Can We Do Without Pragmatic Encroachment?

Brian Weatherson

2005

I argue that interests primarily affect the relationship between credence and belief. A view is set out and defended where evidence and rational credence are not interest-relative, but belief, rational belief, and knowledge are.

1 Introduction

Recently several authors have defended claims suggesting that there is a closer connection between practical interests and epistemic justification than has traditionally been countenanced. Jeremy Fantl and Matthew McGrath (2002) argue that there is a "pragmatic necessary condition on epistemic justification" (77), namely the following.

(PC) S is justified in believing that p only if S is rational to prefer as if p. (77)

And John Hawthorne (2004) and Jason Stanley (2005) have argued that what it takes to turn true belief into knowledge is sensitive to the practical environment the subject is in. These authors seem to be suggesting there is, to use Jonathan Kvanvig's phrase "pragmatic encroachment" in epistemology. In this paper I'll argue that their arguments do not quite show this is true, and that concepts of epistemological justification need not be pragmatically sensitive. The aim here isn't to show that (PC) is false, but rather that it shouldn't be described as a pragmatic condition on *justification*. Rather, it is best thought of as a pragmatic condition on *belief*. There are two ways to spell out the view I'm taking here. These are both massive simplifications, but they are close enough to the truth to show the kind of picture I'm aiming for.

First, imagine a philosopher who holds a very simplified version of functionalism about belief, call it (B).

(B) S believes that p iff S prefers as if p

Our philosopher one day starts thinking about justification, and decides that we can get a principle out of (B) by adding normative operators to both sides, inferring (JB).

(JB) S is justified in believing that p only if S is justified to prefer as if p

Now it would be a mistake to treat (JB) as a pragmatic condition on *justification* (rather than belief) if it was derived from (B) by this simple means. And if our philosopher goes on to infer (PC) from (JB), by replacing 'justified' with 'rational', and inferring the conditional from the biconditional, we still don't get a pragmatic condition on *justification*.

Second, Fantl and McGrath focus their efforts on attacking the following

Evidentialism For any two subjects S and S′, necessarily, if S and S′ have the same evidence for/against p, then S is justified in believing that p iff S′ is, too.

I agree, evidentialism is false. And I agree that there are counterexamples to evidentialism from subjects who are in different practical situations. What I don't agree is that we learn much about the role of pragmatic factors in *epistemology* properly defined from these counterexamples to evidentialism. Evidentialism follows from the following three principles.

Probabilistic Evidentialism For any two subjects S and S′, and any degree of belief α necessarily, if S and S′ have the same evidence for/against p, then S is justified in believing that p to degree α iff S′ is, too.

Threshold View For any two subjects S and S′, and any degree of belief α, if S and S′ both believe p to degree α, then S believes that p iff S′ does too.

Probabilistic Justification For any S, S is justified in believing p iff there is some degree of belief α such that S is justified in believing p to degree α, and in S's situation, believing p to degree α suffices for believing p.

(Degrees of belief here are meant to be the subjective correlates of Keynesian probabilities. See Keynes (1921) for more details. They need not, and usually will not, be numerical values. The Threshold View is so-called because given some other plausible premises it implies that S believes that p iff S's degree of belief in p is above a threshold.)

I endorse Probabilistic Justification, and for present purposes at least I endorse Probabilistic Evidentialism. The reason I think Evidentialism fails is because the Threshold View is false. It is plausible that Probabilistic Justification and Probabilistic Evidentialism are epistemological principles, while the Threshold View is a principle from philosophy of mind. So this matches up with the earlier contention that the failure of Evidentialism tells us something interesting about the role of pragmatics in philosophy of mind, rather than something about the role of pragmatics in epistemology.

As noted, Hawthorne and Stanley are both more interested in knowledge than justification. So my discussion of their views will inevitably be somewhat distorting. I think what I say about justification here should carry over to a theory of knowledge, but space prevents a serious examination of that question. The primary bit of 'translation' I have to do to make their works relevant to a discussion of justification is to interpret their defences of the principle (KP) below as implying some support for (JP), which is obviously similar to (PC).

(KP) If S knows that p, then S is justified in using p as a premise in practical reasoning.

(JP) If *S* justifiably believes that *p*, then *S* is justified in using *p* as a premise in practical reasoning.

I think (JP) is just as plausible as (KP). In any case it is independently plausible whether or not Hawthorne and Stanley are committed to it. So I'll credit recognition of (JP)'s importance to a theory of justification to them, and hope that in doing so I'm not irreparably damaging the public record.

The overall plan here is to use some philosophy of mind, specifically functionalist analyses of belief to respond to some arguments in epistemology. But, as you can see from the role the Threshold View plays in the above argument, our starting point will be the question what is the relation between the credences decision theory deals with, and our traditional notion of a belief? I'll offer an analysis of this relation that supports my above claim that we should work with a pragmatic notion of belief rather than a pragmatic notion of justification. The analysis I offer has a hole in it concerning propositions that are not relevant to our current plans, and I'll fix the hold in section 3. Sections 4 and 5 concern the role that closure principles play in my theory, in particular the relationship between having probabilistically coherent degrees of belief and logically coherent beliefs. In this context, a closure principle is a principle that says probabilistic coherence implies logical coherence, at least in a certain domain. (It's called a closure principle because we usually discuss it by working out properties of probabilistically coherent agents, and show that their beliefs are closed under entailment in the relevant domain.) In section 4 I'll defend the theory against the objection, most commonly heard from those wielding the preface paradox, that we need not endorse as strong a closure principle as I do. In section 5 I'll defend the theory against those who would endorse an even stronger closure principle than is defended here. Once we've got a handle on the relationship between degrees of belief and belief *tout court*, we'll use that to examine the arguments for pragmatic encroachment. In section 6 I'll argue that we can explain the intuitions behind the cases that seem to support pragmatic encroachment, while actually keeping all of the pragmatic factors in our theory of belief. In section 7 I'll discuss how to endorse principles like (PC) and (JP) (as far as they can be endorsed) while keeping a non-pragmatic theory of probabilistic justification. The interesting cases here are ones where agents have mistaken and/or irrational beliefs about their practical environment, and intuitions in those cases are cloudy. But it seems the most natural path in these cases is to keep a pragmatically sensitive notion of belief, and a pragmatically insensitive notion of justification.

2 Belief and Degree of Belief

Traditional epistemology deals with beliefs and their justification. Bayesian epistemology deals with degrees of belief and their justification. In some sense they are both talking about the same thing, namely epistemic justification. Two questions naturally arise. Do we really have two subject matters here (degrees of belief and belief *tout court*) or two descriptions of the one subject matter? If just one subject matter, what relationship is

there between the two modes of description of this subject matter?

The answer to the first question is I think rather easy. There is no evidence to believe that the mind contains two representational systems, one to represent things as being probable or improbable and the other to represent things as being true or false. The mind probably does contain a vast plurality of representational systems, but they don't divide up the doxastic duties this way. If there are distinct visual and auditory representational systems, they don't divide up duties between degrees of belief and belief *tout court*, for example. If there were two distinct systems, then we should imagine that they could vary independently, at least as much as is allowed by constitutive rationality. But such variation is hard to fathom. So I'll infer that the one representational system accounts for our credences and our categorical beliefs. (It follows from this that the question Bovens and Hawthorne (1999) ask, namely what beliefs *should* an agent have given her degrees of belief, doesn't have a non-trivial answer. If fixing the degrees of belief in an environment fixes all her doxastic attitudes, as I think it does, then there is no further question of what she should believe given these are her degrees of belief.)

The second question is much harder. It is tempting to say that S believes that p iff S's credence in p is greater than some salient number r, where r is made salient either by the context of belief ascription, or the context that S is in. I'm following Mark Kaplan (1996) in calling this the threshold view. There are two well-known problems with the threshold view, both of which seem fatal to me.

As Robert Stalnaker (1984, 91) emphasised, any number r is bound to seem arbitrary. Unless these numbers are made salient by the environment, there is no special difference between believing p to degree 0.9786 and believing it to degree 0.9875. But if r is 0.98755, this will be *the difference* between believing p and not believing it, which is an important difference. The usual response to this, as found in Foley (1993 Ch. 4) and Hunter (1996) is to say that the boundary is vague. But it's not clear how this helps. On an epistemic theory of vagueness, there is still a number such that degrees of belief above that count, and degrees below that do not, and any such number is bound to seem unimportant. On supervaluational theories, the same is true. There won't be a *determinate* number, to be sure, but there will a number, and that seems false. My preferred degree of belief theory of vagueness, as set out in Weatherson (2005) has the same consequence. Hunter defends a version of the threshold view combined with a theory of vagueness based around fuzzy logic, which seems to be the only theory that could avoid the arbitrariness objection. But as Williamson (1994) showed, there are deep and probably insurmountable difficulties with that position. So I think the vagueness response to the arbitrariness objection is (a) the only prima facie plausible response and (b) unsuccessful.

The second problem concerns conjunction. It is also set out clearly by Stalnaker.

> Reasoning in this way from accepted premises to their deductive consequences (p, also Q, therefore R) does seem perfectly straightforward. Someone may object to one of the premises, or to the validity of the argument, but one could not intelligibly agree that the premises are each

> acceptable and the argument valid, while objecting to the acceptability of
> the conclusion. (Stalnaker 1984, 92)

If categorical belief is having a credence above the threshold, then one can coherently do exactly this. Let x be a number between r and than $r \frac{1}{2}$, such that for an atom of type U has probability x of decaying within a time t, for some t and U. Assume our agent knows this fact, and is faced with two (isolated) atoms of U. Let p be that the first decays within t, and q be that the second decays within t. She should, given her evidence, believe p to degree x, q to degree x, and $p \wedge q$ to degree x^2. If she believed $p \wedge q$ to a degree greater than r, she'd have to either have credences that were not supported by her evidence, or credences that were incoherent. (Or, most likely, both.) So this theory violates the platitude. This is a well-known argument, so there are many responses to it, most of them involving something like appeal to the preface paradox. I'll argue in section 4 that the preface paradox doesn't in fact offer the threshold view proponent much support here. But even before we get to there, we should note that the arbitrariness objection gives us sufficient reason to reject the threshold view.

A better move is to start with the functionalist idea that to believe that p is to treat p as true for the purposes of practical reasoning. To believe p is to have preferences that make sense, by your own lights, in a world where p is true. So, if you prefer A to B and believe that p, you prefer A to B given p. For reasons that will become apparent below, we'll work in this paper with a notion of preference where *conditional* preferences are primary.[1] So the core insight we'll work with is the following:

> If you prefer A to B given q, and you believe that p, then you prefer A to
> B given $p \wedge q$

The bold suggestion here is that if that is true for all the A, B and q that matter, then you believe p. Put formally, where $Bel(p)$ means that the agent believes that p, and A $\geq_q$ B means that the agent thinks A is at least as good as B given q, we have the following

1. $Bel(p) \leftrightarrow \forall A \forall B \forall q \, (A \geq_q B \leftrightarrow A \geq_{p \wedge q} B)$

In words, an agent believes that p iff conditionalising on p doesn't change any conditional preferences over things that matter.[2] The left-to-right direction of this seems trivial, and the right-to-left direction seems to be a plausible way to operationalise the functionalist insight that belief is a functional state. There is some work to be done if (1) is to be interpreted as a truth though.

---

[1] To say the agent prefers A to B given q is not to say that if the agent were to learn q, she would prefer A to B. It's rather to say that she prefers the state of the world where she does A and q is true to the state of the world where she does B and q is true. These two will come apart in cases where learning q changes the agent's preferences. We'll return to this issue below.

[2] This might seem *much* too simple, especially when compared to all the bells and whistles that functionalists usually put in their theories to (further) distinguish themselves from crude versions of behaviourism. The reason we don't need to include those complications here is that they will all be included in the analysis of *preference*. Indeed, the theory here is compatible with a thoroughly anti-functionalist treatment of preference. The claim is not that we can offer a functional analysis of belief in terms of non-mental concepts, just that we can offer a functionalist reduction of belief to other mental concepts. The threshold view is *also* such a reduction, but it is such a crude reduction that it doesn't obviously fall into any category.

If we interpret the quantifiers in (1) as unrestricted, then we get the (false) conclusion that just about no one believes no contingent propositions. To prove this, consider a bet that wins iff the statue in front of me waves back at me due to random quantum effects when I wave at it. If I take the bet and win, I get to live forever in paradise. If I take the bet and lose, I lose a penny. Letting A be that I take the bet, B be that I decline the bet, *q* be a known tautology (so my preferences given *q* are my preferences *tout court*) and *p* be that the statue does not wave back, we have that I prefer A to B, but not A to B given *p*. So by this standard I don't believe that *p*. This is false – right now I believe that statues won't wave back at me when I wave at them.

This seems like a problem. But the solution to it is not to give up on functionalism, but to insist on its pragmatic foundations. The quantifiers in (1) should be restricted, with the restrictions motivated pragmatically. What is crucial to the theory is to say what the restrictions on A and B are, and what the restrictions on *q* are. We'll deal with these in order.

For better or worse, I don't right now have the option taking that bet and hence spending eternity in paradise if the statue waves back at me. Taking or declining such unavailable bets are not open choices. For any option that is open to me, assuming that statues do not in fact wave does not change its utility. That's to say, I've already factored in the non-waving behaviour of statues into my decision-making calculus. That's to say, I believe statues don't wave.

An action A is a live option for the agent if it is really possible for the agent to perform A. An action A is a salient option if it is an option the agent takes seriously in deliberation. Most of the time gambling large sums of money on internet gambling sites over my phone is a live option, but not a salient option. I know this option is suboptimal, and I don't have to recompute every time whether I should do it. Whenever I'm making a decision, I don't have to add in to the list of choices *bet thousands of dollars on internet gambling sites*, and then rerule that out every time. I just don't consider that option, and properly so. If I have a propensity to daydream, then becoming the centre-fielder for the Boston Red Sox might be a salient option to me, but it certainly isn't a live option. We'll say the two initial quantifiers range over the options that are live and salient options for the agent.

Note that we *don't* say that the quantifiers range over the options that are live and salient for the person making the belief ascription. That would lead us to a form of contextualism for which we have little evidence. We also don't say that an option becomes salient for the agent iff they *should* be considering it. At this stage we are just saying what the agent does believe, not what they should believe, so we don't have any clauses involving normative concepts.

Now we'll look at the restrictions on the quantifier over propositions. Say a proposition is *relevant* if the agent is disposed to take seriously the question of whether it is true (whether or not she is currently considering that question) and conditionalising on that proposition or its negation changes some of the agents *unconditional* preferences over live, salient options.[3] The first clause is designed to rule out wild hypotheses that the

---

[3]Conditionalising on the proposition *There are space aliens about to come down and kill all the people writing*

agent does not take at all seriously. If q is not such a proposition, if the agent is disposed to take it seriously, then it is relevant if there are live, salient A and B such that A $\geq_q$ B $\leftrightarrow$ A $\geq$ B is false. Say a proposition is *salient* if the agent is currently considering whether it is true. Finally, say a proposition is *active* relative to p iff it is a (possibly degenerate) conjunction of propositions such that each conjunct is either relevant or salient, and such that the conjunction is consistent with p. (By a degenerate conjunction I mean a conjunction with just one conjunct. The consistency requirement is there because it might be hard in some cases to make sense of preferences given inconsistencies.) Then the propositional quantifier in (1) ranges over active propositions.

We will expand and clarify this in the next section, but our current solution to the relationship between beliefs and degrees of belief is that degrees of belief determine an agent's preferences, and she believes that p iff the claim (1) about her preferences is true when the quantifiers over options are restricted to live, salient actions, and the quantifier over propositions is restricted to salient propositions. The simple view would be to say that the agent believes that p iff conditioning on p changes none of her preferences. The more complicated view here is that the agent believes that p iff conditioning on p changes none of her conditional preferences over live, salient options, where the conditions are also active relative to p.

3 Impractical Propositions

The theory sketched in the previous paragraph seems to me right in the vast majority of cases. It fits in well with a broadly functionalist view of the mind, and as we'll see it handles some otherwise difficult cases with aplomb. But it needs to be supplemented a little to handle beliefs about propositions that are practically irrelevant. I'll illustrate the problem, then note how I prefer to solve it.

I don't know what Julius Caeser had for breakfast the morning he crossed the Rubicon. But I think he would have had *some* breakfast. It is hard to be a good general without a good morning meal after all. Let p be the proposition that he had breakfast that morning. I believe p. But this makes remarkably little difference to my practical choices in most situations. True, I wouldn't have written this paragraph as I did without this belief, but it is rare that I have to write about Caeser's dietary habits. In general whether p is true makes no practical difference to me. This makes it hard to give a pragmatic account of whether I believe that p. Let's apply (1) to see whether I really believe that p.

1. *Bel*(p) $\leftrightarrow$ $\forall$A$\forall$B$\forall q$ (A $\geq_q$ B $\leftrightarrow$ A $\geq_{p \wedge q}$ B)

Since p makes no practical difference to any choice I have to make, the right hand side is true. So the left hand side is true, as desired. The problem is that the right hand side of (2) is also true here.

---

epistemology papers will make me prefer to stop writing this paper, and perhaps grab some old metaphysics papers I could be working on. So that proposition satisfies the second clause of the definition of relevance. But it clearly doesn't satisfy the first clause. This part of the definition of relevance won't do much work until the discussion of agents with mistaken environmental beliefs in section 7.

2. $Bel(\neg p) \leftrightarrow \forall A \forall B \forall q\ (A \geq_q B \leftrightarrow A \geq_{\neg p \wedge q} B)$

Adding the assumption that Caeser had no breakfast that morning doesn't change any of my practical choices either. So I now seem to *inconsistently* believe both p and $\neg p$. I have some inconsistent beliefs, I'm sure, but those aren't among them. We need to clarify what (1) claims.

To do so, I supplement the theory sketched in section 2 with the following principles.

- A proposition p is *eligible for belief* if it satisfies $\forall A \forall B \forall q\ (A \geq_q B \leftrightarrow A \geq_{p \wedge q} B)$, where the first two quantifiers range over the open, salient actions in the sense described in section 2.
- For any proposition p, and any proposition q that is relevant or salient, among the actions that are (by stipulation!) open and salient with respect to p are *believing that p*, *believing that q*, *not believing that p* and *not believing that q*
- For any proposition, the subject prefers believing it to not believing it iff (a) it is eligible for belief and (b) the agents degree of belief in the proposition is greater than $\frac{1}{2}$.
- The previous stipulation holds both unconditionally and conditional on p, for any p.
- The agent believes that p iff $\forall A \forall B \forall q\ (A \geq_q B \leftrightarrow A \geq_{p \wedge q} B)$, where the first two quantifiers range over all actions that are either open and salient *tout court* (i.e. in the sense of section 2) or open and salient with respect to p (as described above).

This all looks moderately complicated, but I'll explain how it works in some detail as we go along. One simple consequence is that an agent only believes that p iff their degree of belief in p is greater than $\frac{1}{2}$. Since my degree of belief in Caeser's foodless morning is not greater than $\frac{1}{2}$, in fact it is considerably less, I don't believe $\neg p$. On the other hand, since my degree of belief in p is considerably greater than $\frac{1}{2}$, I prefer to believe it than disbelieve it, so I believe it.

There are many possible objections to this position, which I'll address sequentially.

Objection: Even if I have a high degree of belief in p, I might prefer to not believe p because I think that belief in p is bad for some other reason. Perhaps, if p is a proposition about my brilliance, it might be immodest to believe that p.

Reply: Any of these kinds of considerations should be put into the credences. If it is immodest to believe that you are a great philosopher, it is equally immodest to believe to a high degree that you are a great philosopher.

Objection: Belief that p is not an action in the ordinary sense of the term.

Reply: True, which is why this is described as a supplement to the original theory, rather than just cashing out its consequences.

Objection: It is impossible to choose to believe or not believe something, so we shouldn't be applying these kinds of criteria.

Reply: I'm not as convinced of the impossibility of belief by choice as others are, but I won't push that for present purposes. Let's grant that beliefs are always involuntary. So these 'actions' aren't open actions in any interesting sense, and the theory is section 2 was really incomplete. As I said, this is a supplement to the theory in section 2.

This doesn't prevent us using principles of constitutive rationality, such as we prefer to believe p iff our credence in p is over $\frac{1}{2}$. Indeed, on most occasions where we use constitutive rationality to infer that a person has some mental state, the mental state we attribute to them is one they could not fail to have. But functionalists are committed to constitutive rationality (Lewis 1994). So my approach here is consistent with a broadly functionalist outlook.

Objection: This just looks like a roundabout way of stipulating that to believe that p, your degree of belief in p has to be greater than $\frac{1}{2}$. Why not just add that as an extra clause than going through these little understood detours about preferences about beliefs?

Reply: There are three reasons for doing things this way rather than adding such a clause.

First, it's nice to have a systematic theory rather than a theory with an ad hoc clause like that.

Second, the effect of this constraint is much more than to restrict belief to propositions whose credence is greater than $\frac{1}{2}$. Consider a case where p and q and their conjunction are all salient, p and q are probabilistically independent, and the agent's credence in each is 0.7. Assume also that p, q and $p \wedge q$ are completely irrelevant to any practical deliberation the agent must make. Then the criteria above imply that the agent does not believe that p or that q. The reason is that the agent's credence in $p \wedge q$ is 0.49, so she prefers to not believe $p \wedge q$. But conditional on p, her credence in $p \wedge q$ is 0.7, so she prefers to believe it. So conditionalising on p does change her preferences with respect to believing $p \wedge q$, so she doesn't believe p. So the effect of these stipulations rules out much more than just belief in propositions whose credence is below $\frac{1}{2}$.

This suggests the third, and most important point. The problem with the threshold view was that it led to violations of closure. Given the theory as stated, we can prove the following theorem. Whenever p and q and their conjunction are all open or salient, and both are believed, and the agent is probabilistically coherent, the agent also believes $p \wedge q$. This is a quite restricted closure principle, but this is no reason to deny that it is *true*, as it fails to be true on the threshold view.

The proof of this theorem is a little complicated, but worth working through. First we'll prove that if the agent believes p, believes q, and p and q are both salient, then the agent prefers believing $p \wedge q$ to not believing it, if $p \wedge q$ is eligible for belief. In what follows $Pr(x|y)$ is the agent's conditional degree of belief in x given y. Since the agent is coherent, we'll assume this is a probability function (hence the name).

1. Since the agent believes that q, they prefer believing that q to not believing that q (by the criteria for belief).
2. So the agent prefers believing that q to not believing that q given p (From 1 and the fact that they believe that p, and that q is salient).
3. So $Pr(q|p) > \frac{1}{2}$ (from 2)
4. $Pr(q|p) = Pr(p \wedge q|p)$ (by probability calculus)
5. So $Pr(p \wedge q|p) > \frac{1}{2}$ (from 3, 4)
6. So, if $p \wedge q$ is eligible for belief, then the agent prefers believing that $p \wedge q$ to not believing it, given p (from 5)

7. So, if $p \wedge q$ is eligible for belief, the agent prefers believing that $p \wedge q$ to not believing it (from 6, and the fact that they believe that p, and $p \wedge q$ is salient)

So whenever, p, q and $p \wedge q$ are salient, and the agent believes each conjunct, the agent prefers believing the conjunction $p \wedge q$ to not believing it, if $p \wedge q$ is eligible. Now we have to prove that $p \wedge q$ is eligible for belief, to prove that it is actually believed. That is, we have to prove that (5) follows from (4) and (3), where the initial quantifiers range over actions that are open and salient *tout court*.

1. $\forall A \forall B \forall r \, (A \geq_r B \leftrightarrow A \geq_p \wedge r \, B)$
2. $\forall A \forall B \forall r \, (A \geq_r B \leftrightarrow A \geq_q \wedge r \, B)$
3. $\forall A \forall B \forall r \, (A \geq_r B \leftrightarrow A \geq_{p \wedge q \wedge r} B)$

Assume that (5) isn't true. That is, there are A, B and S such that $\neg(A \geq_s B \leftrightarrow A \geq_{p \wedge q \wedge s} B)$. By hypothesis S is active, and consistent with $p \wedge q$. So it is the conjunction of relevant, salient propositions. Since q is salient, this means $q \wedge s$ is also active. Since S is consistent with $p \wedge q$, it follows that $q \wedge s$ is consistent with p. So $q \wedge s$ is a possible substitution instance for r in (3). Since (3) is true, it follows that $A \geq_{q \wedge s} B \leftrightarrow A \geq_{p \wedge q \wedge s} B$. By similar reasoning, it follows that s is a permissible substitution instance in (4), giving us $A \geq_s B \leftrightarrow A \geq_{q \wedge s} B$. Putting the last two biconditionals together we get $A \geq_s B \leftrightarrow A \geq_{p \wedge q \wedge s} B$, contradicting our hypothesis that there is a counterexample to (5). So whenever (3) and (4) are true, (5) is true as well, assuming p, q and $p \wedge q$ are all salient.

4 Defending Closure

So on my account of the connection between degrees of belief and belief *tout court*, probabilistic coherence implies logical coherence amongst salient propositions. The last qualification is necessary. It is possible for a probabilistically coherent agent to not believe the *non*-salient consequences of things they believe, and even for a probabilistically coherent agent to have inconsistent beliefs as long as not all the members of the inconsistent set are active. Some people argue that even this weak a closure principle is implausible. David Christensen (2005), for example, argues that the preface paradox provides a reason for doubting that beliefs must be closed under entailment, or even must be consistent. Here is his description of the case.

> We are to suppose that an apparently rational person has written a long non-fiction book—say, on history. The body of the book, as is typical, contains a large number of assertions. The author is highly confident in each of these assertions; moreover, she has no hesitation in making them unqualifiedly, and would describe herself (and be described by others) as believing each of the book's many claims. But she knows enough about the difficulties of historical scholarship to realize that it is almost inevitable that at least a few of the claims she makes in the book are mistaken. She modestly acknowledges this in her preface, by saying that she believes the

> book will be found to contain some errors, and she graciously invites those
> who discover the errors to set her straight. (Christensen 2005, 33–34)

Christensen thinks such an author might be rational in every one of her beliefs, even though these are all inconsistent. Although he does not say this, nothing in his discussion suggests that he is using the irrelevance of some of the propositions in the author's defence. So here is an argument that we should abandon closure amongst relevant beliefs.

Christensen's discussion, like other discussions of the preface paradox, makes frequent use of the fact that examples like these are quite common. We don't have to go to fake barn country to find a counterexample to closure. But it seems to me that we need two quite strong idealisations in order to get a real counterexample here.

The first of these is discussed in forthcoming work by Ishani Maitra (2010), and is briefly mentioned by Christensen in setting out the problem. We only have a counterexample to closure if the author *believes* every thing she writes in her book. (Indeed, we only have a counterexample if she reasonably believes every one of them. But we'll assume a rational author who only believes what she ought to believe.) This seems unlikely to be true to me. An author of a historical book is like a detective who, when asked to put forward her best guess about what explains the evidence, says "If I had to guess, I'd say ..." and then launches into spelling out her hypothesis. It seems clear that she need not *believe* the truth of her hypothesis. If she did that, she could not later learn it was true, because you can't learn the truth of something you already believe. And she wouldn't put any effort into investigating alternative suspects. But she can come to learn her hypothesis was true, and it would be rational to investigate other suspects. It seems to me (following here Maitra's discussion) that we should understand scholarly assertions as being governed by the same kind of rules that govern detectives making the kind of speech being contemplated here. And those rules don't require that the speaker believe the things they say without qualification. The picture is that the little prelude the detective explicitly says is implicit in all scholarly work.

There are three objections I know to this picture, none of them particularly conclusive. First, Christensen says that the author doesn't qualify their assertions. But neither does our detective qualify most individual sentences. Second, Christensen says that most people would describe our author as believing her assertions. But it is also natural to describe our detective as believing the things she says in her speech. It's natural to say things like "She thinks it was the butler, with the lead pipe," in reporting her hypothesis. Third, Timothy Williamson (2000) has argued that if speakers don't believe what they say, we won't have an explanation of why Moore's paradoxical sentences, like "The butler did it, but I don't believe the butler did it," are always defective. Whatever the explanation of the paradoxicality of these sentences might be, the alleged requirement that speakers believe what they say can't be it. For our detective cannot properly say "The butler did it, but I don't believe the butler did it" in setting out her hypothesis, even though *believing* the butler did it is not necessary for her to say "The butler did it" in setting out just that hypothesis.

It is plausible that for *some* kinds of books, the author should only say things they

believe. This is probably true for travel guides, for example. Interestingly, casual observation suggests that authors of such books are much less likely to write modest prefaces. This makes some sense if those books can only include statements their authors believe, and the authors believe the conjunctions of what they believe.

The second idealisation is stressed by Simon Evnine (1999) in his paper "Believing Conjunctions". The following situation does not involve me believing anything inconsistent.

- I believe that what Manny just said, whatever it was, is false.
- Manny just said that the stands at Fenway Park are green.
- I believe that the stands at Fenway Park are green.

If we read the first claim *de dicto*, that I believe that Manny just said something false, then there is no inconsistency. (Unless I also believe that what Manny just said was that the stands in Fenway Park are green.) But if we read it *de re*, that the thing Manny just said is one of the things I believe to be false, then the situation does involve me being inconsistent. The same is true when the author believes that one of the things she says in her book is mistaken. If we understand what she says *de dicto*, there is no contradiction in her beliefs. It has to be understood *de re* before we get a logical problem. And the fact is that most authors do not have *de re* attitudes towards the claims made in their book. Most authors don't even remember everything that's in their books. (I'm not sure I remember how this section started, let alone this paper.) Some may argue that authors don't even have the capacity to consider a proposition as long and complicated as the conjunction of all the claims in their book. Christensen considers this objection, but says it isn't a serious problem.

> It is undoubtedly true that ordinary humans cannot entertain book-length conjunctions. But surely, agents who do not share this fairly *superficial* limitation are easily conceived. And it seems just as wrong to say of such agents that they are rationally required to believe in the inerrancy of the books they write. (38: my emphasis)

I'm not sure this is undoubtedly true; it isn't clear that propositions (as opposed to their representations) have lengths. And humans can believe propositions that *can* be represented by sentences as long as books. But even without that point, Christensen is right that there is an idealisation here, since ordinary humans do not know exactly what is in a given book, and hence don't have *de re* attitudes towards the propositions expressed in the book.

I'm actually rather suspicious of the intuition that Christensen is pushing here, that idealising in this way doesn't change intuitions about the case. The preface paradox gets a lot of its (apparent) force from intuitions about what attitude we should have towards real books. Once we make it clear that the real life cases are not relevant to the paradox, I find the intuitions become rather murky. But I won't press this point.

A more important point is that we believers in closure don't think that authors should think their books are inerrant. Rather, following Stalnaker (1984), we think

that authors shouldn't unqualifiedly *believe* the individual statements in their book if they don't believe the conjunction of those statements. Rather, their attitude towards those propositions (or at least some of them) should be that they are probably true. (As Stalnaker puts it, they accept the story without believing it.) Proponents of the preface paradox know that this is a possible response, and tend to argue that it is impractical. Here is Christensen on this point.

> It is clear that our everyday binary way of talking about beliefs has im-
> mense practical advantages over a system which insisted on some more
> fine-grained reporting of degrees of confidence … At a minimum, talking
> about people as believing, disbelieving, or withholding belief has at least
> as much point as do many of the imprecise ways we have of talking about
> things that can be described more precisely. (96)

Richard Foley makes a similar point.

> There are *deep* reasons for wanting an epistemology of beliefs, reasons that
> epistemologies of degrees of belief by their very nature cannot possibly
> accommodate. (Foley 1993, 170, my emphasis)

It's easy to make too much of this point. It's a lot easier to triage propositions into TRUE, FALSE and NOT SURE and work with those categories than it is to work assign precise numerical probabilities to each proposition. But these are not the only options. Foley's discussion subsequent to the above quote sometimes suggests they are, especially when he contrasts the triage with "indicat[ing] as accurately as I can my degree of confidence in each assertion that I defend." (171) But really it isn't *much* harder to add two more categories, PROBABLY TRUE and PROBABLY FALSE to those three, and work with that five-way division rather than a three-way division. It's not clear that humans as they are actually constructed have a *strong* preference for the three-way over the five-way division, and even if they do, I'm not sure in what sense this is a 'deep' fact about them.

Once we have the five-way division, it is clear what authors should do if they want to respect closure. For any conjunction that they don't believe (i.e. classify as true), they should not believe one of the conjuncts. But of course they can classify every conjunct as probably true, even if they think the conjunction is false, or even certainly false. Still, might it not be considered something of an idealisation to say rational authors must make this five-way distinction amongst propositions they consider? Yes, but it's no more of an idealisation than we need to set up the preface paradox in the first place. To use the preface paradox to find an example of someone who reasonably violates closure, we need to insist on the following three constraints.

1. They are part of a research community where only asserting propositions you believe is compatible with active scholarship;
2. They know exactly what is in their book, so they are able to believe that one of the propositions in the book is mistaken, where this is understood *de re*; but

3. They are unable to effectively function if they have to effect a five-way, rather than a three-way, division amongst the propositions they consider.

Put more graphically, to motivate the preface paradox we have to think that our inability to have *de re* thoughts about the contents of books is a "superficial constraint", but our preference for working with a three-way rather than a five-way division is a "deep" fact about our cognitive system. Maybe each of these attitudes could be plausible taken on its own (though I'm sceptical of that) but the conjunction seems just absurd.

I'm not entirely sure an agent subject to exactly these constraints is even fully conceivable. (Such an agent is negatively conceivable, in David Chalmers's terminology, but I rather doubt they are positively conceivable.) But even if they are a genuine possibility, why the norms applicable to an agent satisfying that very gerrymandered set of constraints should be considered relevant norms for our state is far from clear. I'd go so far as to say it's clear that the applicability (or otherwise) of a given norm to such an odd agent is no reason whatsoever to say it applies to us. But since the preface paradox only provides a reason for just these kinds of agents to violate closure, we have no reason for ordinary humans to violate closure. So I see no reason here to say that we can have probabilistic coherence without logical coherence, as proponents of the threshold view insist we can have, but which I say we can't have *at least when the propositions involved are salient*. The more pressing question, given the failure of the preface paradox argument, is why I don't endorse a much stronger closure principle, one that drops the restriction to salient propositions. The next section will discuss that point.

I've used Christensen's book as a stalking horse in this section, because it is the clearest and best statement of the preface paradox. Since Christensen is a paradox-mongerer and I'm a paradox-denier, it might be thought we have a deep disagreement about the relevant epistemological issues. But actually I think our overall views are fairly close despite this. I favour an epistemological outlook I call "Probability First", the view that getting the epistemology of partial belief right is of the first importance, and everything else should flow from that. Christensen's view, reduced to a slogan, is "Probability First and Last". This section has been basically about the difference between those two slogans. It's an important dispute, but it's worth bearing in mind that it's a factional squabble within the Probability Party, not an outbreak of partisan warfare.

5 Too Little Closure?

In the previous section I defended the view that a coherent agent has beliefs that are deductively cogent with respect to salient propositions. Here I want to defend the importance of the qualification. Let's start with what I take to be the most important argument for closure, the passage from Stalnaker's *Inquiry* that I quoted above.

> Reasoning in this way from accepted premises to their deductive consequences (p, also Q, therefore R) does seem perfectly straightforward. Someone may object to one of the premises, or to the validity of the argument, but one could not intelligibly agree that the premises are each

> acceptable and the argument valid, while objecting to the acceptability of
> the conclusion. (Stalnaker 1984, 92)

Stalnaker's wording here is typically careful. The relevant question isn't whether we can accept *p*, accept *q*, accept *p* and *q* entail *r*, and reject *r*. As Christensen (2005 Ch. 4) notes, this is impossible even on the threshold view, as long as the threshold is above 2/3. The real question is whether we can accept *p*, accept *q*, accept *p* and *q* entail *r*, and *fail* to accept *r*. And this is always a live possibility on any threshold view, though it seems absurd at first that this could be coherent.

But it's important to note how *active* the verbs in Stalnaker's description are. When faced with a valid argument we have to *object* to one of the premises, or the validity of the argument. What we can't do is *agree* to the premises and the validity of the argument, while *objecting* to the conclusion. I agree. If we are really *agreeing* to some propositions, and *objecting* to others, then all those propositions are salient. And in that case closure, deductive coherence, is mandatory. This doesn't tell us what we have to do if we haven't previously made the propositions salient in the first place.

The position I endorse here is very similar in its conclusions to that endorsed by Gilbert Harman in *Change in View*. There Harman endorses the following principle. (At least he endorses it as true – he doesn't seem to think it is particularly explanatory because it is a special case of a more general interesting principle.)

Recognized Logical Implication Principle   One has reason to believe *p* if one *recognizes* that *p* is logically implied by one's view. (Harman 1986, 17)

This seems right to me, both what it says and its implicature that the reason in question is not a conclusive reason. My main objection to those who use the preface paradox to argue against closure is that they give us a mistaken picture of what we have *to do* epistemically. When I have inconsistent beliefs, or I don't believe some consequence of my beliefs, that is something I have a reason to deal with at some stage, something I have to do. When we say that we have things to do, we don't mean that we have to do them *right now*, or instead of everything else. My current list of things to do includes cleaning my bathroom, yet here I am writing this paper, and (given the relevant deadlines) rightly so. We can have the job of cleaning up our epistemic house as something to do while recognising that we can quite rightly do other things first. But it's a serious mistake to infer from the permissibility of doing other things that cleaning up our epistemic house (or our bathroom) isn't something to be done. The bathroom won't clean itself after all, and eventually this becomes a problem.

There is a possible complication when it comes to tasks that are very low priority. My attic is to be cleaned, or at least it could be cleaner, but there are no imaginable circumstances under which something else wouldn't be higher priority. Given that, should we really leave *clean the attic* on the list of things to be done? Similarly, there might be implications I haven't followed through that it couldn't possibly be worth my time to sort out. Are they things to be done? I think it's worthwhile recording them as such, because otherwise we might miss opportunities to deal with them in the

process of doing something else. I don't need to put off anything else in order to clean the attic, but if I'm up there for independent reasons I should bring down some of the garbage. Similarly, I don't need to follow through implications mostly irrelevant to my interests, but if those propositions come up for independent reasons, I should deal with the fact that some things I believe imply something I don't believe. Having it be the case that all implications from things we believe to things we don't believe constitute jobs to do (possibly in the loose sense that cleaning my attic is something to do) has the right implications for what epistemic duties we do and don't have.

While waxing metaphorical, it seems time to pull out a rather helpful metaphor that Gilbert Ryle (1949) develops in *The Concept of Mind* at a point where he's covering what we'd now call the inference/implication distinction. (This is a large theme of chapter 9, see particularly pages 292-309.) Ryle's point in these passages, as it frequently is throughout the book, is to stress that minds are fundamentally active, and the activity of a mind cannot be easily recovered from its end state. Although Ryle doesn't use this language, his point is that we shouldn't confuse the difficult activity of drawing inferences with the smoothness and precision of a logical implication. The language Ryle does use is more picturesque. He compares the easy work a farmer does when sauntering down a path from the hard work he did when building the path. A good argument, in philosophy or mathematics or elsewhere, is like a well made path that permits sauntering from the start to finish without undue strain. But from that it doesn't follow that the task of coming up with that argument, of building that path in Ryle's metaphor, was easy work. The easiest paths to walk are often the hardest to build. Path-building, smoothing out our beliefs so they are consistent and closed under implication, is hard work, even when the finished results look clean and straightforward. Its work that we shouldn't do unless we need to. But making sure our beliefs are closed under entailment even with respect to irrelevant propositions is suspiciously like the activity of buildings paths between points without first checking you need to walk between them.

For a less metaphorical reason for doubting the wisdom of this unchecked commitment to closure, we might notice the difficulties theorists tend to get into all sorts of difficulties. Consider, for example, the view put forward by Mark Kaplan in *Decision Theory as Philosophy*. Here is his definition of belief.

> You count as believing P just if, were your sole aim to assert the truth (as it pertains to P), and you only options were to assert that P, assert that ¬P or make neither assertion, you would prefer to assert that P. (109)

Kaplan notes that conditional definitions like this are prone to Shope's conditional fallacy. If my sole aim were to assert the truth, I might have different beliefs to what I now have. He addresses one version of this objection (namely that it appears to imply that everyone believes their sole desire is to assert the truth) but as we'll see presently he can't avoid all versions of it.

These arguments are making me thirsty. I'd like a beer. Or at least I think I would. But wait! On Kaplan's theory I can't think that I'd like a beer, for if my sole aim were to assert the truth as it pertains to my beer-desires, I wouldn't have beer desires. And

then I'd prefer to assert that I wouldn't like a beer, I'd merely like to assert the truth as it pertains to my beer desires.

Even bracketing this concern, Kaplan ends up being committed to the view that I can (coherently!) believe that *p* even while regarding *p* as highly improbable. This looks like a refutation of the view to me, but Kaplan accepts it with some equanimity. He has two primary reasons for saying we should live with this. First, he says that it only looks like an absurd consequence if we are committed to the Threshold View. To this all I can say is that *I* don't believe the Threshold View, but it still seems absurd to me. Second, he says that any view is going to have to be revisionary to some extent, because our ordinary concept of belief is not "coherent" (142). His view is that, "Our ordinary notion of belief both construes belief as a state of confidence short of certainty and takes consistency of belief to be something that is at least possible and, perhaps, even desirable" and this is impossible. I think the view here interprets belief as a state less than confidence and allows for as much consistency as the folk view does (i.e. consistency amongst salient propositions), so this defence is unsuccessful as well.

None of the arguments here in favour of our restrictions on closure are completely conclusive. In part the argument at this stage rests on the lack of a plausible rival theory that doesn't interpret belief as certainty but implements a stronger closure principle. It's possible that tomorrow someone will come up with a theory that does just this. Until then, we'll stick with the account here, and see what its epistemological implications might be.

6 Examples of Pragmatic Encroachment

Fantl and McGrath's case for pragmatic encroachment starts with cases like the following. (The following case is not quite theirs, but is similar enough to suit their plan, and easier to explain in my framework.)

> *Local and Express*
>
> There are two kinds of trains that run from the city to the suburbs: the local, which stops at all stations, and the express, which skips the first eight stations. Harry and Louise want to go to the fifth station, so they shouldn't catch the Express. Though if they do it isn't too hard to catch a local back the other way, so it isn't usually a large cost. Unfortunately, the trains are not always clearly labelled. They see a particular train about to leave. If it's a local they are better off catching it, if it is an express they should wait for the next local, which they can see is already boarding passengers and will leave in a few minutes. While running towards the train, they hear a fellow passenger say "It's a local." This gives them good, but far from overwhelming, reason to believe that the train is a local. Passengers get this kind of thing wrong fairly frequently, but they don't have time to get more information. So each of them face a gamble, which they can take by getting on the train. If the train is a local, they will get home a few

minutes early. If it is an express they will get home a few minutes later. For Louise, this is a low stakes gamble, as nothing much turns on whether she is a few minutes early or late, but she does have a weak preference for arriving earlier rather than later. But for Harry it is a high stakes gamble, because if he is late he won't make the start of his daughter's soccer game, which will highly upset her. There is no large payoff for Harry arriving early.

What should each of them do? What should each of them believe?

The first question is relatively easy. Louise should catch the train, and Harry should wait for the next. For each of them that's the utility maximising thing to do. The second one is harder. Fantl and McGrath suggest that, despite being in the same epistemic position with respect to everything except their interests, Louise is justified in believing the train is a local and Harry is not. I agree. (If you don't think the particular case fits this pattern, feel free to modify it so the difference in interests grounds a difference in what they are justified in believing.) Does this show that our notion of epistemic justification has to be pragmatically sensitive? I'll argue that it does not.

The fundamental assumption I'm making is that what is primarily subject to epistemic evaluation are degrees of belief, or what are more commonly called states of confidence in ordinary language. When we think about things this way, we see that Louise and Harry are justified in adopting *the very same degrees of belief*. Both of them should be confident, but not absolutely certain, that the train is a local. We don't have even the appearance of a counterxample to Probabilistic Evidentialism here. If we like putting this in numerical terms, we could say that each of them is justified in assigning a probability of around 0.9 to the proposition *That train is a local*.[4] So as long as we adopt a Probability First epistemology, where we in the first instance evaluate the probabilities that agents assign to propositions, Harry and Louise are evaluated alike iff they do the same thing.

How then can we say that Louise alone is justified in believing that the train is a local? Because that state of confidence they are justified in adopting, the state of being fairly confident but not absolutely certain that the train is a local, counts as believing that the train is a local given Louise's context but not Harry's context. Once Louise hears the other passenger's comment, conditionalising on *That's a local* doesn't change any of her preferences over open, salient actions, including such 'actions' as believing or disbelieving propositions. But conditional on the train being a local, Harry prefers catching the train, which he actually does not prefer.

In cases like this, interests matter not because they affect the degree of confidence that an agent can reasonably have in a proposition's truth. (That is, not because they matter to epistemology.) Rather, interests matter because they affect whether those reasonable

---

[4]I think putting things numerically is misleading because it suggests that the kind of bets we usually use to measure degrees of belief are open, salient options for Louise and Harry. But if those bets were open and salient, they wouldn't *believe* the train is a local. Using qualitative rather than quantitative language to describe them is just as accurate, and doesn't have misleading implications about their practical environment.

degrees of confidence amount to belief. (That is, because they matter to philosophy of mind.) There is no reason here to let pragmatic concerns into epistemology.

7 Justification and Practical Reasoning

The discussion in the last section obviously didn't show that there is no encroachment of pragmatics into epistemology. There are, in particular, two kinds of concerns one might have about the prospects for extending my style of argument to block all attempts at pragmatic encroachment. The biggest concern is that it might turn out to be impossible to defend a Probability First epistemology, particularly if we do not allow ourselves pragmatic concerns. For instance, it is crucial to this project that we have a notion of evidence that is not defined in terms of traditional epistemic concepts (e.g. as knowledge), or in terms of interests. This is an enormous project, and I'm not going to attempt to tackle it here. The second concern is that we won't be able to generalise the discussion of that example to explain the plausibility of (JP) without conceding something to the defenders of pragmatic encroachment.

(JP) If S justifiably believes that p, then S is justified in using p as a premise in practical reasoning.

And that's what we will look at in this section. To start, we need to clarify exactly what (JP) means. Much of this discussion will be indebted to Fantl and McGrath's discussion of various ways of making (JP) more precise. To see some of the complications at issue, consider a simple case of a bet on a reasonably well established historical proposition. The agent has a lot of evidence that supports p, and is offered a bet that returns \$1 if p is true, and loses \$500 if p is false. Since her evidence doesn't support *that* much confidence in p, she properly declines the bet. One might try to reason intuitively as follows. Assume that she justifiably believed that p. Then she'd be in a position to make the following argument.

> p
> If p, then I should take the bet
> So, I should take the bet

Since she isn't in a position to draw the conclusion, she must not be in a position to endorse both of the premises. Hence (arguably) she isn't justified in believing that p. But we have to be careful here. If we assume also that p is true (as Fantl and McGrath do, because they are mostly concerned with knowledge rather than justified belief), then the second premise is clearly false, since it is a conditional with a true antecedent and a false consequent. So the fact that she can't draw the conclusion of this argument only shows that she can't endorse *both* of the premises, and that's not surprising since one of the premises is most likely false. (I'm not assuming here that the conditional is true iff it has a true antecendent or a false consequent, just that it is only true if it has a false antecendent or a true consequent.)

In order to get around this problem, Fantl and McGrath suggest a few other ways that our agent might reason to the bet. They suggest each of the following principles.

> S knows that p only if, for any act A, if *S* knows that if p, then A is the best thing she can do, then *S* is rational to do A. (72)

> *S* knows that p only if, for any states of affairs A and B, if *S* knows that if p, then A is better for her than B, then *S* is rational to prefer A to B. (74)

> **(PC)** *S* is justified in believing that p only if *S* is rational to prefer as if p. (77)

Hawthorne (2004, 174–81) appears to endorse the second of these principles. He considers an agent who endorses the following implication concerning a proposed sell of a lottery ticket for a cent, which is well below its actuarially fair value.

> I will lose the lottery.
> If I keep the ticket, I will get nothing.
> If I sell the ticket, I will get a cent.
> So I ought to sell the ticket. (174)

(To make this fully explicit, it helps to add the tacit premise that a cent is better than nothing.) Hawthorne says that this is intuitively a *bad* argument, and concludes that the agent who attempts to use it is not in a position to know its first premise. But that conclusion only follows if we assume that the argument form is acceptable. So it is plausible to conclude that he endorses Fantl and McGrath's second principle.

The interesting question here is whether the theory endorsed in this paper can validate the true principles that Fantl and McGrath articulate. (Or, more precisely, we can validate the equivalent true principles concerning justified belief, since knowledge is outside the scope of the paper.) I'll argue that it can in the following way. First, I'll just note that given the fact that the theory here implies the closure principles we outlined in section 5, we can easily enough endorse Fantl and McGrath's first two principles. This is good, since they seem true. The longer part of the argument involves arguing that their principle (PC), which doesn't hold on the theory endorsed here, is in fact incorrect.

One might worry that the qualification on the closure principles in section 5 mean that we can't fully endorse the principles Fantl and McGrath endorse. In particular, it might be worried that there could be an agent who believes that *p*, believes that if *p*, then A is better than B, but doesn't put these two beliefs together to infer that A is better than B. This is certainly a possibility given the qualifications listed above. But note that in this position, if those two beliefs were justified, the agent would certainly be *rational* to conclude that A is better than B, and hence rational to prefer A to B. So the constraints on the closure principles don't affect our ability to endorse these two principles.

The real issue is (PC). Fantl and McGrath offer a lot of cases where (PC) holds, as well as arguing that it is plausibly true given the role of implications in practical reasoning. What's at issue is that (PC) is stronger than a deductive closure principle. It is, in effect, equivalent to endorsing the following schema as a valid principle of implication.

p
Given *p*, A is preferable to B
So, A is preferable to B

I call this Practical Modus Ponens, or PMP. The middle premise in PMP is *not* a conditional. It is not to be read as *If p, then A is preferable to B*. Conditional valuations are not conditionals. To see this, again consider the proposed bet on (true) *p* at exorbitant odds, where A is the act of taking the bet, and B the act of declining the bet. It's true that given *p*, A is preferable to B. But it's not true that if *p*, then A is preferable to B. Even if we restrict our attention to cases where the preferences in question are perfectly valid, this is a case where PMP is invalid. Both premises are true, and the conclusion is false. It might nevertheless be true that whenever an agent is justified in believing both of the premises, she is justified in believing the conclusion. To argue against this, we need a *very* complicated case, involving embedded bets and three separate agents, Quentin, Robby and Thom. All of them have received the same evidence, and all of them are faced with the same complex bet, with the following properties.

- *p* is an historical proposition that is well (but not conclusively) supported by their evidence, and happens to be true. All the agents have a high credence in *p*, which is exactly what the evidence supports.
- The bet A, which they are offered, wins if *p* is true, and loses if *p* is false.
- If they win the bet, the prize is the bet B.
- *S* is also an historical proposition, but the evidence tells equally for and against it. All the agents regard *S* as being about as likely as not. Moreover, *S* turns out to be false.
- The bet B is worth \$2 if *S* is true, and worth -\$1 if *S* is false. Although it is actually a losing bet, the agents all rationally value it at around 50 cents.
- How much A costs is determined by which proposition from the partition $\{q, r, s\}$ is true.
- If *q* is true, A costs \$2
- If *r* is true, A costs \$500
- If *t* is true, A costs \$1
- The evidence the agents has strongly supports *r*, though *t* is in fact true
- Quentin believes *q*
- Robby believes *r*
- Thom believes *t*

All of the agents make the utility calculations that their beliefs support, so Quentin and Thom take the bet and lose a dollar, while Robby declines it. Although Robby has a lot of evidence in favour of *p*, he correctly decides that it would be unwise to bet on *p* at effective odds of 1000 to 1 against. I'll now argue that both Quentin and Thom are potential counterexamples to (PC). There are three possibilities for what we can say about those two.

First, we could say that they are justified in believing *p*, and rational to take the bet. The problem with this position is that if they had rational beliefs about the partition

$\{q, r, t\}$ they would realise that taking the bet does not maximise expected utility. If we take rational decisions to be those that maximise expected utility given a rational response to the evidence, then the decisions are clearly not rational.

Second, we could say that although Quentin and Thom are not rational in accepting the bet, nor are they justified in believing that p. This doesn't seem particularly plausible for several reasons. The irrationality in their belief systems concerns whether q, r or t is true, not whether p is true. If Thom suddenly got a lot of evidence that t is true, then all of his (salient) beliefs would be well supported by the evidence. But it is bizarre to think that whether his belief in p is rational turns on how much evidence he has for t. Finally, even if we accept that agents in higher stakes situations need more evidence to have justified beliefs, the fact is that the agents are in a low-risk situation, since t is actually true, so the most they could lose is $1.

So it seems like the natural thing to say is that Quentin and Thom *are* justified in believing that p, and are justified in believing that given p, it maximises expected utility to take the bet, but they are not rational to take the bet. (At least, in the version of the story where they are thinking about which of q, r and t are correct given their evidence when thinking about whether to take the bet they are counterexamples to (PC).) Against this, one might respond that if belief in p is justified, there are arguments one might make to the conclusion that the bet should be taken. So it is inconsistent to say that the belief is justified, but the decision to take the bet is not rational. The problem is finding a premise that goes along with p to get the conclusion that taking the bet is rational. Let's look at some of the premises the agent might use.

- If p, then the best thing to do is to take the bet.

This isn't true (p is true, but the best thing to do isn't to take the bet). More importantly, the agents think this is only true if S is true, and they think S is a 50/50 proposition. So they don't believe this premise, and it would not be rational to believe it.

- If p, then probably the best thing to do is to take the bet.

Again this isn't true, and it isn't well supported, and it doesn't even support the conclusion, for it doesn't follow from the fact that x is probably the best thing to do that x should be done.

- If p, then taking the bet maximises rational expected utility.

This isn't true – it is a conditional with a true antecedent and a false consequent. Moreover, if Quentin and Thom were rational, like Robby, they would recognise this.

- If p, then taking the bet maximises expected utility relative to their beliefs.

This is true, and even reasonable to believe, but it doesn't imply that they should take the bet. It doesn't follow from the fact that doing something maximises expected utility relative to my crazy beliefs that I should do that thing.

- Given p, taking the bet maximises rational expected utility.

This is true, and even reasonable to believe, but it isn't clear that it supports the conclusion that the agents should take the bet. The implication appealed to here is PMP, and in this context that's close enough to equivalent to (PC). If we think that this case is a prima facie problem for (PC), as I think is intuitively plausible, then we can't use (PC) to show that it *doesn't* post a problem. We could obviously continue for a while, but it should be clear it will be very hard to find a way to justify taking the bet even spotting the agents p as a premise they can use in rational deliberation. So it seems to me that (PC) is not in general true, which is good because as we'll see in cases like this one the theory outlined here does not support it.

The theory we have been working with says that belief that p is justified iff the agent's degree of belief in p is sufficient to amount to belief in their context, and they are justified in believing p to that degree. Since by hypothesis Quentin and Thom are justified in believing p to the degree that they do, the only question left is whether this amounts to belief. This turns out not to be settled by the details of the case as yet specified. At first glance, assuming there are no other relevant decisions, we might think they believe that p because (a) they prefer (in the relevant sense) believing p to not believing p, and (b) conditionalising on p doesn't change their attitude towards the bet. (They prefer taking the bet to declining it, both unconditionally and conditional on p.)

But that isn't all there is to the definition of belief *tout court*. We must also ask whether conditionalising on p changes any preferences conditional on any active proposition. And that may well be true. Conditional on r, Quentin and Thom prefer not taking the bet to taking it. But conditional on r and p, they prefer taking the bet to not taking it. So if r is an active proposition, they don't believe that p. If r is not active, they do believe it. In more colloquial terms, if they are concerned about the possible truth of r (if it is salient, or at least not taken for granted to be false) then p becomes a potentially high-stakes proposition, so they don't believe it without extraordinary evidence (which they don't have). Hence they are only a counterexample to (PC) if r is not active. But if r is not active, our theory predicts that they are a counterexample to (PC), which is what we argued above is intuitively correct.

Still, the importance of r suggests a way of saving (PC). Above I relied on the position that if Quentin and Thom are not maximising rational expected utility, then they are being irrational. This is perhaps too harsh. There is a position we could take, derived from some suggestions made by Gilbert Harman in *Change in View*, that an agent can rationally rely on their beliefs, even if those beliefs were not rationally formed, if they cannot be expected to have kept track of the evidence they used to form that belief. If we adopt this view, then we might be able to say that (PC) is compatible with the correct normative judgments about this case.

To make this compatibility explicit, let's adjust the case so Quentin takes q for granted, and cannot be reasonably expected to have remembered the evidence for q. Thom, on the other hand, forms the belief that t rather than r is true in the course of thinking through his evidence that bears on the rationality of taking or declining the bet. (In more familiar terms, t is part of the inference Thom uses in coming to

conclude that he should take the bet, though it is not part of the final implication he endorses whose conclusion is that he should take the bet.) Neither Quentin nor Thom is a counterexample to (PC) thus understood. (That is, with the notion of rationality in (PC) understood as Harman suggests that it should be.) Quentin is not a counterexample, because he is *rational* in taking the bet. And Thom is not a counterexample, because in his context, where *r* is active, his credence in *p* does not amount to belief in *p*, so he is not justified in believing *p*.

We have now two readings of (PC). On the strict reading, where a rational choice is one that maximises rational expected utility, the principle is subject to counterexample, and seems generally to be implausible. On the loose reading, where we allow agents to rely on beliefs formed irrationally in the past in rational decision making, (PC) *is* plausible. Happily, the theory sketched here agrees with (PC) on the plausible loose reading, but not on the implausible strict reading. In the previous section I argued that the theory also accounts for intuitions about particular cases like *Local and Express*. And now we've seen that the theory accounts for our considered opinions about which principles connecting justified belief to rational decision making we should endorse. So it seems at this stage that we can account for the intuitions behind the pragmatic encroachment view while keeping a concept of probabilistic epistemic justification that is free of pragmatic considerations.

8 Conclusions

Given a pragmatic account of belief, we don't need to have a pragmatic account of justification in order to explain the intuitions that whether *S* justifiably believes that *p* might depend on pragmatic factors. My focus here has been on sketching a theory of belief on which it is the belief part of the concept of a justified belief which is pragmatically sensitive. I haven't said much about why we should prefer to take that option than say that the notion of epistemic justification is a pragmatic notion. I've mainly been aiming to show that a particular position is an open possibility, namely that we can accept that whether a particular agent is justified in believing *p* can be sensitive to their practical environment without thinking that the primary epistemic concepts are themselves pragmatically sensitive.[5]

References

Bovens, Luc, and James Hawthorne. 1999. "The Preface, the Lottery, and the Logic of Belief." *Mind* 108 (430): 241–64. doi: 10.1093/mind/108.430.241.

Christensen, David. 2005. *Putting Logic in Its Place*. Oxford: Oxford University Press.

Evnine, Simon. 1999. "Believing Conjunctions." *Synthese* 118: 201–27. doi: 10.1023/A:1005114419965.

Fantl, Jeremy, and Matthew McGrath. 2002. "Evidence, Pragmatics, and Justification." *Philosophical Review* 111: 67–94. doi: 10.2307/3182570.

Foley, Richard. 1993. *Working Without a Net*. Oxford: Oxford University Press.

Harman, Gilbert. 1986. *Change in View*. Cambridge, MA: Bradford.

Hawthorne, John. 2004. *Knowledge and Lotteries*. Oxford: Oxford University Press.

Hunter, Daniel. 1996. "On the Relation Between Categorical and Probabilistic Belief." *Noûs* 30: 75–98. doi: 10.2307/2216304.

Kaplan, Mark. 1996. *Decision Theory as Philosophy*. Cambridge: Cambridge University Press.

Keynes, John Maynard. 1921. *Treatise on Probability*. London: Macmillan.

Lewis, David. 1994. "Reduction of Mind." In *A Companion to the Philosophy of Mind*, edited by Samuel Guttenplan, 412–31. Oxford: Blackwell. doi: 10.1017/CBO9780511625343.019. Reprinted in his *Papers in Metaphysics and Epistemology*, 1999, 291-324, Cambridge: Cambridge University Press. References to reprint.

Maitra, Ishani. 2010. "Assertion, Norms and Games." In *Assertion: New Philosophical Essays*, edited by Jessica Brown and Herman Cappelen, 277–96. Oxford: Oxford University Press.

Ryle, Gilbert. 1949. *The Concept of Mind*. New York: Barnes; Noble.

Stalnaker, Robert. 1984. *Inquiry*. Cambridge, MA: MIT Press.

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

Weatherson, Brian. 2005. "True, Truer, Truest." *Philosophical Studies* 123 (1-2): 47–70. doi: 10.1007/s11098-004-5218-x.

Williamson, Timothy. 1994. *Vagueness*. Routledge.

———. 2000. *Knowledge and its Limits*. Oxford University Press.