p would have to be much higher (close to 0.999) in order to count as a belief. So her belief is not justified.

Reply: The premise here, that if S were ideal she would not believe that p, is true. The conclusion, that she is not justified in believing p, does not follow. It's always a mistake to *identify* what should be done with what is done in ideal circumstances. This is something that has long been known in economics. The *locus classicus* of the

---

[4]Having said all that, the more complicated example at the end of Weatherson (2005) was designed to raise the same problem without the consequence that if p is true, the bet is sure to return a positive amount. In that example, conditionalising on p means the bet has a positive expected return, but still possibly a negative return. So if you like this objection to the simplified argument here, the more complicated argument in that paper might still persuade you.

view that this is a mistake is Lipsey and Lancaster (1956-1957). A similar point has been made in ethics in papers such as Watson (1977) and Kennett and Smith (1996a,b). And it has been extended to epistemology by Williamson (1998).

All of these discussions have a common structure. It is first observed that the ideal is both F and G. It is then stipulated that whatever happens, the thing being created (either a social system, an action, or a cognitive state) will not be F. It is then argued that given the stipulation, the thing being created should not be G. That is not just the claim that we shouldn't *aim* to make the thing be G. It is, rather, that in many cases being G is not the best way to be, given that F-ness will not be achieved. Lipsey and Lancaster argue that (in an admittedly idealised model) that it is actually quite unusual for G to be best given that the system being created will not be F.

It's not too hard to come up with examples that fit this structure. Following Williamson, we might note that I'm justified in believing that there are no ideal cognitive agents, although were I ideal I would not believe this. Or imagine a student taking a ten question mathematics exam who has no idea how to answer the last question. She knows an ideal student would correctly answer an even number of questions, but that's no reason for her to throw out her good answer to question nine. In general, once we have stipulated one departure from the ideal, there's no reason to assign any positive status to other similarities to the idea. In particular, given that S has an irrational view towards q, she won't perfectly match up with the ideal, so there's no reason it's good to agree with the ideal in other respects, such as not believing p.

Stepping back a bit, there's a reason the variable threshold view says that the ideal and justification come apart right here. On the variable threshold view, like on any pragmatic view of mental states, the *identification* of mental states is a somewhat holistic matter. Something is a belief in virtue of its position in a much broader network. But the *evaluation* of belief is (relatively) atomistic. That's why S is justified in believing p, although if she were wiser she would not believe it. If she were wiser, i.e., if she had the right attitude towards q, the very same credence in p would not count as a belief. Whether her state counts as a belief, that is, depends on wide-ranging features of her cognitive system. But whether the state is justified depends on more local factors, and in local respects she is doing everything right.

Objection: Since the ideal agent in S's position would not believe p, it follows that there is no *propositional* justification for p. Moreover, doxastic justification requires propositional justification[5] So S is not doxastically justified in believing p. That is, she isn't justified in believing p.

Reply: I think there are two ways of understanding 'propositional justification'. On one of them, the first sentence of the objection is false. On the other, the second sentence is false. Neither way does the objection go through.

The first way is to say that p is propositionally justified for an agent iff that agent's evidence justifies a credence in p that is high enough to count as a belief *given the agent's other credences and preferences*. On that understanding, p is propositionally

---

[5]See Turri (2010) for a discussion of recent views on the relationship between propositional and doxastic justification. This requirement seems to be presupposed throughout that literature.

justified by S's evidence. For all that evidence has to do to make p justified is to support a credence a little greater than 0.9. And by hypothesis, the evidence does that.

The other way is to say that p is propositionally justified for an agent iff that agent's evidence justifies a credence in p that is high enough to count as a belief *given the agent's preferences and the credences supported by that evidence.* On this reading, the objection reduces to the previous objection. That is, the objection basically says that p is propositionally justified for an agent iff the ideal agent in her situation would believe it. And we've already argued that that is compatible with doxastic justification. So either the objection rests on a false premise, or it has already been taken care of.

Objection: If S is justified in believing p, then S can use p as a premise in practical reasoning. If S can use p as a premise in practical reasoning, and p is true, and her belief in p is not Gettiered, then she knows p. By hypothesis, her belief is true, and her belief is not Gettiered. So she should know p. But in the previous section it was argued that she doesn't know p. So by several steps of modus tollens, she isn't justified in believing p.[6]

Reply: Like the previous objection, this one turns on an equivocation, this time over the neologism 'Gettiered'. Some epistemologists use this to simply mean that a belief is justified and true without constituting knowledge. By that standard, the third sentence is false. Or, at least, we haven't been given any reason to think that it is true. Given everything else that's said, the third sentence is a raw assertion that S knows that p, and I don't think we should accept that.

The other way epistemologists sometimes use the term is to pick out justified true beliefs that fail to be knowledge for the reasons that the beliefs in the original examples from Gettier (1963) fail to be knowledge. That is, it picks out a property that beliefs have when they are derived from a false lemma, or whatever similar property is held to be doing the work in the original Gettier examples. Now on this reading, S's belief that p is not Gettiered. But it doesn't follow that it is known. There's no reason, once we've given up on the JTB theory of knowledge, to think that whatever goes wrong in Gettier's examples is the *only* way for a justified true belief to fall short of knowledge. It could be that there's a practical defeater, as in this case. So the second sentence of the objection is false, and the objection again fails.

But it's worth pausing for a bit to reflect on why we should say that S does not know that p. My main reason for saying this comes from thinking about how we lay out examples in decision theory. Compare the following two examples.

Case 1 S has a counter and two options about what to do with it. If she places it on the table, she will get \$100,000. If she places it in her pocket, she will get nothing. There are no other salient choices, and there are no other effects associated with her action, and she prefers more money to less. What should she do?

---

[6]Compare the 'subtraction argument' on page 99 of ?.

Case 2 *S* has a counter and two options about what to do with it. If she places it on the table, she is betting $100,000 that the ball in the roulette wheel will land on a red number. There is a $^{18}/_{37}$ probability that this will happen, and if she bets, she'll win $100,000 if the ball lands on a red, and lose $100,000 if it does not. If she puts the counter in her pocket, she'll get nothing. There are no other salient choices, and there are no other effects associated with her action, and she prefers more money to less, and the marginal utility of money is constant for her over the financial range being discussed. What should she do?

If we were in a standard decision theory class, the answers would be easy. She should put the counter on the table in the first case, and in her pocket in the second.

But there's nothing in the description of the two cases that rules out their describing the very same facts. If in Case 2 the ball does actually land on a red, then everything that's true in Case 1 will be true in Case 2. And in Case 1, she should put the counter on the table. Presumably it would be a mistake to infer by modus tollens that the ball will not land on a red in Case 2, although a literal reading of the last few paragraphs would suggest that would be a good inference to draw. What is going on?

I think the obvious thing to say is that when we set up cases like Case 1 and Case 2, we don't just stipulate what's true in the story. We also stipulate that the agent stands in some epistemic relationship (broadly construed) to the facts as laid out in the story. So even if the ball in the roulette wheel in Case 2 lands on a red number, that doesn't turn the case into Case 1. That's because for us to be in Case 1, the ball would not only have to land on a red, but *S* would have to stand in the appropriate epistemic relation to that landing.

The big question then is what this salient epistemic relation might happen to be. I think it is knowledge. I don't have anything like a *conclusive* argument for this. But I think there are a couple of reasons that push us that way. If we go with a relation weaker than knowledge, such as justified belief, then we won't be able to explain the fact that agents are meant to be able to rely on the structure of the problem, even in surprising circumstances. Agents are meant to have an indefeasible belief in the structure of the problem as stated, which is hard to do make consistent with the rationality of the agents unless you suppose that they know the problem has that structure. If you go with relation much stronger than knowledge, such as certainty, you'll be left with the problem that decision theory is not a particularly useful tool for modelling real-world interactions, since it is hard to ever be certain of anything. More positively, it's very natural when describing a decision theoretic problem to say that the agent knows various things that are in the setting out of the problem.

Now these problems don't only show up in decision theory textbooks. They also show up in, for example, epistemology papers. The problem with which we opened this section is just such a problem. Now let's say that the standard for including something in the description of the problem is knowledge. Then we could have simplified the problem a lot, as follows.

- If *S* takes the bet, she wins $100 if q and $1 if $\neg q$, and will not lose any money.
- If *S* declines the bet, she gets nothing.

And given that description of the bet, S should obviously take it, whatever her views on q. Since S shouldn't take the bet, that means the bet has been misdescribed. And that means S does not know that p.

4   Hawthorne and Stanley on Reasons

5   Strengths of the Variable Threshold Model

5.1   Knowledge by Indifference and Knowledge by Wealth

Gillian Russell and John Doris (2009) argue that Jason Stanley's account of knowledge leads to some implausible attributions of knowledge. Insofar as my theory agrees with Stanley's about the kinds of cases they are worried about, their objections are also objections to my theory. I'm going to argue that Russell and Doris's objections turn on principles that are *prima facie* rather plausible, but which ultimately we can reject for independent reasons.[7]

Their objection relies on variants of the kind of case Stanley uses heavily in his (2005) to motivate a pragmatic constraint on knowledge. Stanley imagines a character who has evidence which would normally suffice for knowledge that p, but is faced with a decision where A is both the right thing to do if p is true, and will lead to a monumental material loss if p is false. Stanley intuits, and argues, that this is enough that they cease to know that p. I agree, at least as long as the gains from doing A are low enough that doing A amounts to a bet on p at insufficiently favourable odds to be reasonable in the agent's circumstance.

Russell and Doris imagine two kinds of variants on Stanley's case. In one variant the agent doesn't care about the material loss. As I'd put it, the agent's indifference to material odds shortens the odds of the bet. That's because costs and benefits of bets should be measured in something like utils, not something like dollars. As Russell and Doris put it, "you should have reservations ... about what makes [the knowledge claim] true: not giving a damn, however enviable in other respects, should not be knowledge-making." (Russell and Doris, 2009, ??). Their other variant involves an agent with so much money that the material loss is trifling to them. Again, this lowers the effective odds of the bet, so by my lights they may still know that p. But this is somewhat counter-intuitive. As Russell and Doris say, "[m]atters are now even dodgier for practical interest accounts, because *money* turns out to be knowledge making." (Russell and Doris, 2009, ??) And this isn't just because wealth can purchase knowledge. As they say, "money may buy the *instruments* of knowledge ... but here the connection between money and knowledge seems rather too direct." (Russell and Doris, 2009, ??)

The first thing to note about this case is that indifference and wealth aren't really producing knowledge. What they are doing is more like defeating a defeater. Remember that the agent in question had enough evidence, and enough confidence,

---

[7]I think the objections I make here are similar in spirit to those Stanley made in a comments thread on Certain Doubts, though the details are new.

that they would know p were it not for the practical circumstances. I've been proposing a model where practical considerations enter debates about knowledge through two main channels: through the definition of belief, and through distinctive kinds of defeaters. It seems the second channel is particularly relevant here. And we have, somewhat surprisingly, independent evidence to think that indifference and wealth do matter to defeaters.

Consider two variants on Gilbert Harman's 'dead dictator' example (Harman, 1973, 75). In the original example, an agent reads that the dictator has died through an actually reliable source. But there are many other news sources around, defeaters, such that if the agent read them, she would lose her belief.

In the first variant, the agent simply does not care about politics. It's true that there are many other news sources around that are ready to mislead her about the dictator's demise. But she has no interest in looking them up, nor is she at all likely to look them up. She mostly cares about sports, and will spend most of her day reading about baseball. In this case, the misleading news sources are too distant, in a sense, to be defeaters. So she still knows the dictator has died. Her indifference towards politics doesn't generate knowledge - the original reliable report is the knowledge generator - but her indifference means that a would-be defeater doesn't gain traction.

In the second variant, the agent cares deeply about politics, and has masses of wealth at hand to ensure that she knows a lot about it. Were she to read the misleading reports that the dictator has survived, then she would simply use some of the very expensive sources she has to get more reliable reports. Again this suffices for the misleading reports not to be defeaters. Even before the rich agent exercises her wealth, the fact that her wealth gives her access to reports that will correct for misleading reports means that the misleading reports are not actually defeaters. So with her wealth she knows things she wouldn't otherwise know, even before her money goes to work. Again, her money doesn't generate knowledge - - the original reliable report is the knowledge generator - but her wealth means that a would-be defeater doesn't gain traction.

The same thing is true in Russell and Doris's examples. The agent has quite a bit of evidence that p. That's why she knows p. There's a potential practical defeater for p. But due to either indifference or wealth, the defeater is immunised. Surprisingly perhaps, indifference and/or wealth can be the difference between knowledge and ignorance. But that's not because they can be in any interesting sense 'knowledge makers', any more than I can make a bowl of soup by preventing someone from tossing it out. Rather, they can be things that block defeaters, both when the defeaters are the kind Stanley talks about, and when they are more familiar kinds of defeaters.

5.2 Stakes and Odds

In a so far unpublished note, Mark Schroeder (2008) has argued that interest-relative invariantists have erred by stressing variation in stakes as being relevant to knowledge. He argues, using examples of forced choice, that what is really relevant is the odds at which the agent has to make bets. Of course due to the declining marginal utility of material goods, high stakes bets will often be long odds bets. So there's a correlation between stakes and odds. But when the correlation comes apart, Schroeder

argues convincingly that it's the odds and not the stakes that are relevant to knowledge.

The variable threshold view of belief I've been defending agrees with Schroeder's judgments. Interests affect belief because whether someone believes p depends *inter alia* on whether their credence in p is high enough that any bet on p they actually face is a good bet. Raising the stakes of any bet on p does not change that, but changing the odds of the bets on p they face does change it. And that explains why agents don't have knowledge, or even justified belief, in some of the examples that motivate other interest-relative invariantists.

Although I think the variable threshold view gets those cases right, I don't take those examples to be a crucial part of the argument for the view. The core argument is that the view provides a better answer to Sturgeon's challenge of how we should integrate credences and beliefs into a single model. If it turned out that the facts about the examples were less clear than we thought, that wouldn't *undermine* the argument for the variable threshold view, since those facts weren't part of the original argument. But if it turned out that the facts about those examples were quite different to what the variable threshold view predicts, that may *rebut* the view, since it would then be shown to make false predictions.

This kind of rebuttal may be suggested by various recent experimental results, such as the results in May et al. (forthcoming) and Feltz and Zarpentine (2010). I'm going to concentrate on the latter set of results here, though I think that what I say will generalise to related experimental work. Feltz and Zarpentine gave subjects related vignettes, such as the following pair. (Each subject only received one of the pair.)

High Stakes Bridge John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a yawning thousand foot drop. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, 'I know that my truck will make it across the bridge.'

Low Stakes Bridge John is driving a truck along a dirt road in a caravan of trucks. He comes across what looks like a rickety wooden bridge over a three foot ditch. He radios ahead to find out whether other trucks have made it safely over. He is told that all 15 trucks in the caravan made it over without a problem. John reasons that if they made it over, he will make it over as well. So, he thinks to himself, 'I know that my truck will make it across the bridge.' (Feltz and Zarpentine, 2010, ??)

Subjects were asked to evaluate John's thought. And the result was that 27% of the participants said that John does not know that the truck will make it across in **Low Stakes Bridge**, while 36% said he did not know this in **High Stakes Bridge**. Feltz and Zarpentine say that these results should be bad for interest-relativity views. But it is hard to see just why this is so.

Note that the change in the judgments between the cases goes in the direction that the variable threshold view predicts. The change isn't trivial, even if due to the smallish sample size it isn't statistically significant in this sample. But should the variable threshold view have predicted a larger change? To figure this out, we need to ask three questions.

1. What are the costs of the bridge collapsing in the two cases?
2. What are the costs of not taking the bet, i.e., not driving across the bridge?
3. What is the rational credence to have in the bridge's sturdiness given the evidence John has?

None of these are specified in the story given to subjects, so we have to guess a little as to what the subjects' views would be.

Feltz and Zarpentine say that the costs in "High Stakes Bridge [are] very costly— certain death—whereas the costs in Low Stakes Bridge are likely some minor injuries and embarrassment." (Feltz and Zarpentine, 2010, ??) I suspect both of those claims are wrong, or at least not universally believed. A lot more people survive bridge collapses than you may expect, even collapses from a great height.[8] And once the road below a truck collapses, all sorts of things can go wrong, even if the next bit of ground is only 3 feet away. (For instance, if the bridge collapses unevenly, the truck could roll, and the driver would probably suffer more than minor injuries.)

We aren't given any information as to the costs of not crossing the bridge. But given that 15 other trucks, with less evidence than John, have decided to cross the bridge, it seems plausible to think they are substantial. If there was an easy way to avoid the bridge, presumably the *first* truck would have taken it.

But the big issue is the third question. John has a lot of information that the bridge will support his truck. If I've tested something for sturdiness two or three times, and it has worked, I won't even think about testing it again. Consider what evidence you need before you'll happily stand on a particular chair to reach something in the kitchen, or put a heavy television on a stand. Supporting a weight is the kind of thing that either fails the first time, or works fairly reliably. Obviously there could be some strain-induced effects that cause a subsequent failure[9], but John really has a lot of evidence that the bridge will support him.

Given those three answers, it seems to me that it is a reasonable bet to cross the bridge. At the very least, it's no more of an unreasonable bet than the bet I make every day crossing a busy highway by foot. So I'm not surprised that 64% of the subjects agreed that John knew the bridge would hold him. At the very least, that result is perfectly consistent with the variable threshold view, if we make plausible assumptions about how the subjects would answer the three numbered questions above.

---

[8]In the West Gate bridge collapse in Melbourne in 1971, a large number of the victims were underneath the bridge; the people on top of the bridge had a non-trivial chance of survival. That bridge was 200 feet above the water, not 1000, but I'm not sure the extra height would matter greatly. Again from a slightly lower height, over 90% of people on the bridge survived the I-35W collapse in Minneapolis in 2007.

[9]As I believe was the case in the I-35W collapse.

And as I've stressed, these experiments are only a problem for the variable threshold view if the subjects are reliable. I can think of two reasons why they might not be. First, subjects tend to massively discount the costs and likelihoods of traffic related injuries. In most of the country, the risk of death or serious injury through motor vehicle accident is much higher than the risk of death or serious injury through some kind of crime or other attack, yet most people do much less to prevent vehicles harming them than they do to prevent criminals or other attackers harming them.[10] Second, only 73% of this subjects in *this very experiment* said that John knows the bridge will support him in **Low Stakes Bridge**. This is just absurd. Unless the subjects endorse an implausible kind of scepticism, something has gone wrong with the experimental design. Given the fact that the experiment points broadly in the direction of the theory I favour, and that with some plausible assumptions it is perfectly consistent with that theory, and the unreliability of the subjects, I don't think this kind of experimental work threatens the variable threshold view.

April 14, 2010

References

Fantl, Jeremy and McGrath, Matthew. 2009. *Knowledge in an Uncertain World*. Oxford: Oxford University Press.

Feltz, Adam and Zarpentine, Chris. 2010. "Do You Know More When It Matters Less?" *Philosophical Psychology* Retrieved from http://faculty.schreiner.edu/adfeltz/Papers/Know%20more.pdf.

Gettier, Edmund L. 1963. "Is Justified True Belief Knowledge?" *Analysis* 23:121–123. doi:10.2307/3326922.

Ham, Sandra A., Martin, Sarah, and Kohl III, Harold W. 2008. "Changes in the percentage of students who walk or bike to school-United States, 1969 and 2001." *Journal of Physical Activity and Health* 5:205–215.

Harman, Gilbert. 1973. *Thought*. Princeton: Princeton University Press.

Kennett, Jeanette and Smith, Michael. 1996a. "Frog and Toad Lose Control." *Analysis* 56:63–73. doi:10.1111/j.0003-2638.1996.00063.x.

—. 1996b. "Philosophy and Commonsense: The Case of Weakness of Will." In Michaelis Michael and John O'Leary-Hawthorne (eds.), *The Place of Philosophy in the Study of Mind*, 141–157. Norwell, MA: Kluwer.

Lipsey, R. G. and Lancaster, Kelvin. 1956-1957. "The General Theory of Second Best." *Review of Economic Studies* 24:11–32. doi:10.2307/2296233.

---

[10]See the massive drop in the numbers of students walking or biking to school, reported in Ham et al. (2008), for a sense of how big an issue this is.

May, Joshua, Sinnott-Armstrong, Walter, Hull, Jay G., and Zimmerman, Aaron. forthcoming. "Practical Interests, Relevant Alternatives, and Knowledge Attributions: an Empirical Study." *Review of Philosophy and Psychology* doi:10.1007/s13164-009-0014-3.

Russell, Gillian and Doris, John M. 2009. "Knowledge by Indifference." *Australasian Journal of Philosophy* 86:429–437. doi:10.1080/00048400802001996.

Schroeder, Mark. 2008. "Stakes, Withholding and the Subject-Sensitivity of 'Knows'." Retrieved from http://www-rcf.usc.edu/~maschroe/research/Schroeder_Stakes_and_Withholding.pdf.

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

Turri, John. 2010. "On the Relationship between Propositional and Doxastic Justification." *Philosophy and Phenomenological Research* 80:312–326. doi:10.1111/j.1933-1592.2010.00331.x.

Watson, Gary. 1977. "Skepticism about Weakness of Will." *Philosophical Review* 86:316–339. doi:10.2307/2183785.

Weatherson, Brian. 2005. "Can We Do Without Pragmatic Encroachment?" *Philosophical Perspectives* 19:417–443. doi:10.1111/j.1520-8583.2005.00068.x.

Williamson, Timothy. 1998. "Conditionalizing on Knowledge." *British Journal for the Philosophy of Science* 49:89–121. doi:10.1093/bjps/49.1.89.