On the pragmatic and epistemic virtues of inference to the best explanation

Richard Pettigrew
Richard.Pettigrew@bris.ac.uk

October 13, 2020

Abstract

In a series of papers over the past twenty years, and in a new book, Igor Douven has argued that Bayesians are too quick to reject versions of inference to the best explanation that cannot be accommodated within their framework. In this paper, I survey Douven's worries and attempt to answer them using a series of pragmatic and purely epistemic arguments that I take to show that Bayes' Rule really is the only correct way to respond to your evidence.

When we were ten years old, my friend Robert thought we saw a ghost. We were sleeping over at his house. Before bedtime, I cleaned my teeth and came back to his bedroom and we chatted before going to sleep. I knew I'd left the bathroom empty. As we chatted, we saw a shadow pass the door to our room on the way to the bathroom. His father, surely. Then we saw the same shadow pass back again in the other direction. And then it happened. A second shadow passed from the bathroom in the direction of Robert's parents' bedroom. But how could that be? The bathroom was empty when I'd left it. We'd seen one shadow pass and return. Whose was the second shadow? I maintained we must have been distracted when his mother's shadow passed to go to the bathroom after his father. But Robert concluded it must be a ghost, and he believes that to this day.

It's natural to say that both Robert and I used inference to the best explanation first to arrive at and then to justify our different conclusions. We shared the same evidence: my report that the bathroom was empty when I left it; the layout of the house; the second shadow moving away from the bathroom. I thought the best explanation was that the shadow belonged to Robert's mother, and we'd simply missed her passing to go to the bathroom; Robert favoured an explanation that posited ghosts.

At the name suggests, philosophers think of inference to the best explanation as a rule of inference. Indeed, it is often listed as one of the three species of inference: deduction, induction, and inference to the best explanation, also known as abduction (Lipton, 2004; Douven, 2017, 2021).

Inference to the best explanation (rule of inference)

From

(P1) E; and

(P2) H is the best explanation of E;

infer

(C) H

As such, it gives rise to a norm that governs our beliefs:

Inference to the best explanation (norm for beliefs)

If you believe (P1) and you believe (P2), you should believe (C).

But we can also think of inference to the best explanation as a norm that governs how we change our degrees of belief or credences when we receive new evidence:

Inference to the best explanation (norm for credences)

You should be more confident in better explanations of your total evidence than in poorer ones.[1]

So, if H_1 is a better explanation of E than H_2, and if p is our prior credence function and p_E is our posterior after learning E, then $p_E(H_1)$ should be greater than $p_E(H_2)$. We might even add that our posterior should track the extent to which one hypothesis is better than another, so that if H_1 is a much better explanation than H_2, then $p_E(H_1)$ should be much greater than $p_E(H_2)$, while if H_1 is only slightly better, then $p_E(H_1)$ should only be slightly greater than $p_E(H_2)$.

Now of course there are other norms we take to govern our credences, and they include norms that govern how to set our posterior given our prior and our evidence. So we might worry that the explanationist norms just sketched will conflict with them. The norms I have in mind are the Bayesian ones. There are at least two, though we will meet a third later:

Probabilism Your credences at any given time should satisfy the probability axioms.

That is, if your credence function p is defined on an algebra of propositions $\mathcal{F}$, then

(i) $0 \leq p(X) \leq 1$ for all propositions X in $\mathcal{F}$;

[1] Though see Lange (2020) for a more nuanced understanding of how the explanatory quality of hypotheses (or 'loveliness', as it has come to be called) relates to our posterior credences in them.

(ii) $p(\bot) = 0$ and $p(\top) = 1$, where $\bot$ is a contradiction and $\top$ a tautology;

(iii) $p(X \vee Y) = p(X) + p(Y) - p(XY)$ for all X, Y in $\mathcal{F}$.

Bayes' Rule When you receive new evidence, you should update your credences by conditioning your prior credences on your total evidence at that point.[2]

That is, if p is your prior, p_E your posterior when your total evidence is E, and $p(E) > 0$, then it ought to be that

$$p_E(X) = p_E^{\beta}(X) := p(X|E) := \frac{p(XE)}{p(E)}$$

Given an updating rule α, we write p_E^{α}, for the result of updating prior p on evidence E using rule α. We write β for Bayes' rule. So $p_E^{\beta}(X)$ is the result of updating p on E using Bayes' Rule.

Now, it is common to point out that Bayes' Theorem allows us to write Bayes' Rule in a couple of more useful ways:

Bayes' Rule (combined with Bayes' Theorem) If $p(E) > 0$, it ought to be that

$$p_E(X) = p_E^{\beta}(X) := p(X|E) = \frac{p(E|X)p(X)}{p(E|X)p(X) + p(E|\overline{X})p(\overline{X})}$$

And, more generally, if $H_1, \ldots, H_n$ is a set of mutually exclusive and exhaustive hypotheses, then it ought to be that

$$p_E(H_i) = p_E^{\beta}(H_i) := p(H_i|E) = \frac{p(E|H_i)p(H_i)}{\sum_{j=1}^{n} p(E|H_j)p(H_j)}$$

So, if I entertain a set of hypotheses that form a partition, my posterior confidence in each hypothesis is obtained by asking how likely the evidence is given that hypothesis, weighting that by how likely I thought the hypothesis was prior to receiving the evidence, and then normalizing the results.

Now, if H_1 is a better explanation for E than H_2, then Bayes' Rule tells us that

$$p_E(H_1) > p_E(H_2) \quad \text{iff} \quad p(E|H_1)p(H_1) > p(E|H_2)p(H_2)$$

So there are two straightforward ways to accommodate inference to the best explanation within Bayesianism:

[2] In fact, one of the nice features of Bayes' Rule is that you get the same result if you update by conditioning on your total evidence or just on your new evidence. After all, if $q(X) = p(X|E)$, then $q(X|F) = p(X|EF)$. But this is not true for the explanationist's rival updating rule ε, which we'll describe below. So it will be easier to state all candidate updating rules as operating on total evidence.

(1) Set $p(H_1) > p(H_2)$ and $p(E|H_1) \approx p(E|H_2)$. That is, assign a higher unconditional prior to more explanatory hypotheses.

(2) Set $p(E|H_1) > p(E|H_2)$ and $p(H_1) \approx p(H_2)$. That is, assign a higher likelihood to the evidence conditional on the more explanatory hypothesis.

Either of these might account for my conclusion or my friend Robert's when we saw that second shadow passing away from the bathroom that night. It might have been that we roughly agreed on the likelihood of our evidence given each hypothesis, but disagreed on the prior probability of the hypothesis: Robert might just have been antecedently much more confident that ghosts exist, and much less confident that we were distracted enough to miss his mother's shadow as she passed to go to the bathroom. Or we might have both agreed that it is very unlikely that ghosts exist and reasonably likely that we were distracted, but disagreed on how likely each hypothesis made our evidence: Robert might just have thought that, if ghosts were to exist, this is quite a likely way they'd show themselves. Or, of course, it might be a bit of both.

In general, we can better accommodate some cases of inference to the best explanation using (1), and some using (2), and some using a combination. You might, for instance, have two empirically equivalent hypotheses, such as the realist's hypothesis that the external world exists and is as we perceive it to be (H_1) and the sceptic's hypothesis that our experience of the external world is an illusion imposed on us by some powerful deceivers trying to trick us into thinking that it is the way we perceive it to be (H_2). In that case, providing neither is stronger than the other, it's plausible that $p(E|H_1) = p(E|H_2)$. Indeed, if both hypotheses entail E, then $p(E|H_1) = 1 = p(E|H_2)$. In that case, we can only ensure that one receives higher posterior probability than the other by assigning it higher prior unconditional probability. So, if you want to use inference to the best explanation to justify your higher posterior in realism, you'd better set $p(H_1) > p(H_2)$. That is, you must use (1).

But sometimes (1) won't do. I set an urn in front of you that contains three balls. I tell you that either two balls are violet and one green (H_1) or two balls are green and one violet (H_2). You will draw a ball at random, look at its colour, and update your credences in the two hypotheses in the light of your evidence. So there are two possible pieces of evidence you might receive: you might draw a violet ball (E_1) or you might draw a green one (E_2). Intuitively, H_1 explains E_1 better than H_2 does, while H_2 explains E_2 better than H_1 does. So, the credal version of inference to the best explanation demands that

$$p_{E_1}(H_1) > p_{E_1}(H_2) \quad \text{and} \quad p_{E_2}(H_2) > p_{E_2}(H_1)$$

But we can't ensure that only by setting $p(H_1) > p(H_2)$ or $p(H_2) > p(H_1)$. Instead, we must set $p(E_1|H_1) > p(E_1|H_2)$ and $p(E_2|H_2) > p(E_2|H_1)$. In fact, that seems reasonable anyway. Indeed, it is mandated by a norm that is often added to Probabilism and Bayes' Rule to give a slightly stronger version of Bayesianism, namely, David Lewis's Principal Principle (Lewis, 1980).

Principal Principle It ought to be the case that

$$p(X \mid \text{the chance of } X \text{ is } r) = r$$

In the case we're considering the Principal Principle demands:

$$\begin{array}{rclcrcl} p(E_1|H_1) & = & \frac{2}{3} & \quad & p(E_1|H_2) & = & \frac{1}{3} \\ p(E_2|H_1) & = & \frac{1}{3} & \quad & p(E_2|H_2) & = & \frac{2}{3} \end{array}$$

If $p(H_1) = p(H_2)$, then by Bayes' Rule we have:

$$p_{E_1}(H_1) > p_{E_1}(H_2) \quad \text{and} \quad p_{E_2}(H_2) > p_{E_2}(H_1)$$

as we wished. And we obtained that using (2).

The upshot of the preceding discussion is that Bayesianism can accommodate much of what the credal version of inference to the best explanation demands.[3] And, as Jonathan Weisberg (2009) points out, it could go further and mandate it if we were to embrace a less subjectivist and more objectivist version of Bayesianism; one that limits the rational priors in such a way that, whenever H_1 better explains E than H_2 does, $p(H_1|E) > p(H_2|E)$.

Nonetheless, some think that this strategy does not go far enough. For instance, think again about the mystery urn from above. If I have equal priors in the two hypotheses about the colour distribution in the urn, and if I update using Bayes' Rule, here are my posteriors if I draw a violet ball:

$$p_{E_1}(H_1) = p^{\beta}_{E_1}(H_1) = \frac{p(E_1|H_1)p(H_1)}{p(E_1|H_1)p(H_1) + p(E_1|H_2)p(H_2)} = \frac{\frac{2}{3}\frac{1}{2}}{\frac{2}{3}\frac{1}{2} + \frac{1}{3}\frac{1}{2}} = \frac{2}{3}$$

and

$$p_{E_1}(H_2) = p^{\beta}_{E_1}(H_2) = \frac{p(E_1|H_2)p(H_2)}{p(E_1|H_1)p(H_1) + p(E_1|H_2)p(H_2)} = \frac{\frac{1}{3}\frac{1}{2}}{\frac{2}{3}\frac{1}{2} + \frac{1}{3}\frac{1}{2}} = \frac{1}{3}$$

So $p_{E_1}(H_1) > p_{E_1}(H_2)$, as we hoped. But you might think that, while Bayes' Rule results in higher posterior confidence in H_1 upon learning E_1,

[3] For more detailed accounts that fit inference to the best explanation inside Bayesianism, see (Okasha, 2000), (McGrew, 2003), and (Lipton, 2004). For an argument that it cannot fit even with probabilism let alone Bayes' Rule, see (Climenhaga, 2017).

it doesn't make that posterior confidence high enough. You might think that, upon seeing the violet ball, you should be even more confident in H_1 than Bayes' Rule mandates, and even less confident in H_2. As I noted above, Bayes' Rule says that my posterior confidence in each hypothesis from should be obtained by asking how likely the evidence is given that hypothesis, weighting that by how likely I thought the hypothesis was prior to receiving the evidence, and then normalizing the resulting credences. You might think instead that I should ask how likely the evidence is given the hypothesis, weight that by how likely I thought the hypothesis was prior to learning the evidence, *then add a little boost to that weighted likelihood if the hypothesis is one of best explanations of the evidence*, and then normalize. That is, instead of updating by Bayes' Rule, we should use the Explanationist's Rule, which we write as ε and which says that it ought to be the case that:

$$p_{E_1}(H_1) = p^{\varepsilon}_{E_1}(H_1) := \frac{p(E_1|H_1)p(H_1) + c}{p(E_1|H_1)p(H_1) + p(E_1|H_2)p(H_2) + c} = \frac{\frac{2}{3}\frac{1}{2} + c}{\frac{2}{3}\frac{1}{2} + \frac{1}{3}\frac{1}{2} + c} = \frac{2+6c}{3+6c}$$

since H_1 best explains E_1, and

$$p_{E_1}(H_2) = p^{\varepsilon}_{E_1}(H_2) := \frac{p(E_1|H_2)p(H_2)}{p(E_1|H_1)p(H_1) + p(E_1|H_2)p(H_2) + c} = \frac{\frac{1}{3}\frac{1}{2}}{\frac{2}{3}\frac{1}{2} + \frac{1}{3}\frac{1}{2} + c} = \frac{1}{3+6c}$$

since H_2 does not best explain E_1. So $c > 0$ is a boost that is awarded to the best explanation over and above what is already given by Bayes' Rule. If $c > 0$, then the explanationist demands that

$$p_{E_1}(H_1) = p^{\varepsilon}_{E_1}(H_1) = \frac{2+6c}{3+6c} \quad \text{and} \quad p_{E_1}(H_2) = p^{\varepsilon}_{E_1}(H_2) = \frac{1}{3+6c}$$

And

$$p^{\varepsilon}_{E_1}(H_1) = \frac{2+6c}{3+6c} > \frac{2}{3} = p^{\beta}_{E_1}(H_1)$$

and

$$p^{\varepsilon}_{E_1}(H_2) = \frac{1}{3+6c} < \frac{1}{3} = p^{\beta}_{E_1}(H_1)$$

which is exactly what this more extreme version of explanationism demands, dissatisfied as it is with the boost given to H_1 and the reduction given to H_2 by Bayesianism. In what follows, I will call this more extreme, non-Bayesian version of explanationism simply explanationism, since the less extreme version is simply Bayesianism.

The explanationist update rule we just described is a particular case of the following rule, which van Fraassen (1989, Chapter 6) sketched in his early discussion of the tension between inference to the best explanation and Bayesianism, and which Igor Douven (2013, 2021) has made precise and explored in great detail:

> **Explanationist's Rule** If $H_1, \ldots, H_n$ is a set of mutually exclusive and exhaustive hypotheses, then it ought to be that
>
> $$p_E(H_i) = p^{\varepsilon}_E(H_i) := \frac{p(E|H_i)p(H_i) + f(H_i, E)}{\sum_{j=1}^{n} \left(p(E|H_j)p(H_j) + f(H_i, E)\right)}$$
>
> where $f(H_i, E)$ gives a reward to H_i if it is one of the best explanations on E from among $H_1, \ldots, H_n$.[4]

In Douven's version of the rule, each time you apply it, there is some fixed positive amount c of reward that we distribute evenly between the best explanations of the total evidence gathered so far. So, if there are k best explanations of E, then $f(H_i, E) = \frac{c}{k}$ if H_i is among them, and $f(H_i, E) = 0$ if it is not.

As we've already seen, typically, $p^{\varepsilon}_E \neq p^{\beta}_E$. So Bayesianism conflicts with this non-Bayesian version of explanationism. Which should we use? That is the question that will engage us for the rest of the paper. And it is a question of no small moment. Bayesianism is a central statistical tool in contemporary science, from epidemiology to particle detection; but inference to the best explanation is often advertised as a central component of the scientific method. If they do conflict and if we must choose one over the other, there will be work to do.

Van Fraassen defended Bayesianism against this extreme version of explanationism by appealing to David Lewis' betting argument for Bayes' Rule. Igor Douven has considered that argument, as well as other pragmatic considerations and also accuracy-based arguments for Bayes' Rule. He thinks that none decisively establish Bayes' Rule, and presents considerations in favour of the non-Bayesian explanationist rule, at least in certain situations. His goal is to reject the dominance of Bayesianism, rather

[4] Recall: p_E is your posterior when your total evidence is E, and p^{ε}_E is the posterior the explanationist's rule demands in that situation. We can now see why it is important to specify that update rules go to work on the prior and the *total* evidence and not just the *new* evidence. In our urn example, suppose you first draw a violet ball and replace it; you update using the explanationist's rule ε; next, you draw a green ball and replace it; you update again using the explanationist's rule ε. For the first update, your new evidence and total evidence are the same—the first ball drawn is violet—and both are best explained by H_1, so that gets the boost. For your second update, your new evidence is that the second draw was green: this is best explained by H_2; so that would then get the boost all to itself. But your total evidence is that the first draw was violet and the second was green: this is equally well explained by both hypotheses; so they would share the boost equally between themselves. So we get two different rules depending on whether they act on the new evidence or the total evidence. That distinguishes the explanationist approach from the Bayesian one. The explanationist rule that Douven considers is the one that acts on the total evidence, and that is the version of the rule I'll consider throughout.

than to establish the dominance of explanationism. He allows that Bayes' Rule may be the right way to go in certain situations, but sees no reason to think it is always the right way to update. In the remainder of the paper, I'll consider Douven's arguments, describe further arguments in favour of Bayes' Rule, some pragmatic and some purely epistemic. I'll argue that they provide compelling responses to Douven's concerns. I conclude that the dominance of Bayes' Rule should continue.

1 Pragmatic arguments for Bayes' Rule

I'll start in this section with the argument for Bayes' Rule to which van Fraassen appealed when he first argued against non-Bayesian versions of inference to the best explanation. I'll then consider Igor Douven's responses to that argument, and that will lead me to introduce two further pragmatic arguments for Bayes' Rule.

1.1 Lewis' sure loss argument for Bayes' Rule

Van Fraassen took the sure loss argument for Bayes' Rule from David Lewis, who had presented it in a seminar at Princeton in the 1970s, but didn't publish it himself until 1999 (Lewis, 1999). It's a betting argument of the sort that Frank Ramsey and Bruno de Finetti provided for probabilism (Ramsey, 1926 [1931]; de Finetti, 1937 [1980]). So it starts with the same basic premise as those: if your credence in a proposition X is p, then for any stake S, whether positive or negative and regardless how large, you are rationally required to accept any bet that gains you more than £$(1-p) \times S$ if X is true and loses you less than £$p \times S$ if X is false. Lewis shows that, if you do not plan to update by Bayes' Rule, there is a series of bets each of which your prior rationally requires you to accept and a series of bets each of which your posterior requires you to accept that, taken together, will lose you money for sure.[5] Lewis contends that planning to update in a way that makes you vulnerable to such a sure loss is irrational.

Igor Douven provides three responses to van Fraassen's argument:

(1) First, he suggests that we can have the best of both worlds by setting our priors in such a way that following Bayes' Rule when we update gives us posteriors that agree with the explanationist's updating rule but avoid the sure loss (Douven, 1999). We'll consider this in Section 1.2.

(2) Second, he argues that, while it is certainly a consideration against an updating rule that it renders you vulnerable to a sure loss, we cannot conclude that it renders you irrational without considering

[5] R. A. Briggs (2009) gives a particularly clear presentation of the argument.

whether there are considerations in its favour that compensate for this flaw; and he argues that there are such considerations (Douven, 2013, 2021). This is the topic of Sections 1.4-1.6.

(3) Third, he suggests that we cannot establish any credal norm by paying attention only to pragmatic considerations. We must instead show that there is an epistemic flaw in updating rules other than Bayes' Rule (Douven, 2013; Douven & Wenmackers, 2017; Douven, 2021). That will bring us to the accuracy arguments in Section 2, and their extension into questions of social epistemology in Section 3 and choices between different intellectual trajectories in Section 4.

1.2 Avoiding the sure loss

You are about to learn something. You know that it will be a proposition in the partition $E_1, \ldots, E_m$. You consider each of the mutually exclusive and exhaustive hypotheses $H_1, \ldots, H_n$. I give you a prior p and an updating rule α. Together, these determine, for each possible piece of evidence E_j, a posterior credence function $p^{\alpha}_{E_j}$ that the rule α says you should adopt if you learn E_i. Then it's possible to pick an alternative prior q in such a way that updating q on E_j using Bayes' Rule β will agree with updating p on E_j using α. That is, $p^{\alpha}_{E_j} = q^{\beta}_{E_j}$, for each possible piece of evidence E_j.

Here's the trick: first, pick your alternative priors in the different possible pieces of evidence; that is, pick $q(E_1), \ldots, q(E_m)$; then set your alternative priors in the conjunctions of hypotheses with evidence as follows:

$$q(H_i E_j) = p^{\alpha}_{E_j}(H_i) q(E_j)$$

That then completely determines your alternative prior credence function q, and it's easy to show that, defined in this way, q is a probability function.[6] What's more:

$$q^{\beta}_{E_j}(H_i) = q(H_i | E_j) = \frac{q(H_i E_j)}{q(E_j)} = \frac{p^{\alpha}_{E_j}(H_i) p(E_j)}{p(E_j)} = p^{\alpha}_{E_j}(H_i)$$

as required. So, in particular, if p is the prior to which we wish to apply the explanationist's updating rule ε, we can pick an alternative prior q in such a way that $q^{\beta}_{E_j} = p^{\varepsilon}_{E_j}$, for any possible piece of evidence E_j. Providing we then use q as our prior, we can then update by Bayes' Rule to $p^{\varepsilon}_{E_j} = q^{\beta}_{E_j}$ and thereby sidestep the sure loss argument against the explanationist.

[6] It suffices to show that

$$\sum_{i=1}^{n} \sum_{j=1}^{m} q(H_i E_j) = \sum_{i=1}^{n} \sum_{j=1}^{m} p^{\alpha}_{E_j}(H_i) q(E_j) = \sum_{j=1}^{m} q(E_j) \sum_{i=1}^{n} p^{\alpha}_{E_j}(H_i) = \sum_{j=1}^{m} q(E_j) = 1$$

However, pushing down the lump in the carpet here just causes it to pop up unwanted elsewhere. In this case, using this trick leads to a prior that violates the Principal Principle, the extra norm of Bayesianism that we met above.

Return once more to our urn and the case in which we draw just a single ball. H_1 says that the chance of drawing a violet ball is two-thirds, while H_2 says the same for drawing a green ball. E_1 is your evidence if you draw a violet ball, and E_2 is your evidence if you draw a green ball. The Principal Principle demands that, if q is your prior, then

$$q(E_1|H_1) = \frac{2}{3} = q(E_2|H_2)$$

But according to the construction of the prior we've just described,

$$\begin{aligned} q(E_1|H_1) = \frac{q(H_1E_1)}{q(H_1E_1) + q(H_1E_2)} = \\ \frac{p^{\varepsilon}_{E_1}(H_1)q(E_1)}{p^{\varepsilon}_{E_1}(H_1)q(E_1) + p^{\varepsilon}_{E_2}(H_1)q(E_2)} = \\ \frac{\frac{2+6c}{3+6c}q(E_1)}{\frac{2+6c}{3+6c}q(E_1) + \frac{1}{3+6c}q(E_2)} = \\ \frac{(2+6c)q(E_1)}{(2+6c)q(E_1) + q(E_2)} \end{aligned}$$

and

$$\begin{aligned} q(E_2|H_2) = \frac{q(H_2E_2)}{q(H_2E_1) + q(H_2E_2)} = \\ \frac{p^{\varepsilon}_{E_2}(H_2)q(E_2)}{p^{\varepsilon}_{E_1}(H_2)q(E_1) + p^{\varepsilon}_{E_2}(H_2)q(E_2)} = \\ \frac{\frac{2+6c}{3+6c}q(E_2)}{\frac{1}{3+6c}q(E_1) + \frac{2+6c}{3+6c}q(E_2)} = \\ \frac{(2+6c)q(E_2)}{q(E_1) + (2+6c)q(E_2)} \end{aligned}$$

Now, if q satisfies the Principal Principle, then $q(E_1|H_1) = \frac{2}{3} = q(E_2|H_2)$. And if that's the case, then it is easy to see from the equations we've just set down that $q(E_1) = q(E_2)$. But then, by those same equations,

$$q(E_1|H_1) = \frac{2+6c}{3+6c} \quad \text{and} \quad q(E_2|H_2) = \frac{2+6c}{3+6c}$$

But then $q(E_1|H_1) = \frac{2}{3} = q(E_2|H_2)$ only if $c = 0$. So q satisfies the Principal Principle only if $c = 0$ and the explanationist's rule is just Bayes' Rule.

So, if we want to give an extra boost to the best explanation of our total evidence over and above what Bayes' Rule already gives it, and we wish to avoid Lewis' sure loss argument against violations of Bayes' Rule, we must pick a prior that violates the Principal Principle. And while the Converse Dutch Book Theorem ensures that there is no sure loss argument against violations of the Principal Principle that satisfy probabilism, there is an expected loss argument against it (Pettigrew, 2020, Section 2.8). It turns on the following fact: if you violate the Principal Principle, there is a set of bets that your credences will require you to enter into such that, whatever the objective chances are, those chances will expect you to lose money from those bets.

Douven himself recognises that the prior he constructs to match with a non-Bayesian updating rule might leave it vulnerable to some sort of betting argument. But he contends that such vulnerability is no threat to your rationality. After all, you could see your sure loss or expected loss coming, and simply refuse to enter into the final bet that locks you in to that loss (Douven, 1999, S429-S434).

One problem with this response is that, if it works against the expected loss argument for the Principal Principle, it also works against the sure loss argument for probabilism, since the sure loss there is just as visible to the person who violates probabilism as it is to the imagined bookie. However, the real problem with Douven's argument is that this 'look before you leap' strategy works against neither argument. Suppose that you satisfy probabilism but violate the Principal Principle, which is what Douven's strategy requires of you. And suppose that, faced with a decision problem, rationality requires you to choose by maximizing expected utility. Then it turns out that you should accept each bet offered in the expected loss argument for the Principal Principle, since each maximises expected utility for you; and this is true even if you take into account the bets that you've already accepted (Pettigrew, 2020, Section 3.4). So even at the final stage of the expected loss argument, where there is just one more bet to consider, and you know what you've already accepted and you can see that accepting this final bet locks you in to an expected loss, accepting it still has greater expected utility from the point of view of your credence function than rejecting it. So even if you do look before you leap, and even if you do see what awaits you should you leap, your credences still rationally require you to leap. Indeed, it is this that renders them irrational.

1.3 Further pragmatic arguments for Bayes' Rule

This brings us to Douven's second objection to van Fraassen's argument. The sure loss argument for Bayes' Rule presents vulnerability to a sure loss as a flaw that renders an updating rule irrational. But it is a very peculiar sort of flaw. On the one hand, when it manifests, it will lose you money for sure, and there is no limit to the amount of money it will lose you, since the stake S of the bets may be set as high as you like. But, on the other hand, the set of choices you must face in order that the flaw becomes manifest is very specific and quite unlikely to arise. So, if you think other decision problems are more likely, and if the credences your updating rule bequeaths to you serve you better when you face those than the credences that Bayes' Rule demands, then you might well think that this outweighs the flaw of vulnerability to a sure loss.

I'm very sympathetic to the starting point of this argument. I agree that vulnerability to a sure loss does not, on its own, render credences irrational. But I think the prospects are bleak for finding some virtue of alternative updating rules that compensates for this flaw. The reason is that the sure loss argument is by no means the only argument for Bayes' Rule that appeals to how well your credences serve you as a basis for decision-making. In the following two sections, I'll describe two more.

1.4 The expected pragmatic utility argument for Bayes' Rule

The first is due to Peter M. Brown (1976) and it is perhaps best seen as a generalization of I. J. Good's Value of Information Theorem (Good, 1967).[7] The set up is this. I am about to learn some evidence. After I learn this new evidence, I'll face a decision—that is, I'll have to choose between a set of available acts. I'll make this choice by maximising expected utility from the point of view of my credences at that time. How, then, should I plan to update my credences, knowing that I'll use them to make this decision? Good showed that, if your only two options are to use Bayes' Rule to update or to simply stick with your prior when the evidence comes in, then your prior expects Bayes' Rule to produce posteriors that guide your choice after the evidence comes in better than sticking with your prior does. Brown generalizes this by showing that your prior expects Bayes' Rule to produce posteriors that guide your actions better than *any* available updating rule.

Suppose:

- your prior is p;
- the evidence you're about to receive will be a proposition from the set $\mathcal{E}$, and suppose that those propositions are mutually exclusive and exhaustive—that is, $\mathcal{E}$ is a partition; if w is a possible world, E_w is the unique proposition in $\mathcal{E}$ that is true at w;
- α is an updating rule that tells you to adopt p^α_E if you start with p and learn E from $\mathcal{E}$; we write p^α_w for $p^\alpha_{E_w}$—that is, p^α_w is the posterior you

[7] It's pretty clear that Savage already knew Good's theorem when he wrote *The Foundations of Statistics* (Savage, 1954, Section 7.3).

would end up with if you were to update the prior p on the evidence you would receive from $\mathcal{E}$ at world w;

- if a is an act and w is a possible world, $a(w)$ is the utility of a at w; and
- if q is a credence function, a^q is an act that maximizes expected utility by the lights of q, so that

$$\sum_{w \in W} q(w)a^q(w) \geq \sum_{w \in W} q(w)a(w)$$

for all acts a.

Then the expected utility of updating your prior p using the rule α is:

$$\mathrm{Exp}_p(\text{Use rule } \alpha) = \sum_{w \in W} p(w)a^{p^\alpha_w}(w)$$

In particular, the expected utility of updating using Bayes' Rule β is:

$$\mathrm{Exp}_p(\text{Use rule } \beta) = \sum_{w \in W} p(w)a^{p^\beta_w}(w)$$

Now, for any world w, by the definition of $a^{p^\beta_w}$,

$$\sum_{w' \in W} p^\beta_w(w')a^{p^\beta_w}(w') \geq \sum_{w' \in W} p^\beta_w(w')a^{p^\alpha_w}(w')$$

So, since $p^\beta_w(w') = \frac{p(w')}{p(E_w)}$ if w' is in E_w and $p^\beta_w(w') = 0$ if w' is not in E_w,

$$\sum_{w' \in E_w} p(w')a^{p^\beta_w}(w') \geq \sum_{w' \in E_w} p(w')a^{p^\alpha_w}(w')$$

But of course, if w' is in E_w, then $E_{w'} = E_w$ and $p^\beta_{w'} = p^\beta_w$ and $p^\alpha_{w'} = p^\alpha_w$. So

$$\sum_{w \in W} p(w)a^{p^\beta_w}(w) \geq \sum_{w \in W} p(w)a^{p^\alpha_w}(w)$$

So

Theorem 1 (Expected pragmatic argument) *For any prior p,*

$$\mathrm{Exp}_p(\textit{Use rule } \beta) \geq \mathrm{Exp}_p(\textit{Use rule } \alpha)$$

And, if there is a world w such that (i) $a^{p^\beta_w} \neq a^{p^\alpha_w}$ *and* (ii) $p(w) > 0$, *then this inequality is strict.*

That is, if you give any prior credence to ending up with a posterior that chooses differently from how the Bayesian's posterior will choose, then your prior expects updating using β to be strictly better. So, if we must make a choice after receiving some evidence, our prior expects us to make that choice best if we choose using the posteriors we get by updating using Bayes' Rule.

Of course, we are not often in the precise situation covered by this result. Rarely do we know which decisions we will face using the posteriors that our updating rule bestows on us when we deploy it on our next piece of evidence. What's more, an updating rule doesn't just give you the credences you will use to make decisions after you receive this piece of evidence. It also gives you the credences you will update when you receive the next piece of evidence after that. And then the credences you will update when you receive the next piece of evidence after that. And so on. So we should be concerned not only with the choices that our updated credences mandate, but also the choices that our updated updated credences mandate and our updated updated update credences, and so on.

Fortunately, Brown's reasoning goes through even for this more complex but more realistic situation, provided we grant a certain assumption, which we'll explain below. Here's the setup. Suppose p is your prior. Suppose $t_1, \ldots, t_n$ are the times during your epistemic life. For each $1 \leq i \leq n$,

- Your total evidence at t_i is a proposition in the partition $\mathcal{E}_i$; Let $E_{w,i}$ be the total evidence from $\mathcal{E}_i$ that you will have at time t_i at world w.[8]

- If α is an updating rule and w is a possible world, $p^{\alpha}_{w,i}$ is the credence function you reach in world w by time t_i if you start with prior p and successively apply α to the total evidence you'll have at that world at each time $t_1, \ldots, t_i$.

- The decision problem you will face at t_i comes from the set $\mathcal{D}_i$. We can assume without loss of generality that you just face a single decision problem at each time t_i. If you face two, we just combine them into a single composite one.[9] Let $D_{w,i}$ be the decision problem in $\mathcal{D}_i$ that you face at time t_i in world w.

- $0 < \lambda_i < 1$ is the weight that records how much you care about the pragmatic utility your credences obtain for you at time t_i.

[8]Since we assume that total evidence is cumulative, so that your total evidence at a later time is strictly stronger than your total evidence at an earlier time, it follows that each $\mathcal{E}_{i+1}$ is a fine-graining of $\mathcal{E}_i$—that is, each proposition in $\mathcal{E}_i$, there is a set of propositions in $\mathcal{E}_{i+1}$ that partitions it.

[9]Here's how to do that: Suppose decision problem D consists of available acts A and D' consists of available acts A'. Then define $D \times D'$ to be the decision problem with available acts $A \times A' = \{(a, a') : a \in A \ \& \ a' \in A'\}$, where $(a, a')(w) = a(w) + a'(w)$.

- Given credence function q and decision problem D, let a^q_D be the act in D that maximises expected utility from the point of view of q.

Then:

$$\text{Exp}^{\Lambda}_p(\text{Use rule } \alpha) = \sum_{w \in W} p(w) \sum_{t_i} \lambda_i a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w)$$

Now we introduce the assumption we must make if we are to extend Brown's proof: for any time t_i, for any total body of evidence E in $\mathcal{E}_i$, and for any decision problem D in $\mathcal{D}_i$, $p(w|ED) = p(w|E)$. Assuming that, we can prove:

$$\text{Exp}^{\Lambda}_p(\text{Use rule } \beta) \geq \text{Exp}^{\Lambda}_p(\text{Use rule } \alpha)$$

After all, for any world w and any time t_i, by the definition of $a^{p^{\beta}_{w,i}}_{D_{w,i}}$,

$$\sum_{w' \in W} p^{\beta}_{w,i}(w') a^{p^{\beta}_{w,i}}_{D_{w,i}}(w') \geq \sum_{w' \in W} p^{\beta}_{w,i}(w') a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w')$$

So

$$\sum_{w' \in W} p(w'|E_{w,i}) a^{p^{\beta}_{w,i}}_{D_{w,i}}(w') \geq \sum_{w' \in W} p(w'|E_{w,i}) a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w')$$

But, by our assumption, $p(w'|E_{w,i}) = p(w'|E_{w,i}D_{w,i})$, so

$$\sum_{w' \in W} p(w'|E_{w,i}D_{w,i}) a^{p^{\beta}_{w,i}}_{D_{w,i}}(w') \geq \sum_{w' \in W} p(w'|E_{w,i}D_{w,i}) a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w')$$

So,

$$\sum_{w' \in E_{w,i}D_{w,i}} p(w') a^{p^{\beta}_{w,i}}_{D_{w,i}}(w') \geq \sum_{w' \in E_{w,i}D_{w,i}} p(w') a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w')$$

But of course, if w' is in $E_{w,i}D_{w,i}$, then $E_{w',i} = E_{w,i}$, $D_{w',i} = D_{w,i}$, and thus $p^{\beta}_{w',i} = p^{\beta}_{w,i}$ and $p^{\alpha}_{w',i} = p^{\alpha}_{w,i}$. So

$$\sum_{w \in W} p(w) \sum_{t_i} \lambda_i a^{p^{\beta}_{w,i}}_{D_{w,i}}(w) \geq \sum_{w \in W} p(w) \sum_{t_i} \lambda_i a^{p^{\alpha}_{w,i}}_{D_{w,i}}(w)$$

And thus

Theorem 2 (Longitudinal expected pragmatic argument) *For any prior p,*

$$\text{Exp}^{\Lambda}_p(\textit{Use rule } \beta) \geq \text{Exp}^{\Lambda}_p(\textit{Use rule } \alpha)$$

And, if there is a time t_i and a world w such that (i) $a^{p^{\beta}_{w,i}} \neq a^{p^{\alpha}_{w,i}}$ *and* (ii) $p(w) > 0$, *then this inequality is strict.*

That is, if you give any prior credence to ending up at some point with a posterior that chooses differently from how the Bayesian's posterior will choose at that point, then your prior expects updating using β to be strictly better.

The problem with the sure loss argument for Bayes' Rule is that it declares any alternative updating rule irrational just because there is a very specific decision problem you might face where your priors, together with the credences issued by that updating rule, serve you very badly indeed—to wit, they lead you to accept a sure loss. Douven's worry is that, while this is certainly a strike against non-Bayesian updating rules, it is a shortcoming for which they might compensate in other ways. The foregoing expected pragmatic utility argument pours cold water on that hope. Whichever series of decision problems you might face at whatever stage of your epistemic life, and almost whatever prior credences you have in facing those decisions, you will be served best by updating using Bayes' Rule. Or at least that is what your prior expects.

Now Douven notes that we surely care more about the *actual* pragmatic utility of adopting a particular updating rule than about its *expected* pragmatic utility. So does the foregoing argument tell us nothing until we find out which rule maximizes actual pragmatic utility? Surely not. This objection mistakes the reason we care about expected pragmatic utility. We care about it precisely because we care about actual pragmatic utility. It is our best way of choosing options when maximizing actual pragmatic utility is our aim but our ignorance of what the actual world is like prevents us from maximizing that directly. When I have a headache and choose which painkiller to take, I ask myself which will minimize my expected pain. I do this not because I care about expected pain in itself, but because I care about my actual pain, and I think minimizing expected pain is my best shot at minimizing that.

If we know more about the actual world than is encoded in our prior, then we should incorporate that new information into our prior and then do whatever maximizes expected pragmatic utility from the point of view of this new updated prior. And, again, the advice will be to follow Bayes' Rule, but this time applied to our updated prior. It is no surprise that, if we know more about the actual world than our prior does, we can find updating rules that actually outperform what our prior expects to do best. If I know that, in fact, the urn contains two violet balls and one green ball, while my prior assigns only credence 0.5 to that hypothesis, then I can simply update by setting my credence in that hypothesis to 1, regardless of the further evidence I observe, and that will actually outperform Bayes' Rule as applied to my prior. But that is no objection to the expected pragmatic utility argument for Bayes' Rule.

1.5 The pragmatic utility dominance argument for Bayes' Rule

In any case, Brown's expected pragmatic utility argument suggests a new argument for Bayes' Rule that might better satisfy someone who feels that maximising expected utility is not our best shot at maximising actual utility. The idea behind the argument is drawn from a brief section in Savage's seminal 1971 paper on eliciting credences using scoring rules (Savage, 1971, Section 7). It is developed a little by Schervish (1989), and made more explicit by Levinstein (2017), and then by Pettigrew (2020, Section 6.3).

In the extended version of Brown's argument that we presented above, we assess update rules by the utility they obtain for us when we use the credences they provide to make choices. Why not then use this to give an explicit measure of the pragmatic utility of a credence function at a world? Here's the idea. As above, we can assume that, for each credence function you have during your epistemic life, you will face only one decision problem armed with it. Let's focus on a particular credence function q for the moment. Let C be an objective chance distribution that says how likely a particular decision problem is to be the one you face armed with q. That is, C is a probability distribution over the space of possible decision problems, where such a problem is specified by (i) the number of acts available to the chooser, which we can assume is bounded by some very large finite number; and (ii) for each act and each world, the utility of that act at that world, which might take any real number. If there were just finitely many decision problems D in the set $\mathcal{D}$ of decision problems you might face with q, then we'd define the pragmatic utility relative of q at a possible world w relative to the chance distribution C as follows:

$$\mathfrak{U}_C(q, w) = \sum_{D \in \mathcal{D}} C(D) a^q_D(w)$$

where $C(D)$ is the chance that you'll face decision problem D armed with credence function q. Thus, $\mathfrak{U}_C(q, w)$ is the expected utility at world w of the choice you will make when you use q to make your decisions, where the probability that goes into the expectation is not the probability of the world, since that is fixed to be w, but the probability that you will face that decision problem. Of course, in reality, there are infinitely many decision problems you might face armed with a particular credence function, and so we must use an integral rather than a summation:

$$\mathfrak{U}_C(q, w) = \int_{\mathcal{D}} a^q_D(w) dC$$

Building on Savage's observation, we can see that, if we make certain assumptions about C, $\mathfrak{U}$ has two important properties:

Continuity of $\mathfrak{U}_C$

For each world w, $\mathfrak{U}_C(q, w)$ is a continuous function of q.

We ensure this by ensuring that C is continuous over the space of possible decision problems.

Strict Propriety of $\mathfrak{U}$

$\mathfrak{U}_C$ is strictly proper. That is, for any two probability functions $p \neq q$,

$$\text{Exp}_p(\mathfrak{U}_C(p)) = \sum_{w \in W} p(w)\mathfrak{U}_C(p, w) > \sum_{w \in W} p(w)\mathfrak{U}_C(q, w) = \text{Exp}_p(\mathfrak{U}_C(q))$$

So p expects itself to do better as a guide to action than it expects any other credence function to do.

Without any restrictions on C, we have

$$\begin{aligned}\text{Exp}_p(\mathfrak{U}_C(q)) = \sum_{w \in W} p(w)\mathfrak{U}_C(q, w) = \\ \sum_{w \in W} p(w) \int_{\mathcal{D}} a^q_D(w) dC = \\ \int_{\mathcal{D}} \sum_{w \in W} p(w) a^q_D(w) dC = \int_{\mathcal{D}} \text{Exp}_p(a^q_D) dC\end{aligned}$$

Now, by the definition of a^p and a^q,

$$\text{Exp}_p(a^p_D) \geq \text{Exp}_p(a^q_D)$$

So we have

$$\text{Exp}_p(\mathfrak{U}_C(p)) \geq \text{Exp}_p(\mathfrak{U}_C(q))$$

What's more, if there is a set $\mathcal{D}_{p,q} \subseteq \mathcal{D}$ of decision problems to which C assigns positive credence such that $a^p_D \neq a^q_D$ for all D in $\mathcal{D}_{p,q}$,

$$\int_{\mathcal{D}} \text{Exp}_p(a^p) dC > \int_{\mathcal{D}} \text{Exp}_p(a^q) dC$$

and

$$\text{Exp}_p(\mathfrak{U}_C(p)) > \text{Exp}_p(\mathfrak{U}_C(q))$$

So $\mathfrak{U}_C$ is strictly proper if there is such a set $\mathcal{D}_{p,q}$ for any two probabilistic credence functions p and q.

We can now give the second pragmatic argument for Bayes' Rule. The setup is very similar to before. As before, $t_1, \ldots, t_n$ are the times in your epistemic life at which you'll receive new evidence, and t_0 is the time at which you have your prior. For each $1 \leq i \leq n$:

- $\mathcal{E}_i$ is the partition from among which your total evidence at t_i will come;

- $p^{\alpha}_{w,i}$ is the credence function you will reach in world w by time t_i if you use α to successively update your credences on the total bodies of evidence you'll have at the various times in that world leading up to t_i;
- $\mathcal{D}_i$ is the set of possible decision problems you might face at t_i;
- C_i is the objective chance function over $\mathcal{D}_i$ that determines how likely you are to face a particular decision problem in that set at t_i;
- $0 < \lambda_i < 1$ is the weight that records how much you care about the pragmatic utility your credences obtain for you at time t_i.

Then the pragmatic utility of a prior p together with an updating rule α is

$$\mathfrak{U}((p, \alpha), w) = \lambda_0 \mathfrak{U}_{C_0}(p, w) + \lambda_1 \mathfrak{U}_{C_1}(p^{\alpha}_{w,1}, w) + \ldots + \lambda_n \mathfrak{U}_{C_n}(p^{\alpha}_{w,n}, w)$$

But then it is possible to show the following:

Theorem 3 (Longitudinal pragmatic dominance argument) *Suppose each $\mathfrak{U}_{C_i}$ is continuous and strictly proper. Then:*

(I) *If p is your prior and α is not Bayes' Rule, then there is a prior q such that, for all w,*
$$\mathfrak{U}((p, \alpha), w) < \mathfrak{U}((q, \beta), w)$$

(II) *If p is your prior, there is no q, α such that, for all w,*
$$\mathfrak{U}((p, \beta), w) < \mathfrak{U}((q, \alpha), w)$$

This is a straightforward generalization of the mathematical result at the heart of Briggs and Pettigrew's paper, even though we here consider $\mathfrak{U}$ as a measure of pragmatic utility, while Briggs and Pettigrew consider it a measure of accuracy (Briggs & Pettigrew, 2020). So, if you plan to update at any point in your epistemic life using any updating rule other than Bayes' Rule, then there is an alternative prior you might have such that you would obtain greater pragmatic utility over the course of your epistemic life if you were to adopt that alternative prior and update it using Bayes' Rule. And if you plan always to update by Bayes' Rule, that won't happen.

The upshot: not only does Brown's result show that Bayes' Rule is the updating rule that your prior expects to obtain for you the most pragmatic utility; it is also the only rule that renders your priors and posteriors, taken together, undominated when their value at a world is measured by the expected utility of the actions they'll lead you to choose. These two arguments address slightly different issues. Brown's argument imagines that you have a prior and you are in the market for an updating rule, and it tells you how to choose that. It tells you to choose Bayes' Rule. The argument inspired by Savage's remarks and based on Briggs and Pettigrew's theorem, on the other hand, imagines that you have yet to pick any of your credal furniture and you are therefore shopping around for both prior and updating rule. It tells you to pick a prior and Bayes' Rule. Together, they are a formidable obstacle to any claim that rules other than Bayes' Rule bring pragmatic benefits.

1.6 Getting to the truth faster

Nonetheless, Douven thinks there is a pragmatic virtue of the explanationist's rule that might save it from irrationality. He does not consider the two pragmatic arguments just described, so we can't know whether he thinks those virtues outweigh the flaws those arguments identify, but let's consider the matter ourselves.

In short, Douven claims that the explanationist's update rule ε might lead us to converge to the truth more quickly than Bayes' Rule β (Douven, 2013, 2021). He uses computer simulations of their performance to support this conclusion. The example he uses is a slight variation of the urn case we described above. Instead of three balls, there are ten in this urn; but, as before, all are coloured violet or green; and, as before, you know nothing of the distribution. There might be no violet and ten green (H_0), one violet and nine green (H_1), and so on up to nine violet and one green (H_9), and ten violet and no green (H_{10}). So H_i is the hypothesis that there are exactly i violet balls in the urn. In this context, the explanationist's updating rule works on a prior p like this:

$$p^{\varepsilon}_{E}(H_i) := \frac{p(E|H_i)p(H_i) + f(H_i, E)}{\sum_{j=1}^{n} \left(p(E|H_j)p(H_j) + f(H_j, E)\right)}$$

where $f(H_i, E)$ divides a fixed award c between the best explanations of E; that is, in this case, where H_i is a better explanation for E than H_j iff $p(E|H_i) > p(E|H_j)$ and at most two hypotheses can be the best explanations,

$$f_c(H_i, E) = \begin{cases} c & \text{if } P(E|H_i) > P(E|H_j) \text{ for all } j \neq i \\ \frac{1}{2}c & \text{if } P(E|H_i) = P(E|H_j) > P(E|H_k) \text{ for all } k \neq j, i \\ 0 & \text{otherwise} \end{cases}$$

Douven begins by making precise what he means by converging to the truth. He sets a threshold—in particular, he picks 0.9, though it seems plausible that we'd see the same phenomenon for other values. And he says that a credence function leads us to assert the truth of the hypothesis if it assigns it credence above that threshold. Then, for each possible composition of the urn, he asks a computer to simulate drawing a ball, looking at it, and replacing it 500 times in a row; and he asks it to do that 1,000 times. He also asks the computer to start with two uniform priors over the eleven hypotheses $H_1, \ldots, H_{10}$ about the urn's contents and then to update one of them after each draw using Bayes' Rule and to update the other after each draw using the explanationist's rule. Then, for each draw from the urn, he looks at the proportion of those 1,000 sequences of 500 draws at which updating using Bayes' Rule leads to credences in the true hypothesis that first cross the 0.9 threshold at that draw, and the proportion at which updating using the explanationist's rule leads to credences in the true hypothesis that first cross the 0.9 threshold at that draw. That is, for each draw, he asks how likely it is that Bayes' Rule 'gets it right' for the first time at that draw, and how likely it is that the explanationist's rule 'gets it right' for the first time at that draw. And he asks for which of the two updating rules does the draw at which it first 'gets it right' occur earliest. For each possible bias, his simulations show, it is the explanationist's rule. He then asks the same question but not for the draw with the highest chance of your rule getting it right, but for the draw with the highest chance of your rule getting it right *and remaining right*, where by that he means that the credence in the true hypothesis crosses the threshold *and stays there for the remainder of the draws*. And again it is the explanationist's rule.

One problem with this argument is that it isn't clear how impressive explanationism's victory is here. Consider, for instance, the following update rule. Whatever prior it is given, it recommends no change until you have seen ten draws from the urn. Then, if the first ten draws contain exactly i purple balls, assign credence 1 to hypothesis H_i, which says the urn contains exactly i purple balls, and 0 to all the others. Then never change your credences again, whatever further draws you witness. Now, for each bias H_i, run the same 1000 versions of the sequence of 500 draws from the urn. At which draw is it most likely this rule will lead to credence greater than 0.9 in the true hypothesis for the first time? And at which draw is it most likely to do that and then never fall below that threshold again? Well, it's the tenth draw for both, of course. Granted, it might not get it right at that toss. And indeed if it doesn't it never will. But if it's going to do it, it's going to do it then. Indeed, consider the rule that behaves exactly like this, but instead of assigning credence 1 to H_i when you see i purple balls among the first ten draws, it instead assigns 1 to H_1 if no purple balls are drawn, it assigns 1 to H_2 if exactly one purple ball is drawn, it assigns 1 to H_3 if exactly two purple balls are drawn, and so on until it assign 1 to H_{10} if exactly nine purple balls are drawn, and 1 to H_0 if ten purple balls are drawn. Again, it's most likely to get it right for the first time at the tenth draw, and most likely to get it right and stay right at that same draw. So these two rules outperform both Bayes' Rule and the explanationist's rule according to the measure that Douven introduces. But these are, of course, terrible rules. And that suggests that we should not care much about this measure of convergence to the truth.

Now, you might wonder what this all has to do with pragmatic utility. Here is Douven:

> [I]magine that the hypotheses concern some scientifically interesting quantity—such as the success rate of a medical treatment, or the probability of depressive relapse—rather than the bias of a coin, and the tosses are observations or experiments aimed at determining that quantity. Which researcher would not want to use an update rule that increases her chances of being in a position to make public a scientific theory, or a new medical treatment, before the (Bayesian) competition is? (Douven, 2021, 103)

Well, here is one answer: a researcher who wishes to update in a way that gives her posterior credences that her prior expects will lead her to the best choice when she uses them to face decisions. And Brown's expected pragmatic utility argument from above says that Bayes' Rule does this. We might suppose that the researcher will receive a stream of data, some parcel at each of a number of successive times. At each time, they'll face the same decision: make the new treatment public, or don't. The decision whether to announce a new treatment is always difficult. If you announce early and it's safe and effective, you prevent lots of suffering. If you announce early and it's safe but ineffective, you prevent no suffering, but equally you cause none, but perhaps you precipitate some loss of faith in medical science. And so on. So our researcher tries to quantify the utility of these different outcomes, assign credences to the different possibilities, and choose. But if they know that, at each time, they'll choose whether or not to make their treatment public by maximising their expected utility by the lights of their credences at that time, we know from the result of the previous section that they should update using Bayes' Rule. They will expect their future choices to have lower utility if they update in any other way.

This point is relevant also to a game Douven describes that pits the Bayesian against the explanationist, and is intended as another way to find out which converges to the truth faster. Again, in this game, balls are drawn and replaced from an urn containing ten balls, each violet or green; again, we don't know how many of each colour. After each draw, the Bayesian and the explanationist update the uniform prior using their favoured rule. And, at each point, they raise their hand if their credence in one of the hypotheses has risen above 0.9. The scoring system is then somewhat elaborate. Before we consider it, let's consider a slightly different, but much simpler system. If a player does not raise their hand on a given draw, they add 0 points to their total; if they do raise it and the hypothesis in which their credence is above 0.9 is true, they receive 1 point; if they raise it and the hypothesis is false, they lose 1 point. Now, I don't know which player will typically win this game, but it wouldn't surprise me at all if it is the explanationist. Nonetheless, I don't think this would tell against the Bayesian. After all, given that reward structure, raising you hand exactly when your credence in a hypothesis rises above 0.9 is just not what the Bayesian would choose to do. Rather, they would raise their hand whenever they would maximise expected utility by doing so, and that would be whenever their credence in a hypothesis rose above 0.5. After all, they would then have greater than 0.5 credence that they would obtain 1 point by raising their hand, and less than 0.5 credence that they would lose 1 point by doing so. And so raising their hand will have positive expected utility, which is greater than the guaranteed utility of 0 that keeping their hand lowered will have.

So it is no great criticism of Bayesianism that their credences lead them to have fewer points at the end of a game in which they wouldn't have chosen to play the way they were forced to play. One lesson from the literature on epistemic consequentialism is that, for any epistemic behaviour whatsoever, however rational, we can create a way of scoring epistemic states on which it performs poorly (Jenkins, 2007; Greaves, 2013; Ahlstrom-Vij & Dunn, 2018; Jenkins & Elstein, ta). Imagine this game: at successive stagess, I work through the propositions to which you assign credences and I ask you to report your credence in that proposition and in its negation. The points you receive at each turn is the difference between 1 and the sum of your credences in that turn's proposition and its negation. Many incoherent agents will win this game against a coherent agent. But that does not tell against probabilism.

Now, in Douven's version of this game, things are slightly more complicated. Nonetheless, the same problem arises. For him, the points you receive at a given draw depend not only on your credences, but also on your opponent's. Here are possibilities:

	Raise & Right	Raise & Wrong	Don't Raise
Raise & Right	$1, 1$	$0, 2$	$0, 1$
Raise & Wrong	$2, 0$	$0, 0$	$1, 0$
Don't Raise	$1, 0$	$0, 1$	$0, 0$

In the simpler version of the game, when each player considered whether raising their hand or keeping it lowered would maximise expected utility, they had only to consider their credences in the different hypotheses, since their score was dependent only on which of those was true. In this game, since the points they receive depend not only on which hypothesis is true but also on whether their opponent raises their hand as well, they must attend to their opponent's credences too. But the problem with the argument is the same: in each case, neither Bayesian nor explanationist would choose to raise their hand exactly when their credence in a hypothesis rises above

0.9. So it is no pragmatic argument against the Bayesian that their update rule, coupled with a non-Bayesian decision rule governing their play in this game, performs worse than the explanationist updating rule coupled with the same non-Bayesian decision rule.

2 Accuracy arguments for Bayes' Rule

So much, then, for the practical benefits of updating either using Bayes' Rule or using the explanationist's alternative or in some other way entirely. Alongside these practical arguments for Bayes' Rule, there are also purely epistemic arguments. These appeal not to how well the update credences guide action but how accurately they represent the world. The idea is this: just as full beliefs represent the world accurately by being true, credences represent the world more accurately the closer they are to the truth, where a credence in a true proposition is more accurate the higher it is and a credence in a false proposition is more accurate the lower it is. Assuming veritism, which says that accuracy is the fundamental source of purely epistemic value, we can give accuracy arguments for norms that govern credences by showing that, if you violate the norm, your credences are somehow suboptimal from the point of view of accuracy. Accuracy arguments have been given for probabilism (Joyce, 1998, 2009; Pettigrew, 2016a), Bayes' Rule (Greaves & Wallace, 2006; Leitgeb & Pettigrew, 2010; Briggs & Pettigrew, 2020), the principal principle (Pettigrew, 2013), the principle of indifference (Leitgeb & Pettigrew, 2010; Pettigrew, 2016b), and norms governing peer disagreement (Levinstein, 2015), higher-order evidence (Schoenfield, 2016), and the permissibility of rationality (Horowitz, 2014; Schoenfield, 2019), among many others. If the accuracy arguments for Bayes' Rule succeed, they tell against explanationism. So in this section we consider them and ask whether they do, indeed, succeed.

The two accuracy arguments for Bayes' Rule closely mirror the pragmatic arguments, and indeed they appeal to almost exactly the same mathematics. Just as the expected pragmatic argument showed that you will expect the credences demanded by Bayes' Rule to serve you best as a guide to future decision-making, so the expected epistemic utility argument shows that you will expect the credences demanded by that rule to most accurately represent the world. And just as the pragmatic utility dominance argument showed that updating your prior using anything other than Bayes' Rule will ensure there is an alternative prior and updating rule that, together, better serve as a guide to choice, so the epistemic utility dominance argument shows updating your prior by anything other than Bayes' Rule will ensure there is an alternative prior and updating rule that, together, more accurately represent the world.

At the heart of accuracy arguments for credal norms lie the measures of accuracy. Mathematically, they are very much like the measures of pragmatic utility we introduced above. Each such measure $\mathfrak{A}$ takes a credence function q and a possible world w and returns $\mathfrak{A}(q, w)$, which measures the accuracy of q at w. And indeed we assume of these measures the same properties that we showed our pragmatic utility measures had above:

Continuity of $\mathfrak{A}$

For each world w, $\mathfrak{A}(q, w)$ is a continuous function of q.

Strict Propriety of $\mathfrak{A}$

$\mathfrak{A}$ is strictly proper. That is, for any two probability functions $p \neq q$, p expects itself to be more accurate than it expects q to be.

$$\mathrm{Exp}_p(\mathfrak{A}(p)) > \mathrm{Exp}_p(\mathfrak{A}(q))$$

Both properties are assumed in nearly all discussions of epistemic utility, and I won't rehearse arguments in their favour here. There are many many accuracy measures that boast them both, but it will suffice to mention just the most popular pair. Suppose p is a credence function defined on a set of propositions $\mathcal{F}$. Given a proposition X in $\mathcal{F}$ and a possible world w, let $w(X) = 1$ if X is true at w and $w(X) = 0$ if X is false at w. Then:

Brier score First, define the quadratic scoring rule:

- $\mathfrak{q}(0, x) = x^2$
- $\mathfrak{q}(1, x) = (1 - x)^2$

Then define the Brier score of p at w:

$$\mathfrak{B}(p, w) = \sum_{X \in \mathcal{F}} \mathfrak{q}(w(X), p(X))$$

Additive log score First, define the logarithmic scoring rule:

- $\mathfrak{l}(0, x) = x$
- $\mathfrak{l}(1, x) = \log x + x$

Then define the additive log score of p at w:

$$\mathfrak{L}(p, w) = \sum_{X \in \mathcal{F}} \mathfrak{l}(w(X), p(X))$$

We then have the following mathematical results. The setup is the same as in the pragmatic case, except that we don't assume that there are any decision problems you might face using the credences you obtain from your prior and your updating rule. So $t_1, \ldots, t_n$ are the times in your epistemic future, and t_0 is the time at which you have your prior. At each time t_i your total evidence will come from the partition $\mathcal{E}_i$. $\mathfrak{A}_i$ measures the accuracy of your credences at time t_i. And $0 < \lambda_i < 1$ is the weight that encodes how much you care about the accuracy of your credences at t_i. Then we define the longitudinal accuracy of an updating rule α applied to a prior p as follows:

$$\mathfrak{A}(\alpha, w) = \lambda_1 \mathfrak{A}_1(p^\alpha_{w,1}, w) + \ldots + \lambda_n \mathfrak{A}_n(p^\alpha_{w,n}, w)$$

Then,

Theorem 4 (Longitudinal expected accuracy argument) *For any prior p,*

$$\sum_{w \in W} p(w)\mathfrak{A}(\beta, w) \geq \sum_{w \in W} p(w)\mathfrak{A}(\alpha, w)$$

And the inequality is strict if there is a world w and a time t_i such that (i) $p^\beta_{w,i} \neq p^\alpha_{w,i}$ *and* (ii) $p(w) > 0$.

This is the accuracy analogue of the expected pragmatic utility argument for Bayes' Rule. It generalizes the argument by Hilary Greaves and David Wallace that applies when you learn just once (Oddie, 1997; Greaves & Wallace, 2006). It shows that any prior will expect Bayes' Rule to produce more accurate credences than it expects any other rule to produce.

And we define the longitudinal accuracy of a prior and updating rule together as follows:

$$\mathfrak{A}((p, \alpha), w) = \lambda_0 \mathfrak{A}_0(p, w) + \lambda_1 \mathfrak{A}_1(p^\alpha_{w,1}, w) + \ldots + \lambda_n \mathfrak{A}_n(p^\alpha_{w,n}, w)$$

But then it is possible to show the following analogue of the pragmatic utility dominance argument for Bayes' Rule.

Theorem 5 (Longitudinal accuracy dominance argument) *Suppose each $\mathfrak{U}_i$ is continuous and strictly proper.*

(I) *If p is your prior and α is not Bayes' Rule, then there is a prior q such that, for all w,*

$$\mathfrak{A}((p, \alpha), w) < \mathfrak{A}((q, \beta), w)$$

(II) *If p is your prior, there is no q, α such that, for all w,*

$$\mathfrak{A}((p, \beta), w) < \mathfrak{A}((q, \alpha), w)$$

That is, if you are picking priors and updating rules together, only if you pick a prior and Bayes' Rule will you avoid accuracy dominance—that is, if you pick a prior and any other rule, there will be an alternative prior such that picking that and Bayes' Rule would produce more total accuracy in all worlds. This generalises Briggs and Pettigrew's argument (Briggs & Pettigrew, 2020).

Before we move on to Douven's simulation results concerning the accuracy of the two competing updating rules, it's worth noting how these results answer one of his concerns. He writes:

> The general problem for the inaccuracy-minimization approach this points to is that [minimizing accuracy] permits of a number of different interpretations. For instance, it can be interpreted as demanding that every single update minimize expected inaccuracy [...] or that every update minimize actual inaccuracy, or that every update be aimed at realizing the long-term project of coming to have a minimally inaccurate representation of the world, even if individual updates do not always minimize inaccuracy or expected inaccuracy. (Douven, 2021, 108)

In the longitudinal versions of the expected accuracy and accuracy dominance arguments we just described, we needn't weight all moments in the individual's epistemic life equally. If we are interested primarily in our long-run accuracy, we can give the lion's share of the weight to later points in our life. On the other hand, if we want to maximise getting quick results, perhaps in time for a big decision at the end of the week, we can shift it all to the times that lie within the next few days. But whatever we do, the results will be the same: Bayes' Rule is the uniquely best rule.[10]

Nonetheless, Douven thinks there is still a sense in which explanationism does better than Bayesianism from an accuracy point of view (Douven, 2021, Section 4.3). Again, he uses computer simulations to make his point. This time, he considers sequences of 1,000 draws from our urn. For each bias, he asks his computer to produce 1,000 such sequences and to update one uniform prior using Bayes' Rule after each draw, and one uniform prior using the explanationist's rule. After 100, 250, 500, 750, and 1000 draws, he compares the accuracy of the results of these two updating rules using the Brier score. He shows that, for any bias and any one of these five staging posts, the explanationist is more likely to have produced the more accurate credences of the two; and much more likely for the more extreme biases.

How can we reconcile this fact with the expected accuracy argument above? Well, as Douven himself notes, in those many cases where the explanationist does better than the Bayesian, they do only slightly better, while in the cases where the Bayesian prevails, they do a lot better. So, in expectation, Bayes' Rule is superior, even though in most cases, the explanationist's rule is better. Douven contends that, while this doesn't tell decisively in favour of the of the explanationist, it does undermine the claim that accuracy considerations tell decisively in favour of Bayes' Rule. If we care about being more accurate most of the time, rather than having greatest expected accuracy, we should be explanationists. And caring in this way is reasonable.

This is an interesting result, and it should give Bayesian's pause. But is it really reasonable to care about probability of comparative performance and ignore the distribution of absolute performance? Let's think how that

[10]This answers Douven's objection in footnote 74 (Douven, 2021).

pattern of caring would play out in a practical decision. Suppose, for instance, I think there are three possible outcomes of a new treatment for a particular medical condition: on the first, it alleviates the condition a very small amount; on the second, it alleviates the condition a small amount; on the third, it exacerbates the condition greatly and indeed produces new complications far far worse than the original condition. If I'm equally confident in these three possibilities, it doesn't seem at all reasonable to favour administering the drug, even though doing so is better in the majority of cases. Indeed, the very purpose of expected utility theory is to give us the means to navigate this sort of problem.

You might reasonably respond to this by pointing out that many decision theorists now think that maximizing expected utility isn't rationally mandated. Responding to examples like the Allais paradox, they hold that it is rationally permissible to use decision rules that give greater weight to worst-case scenarios than expected utility gives, and it is rationally permissible to use decision rules that give greater weight to best-case scenarios than expected utility gives (Allais, 1953; Quiggin, 1993). Perhaps the best example of such rules are given by Lara Buchak's risk-weighted expected utility theory (Buchak, 2013). I don't have a settled view on the rational permissibility of these rules. But I do know that they all obey the dominance or Pareto principle, which says that an option that is worse in every possible state of the world should be dispreferred. And of course the accuracy dominance result appeals not to expected utility theory, but only to such a dominance principle.

3 Updating in a social setting

So far, the setting for the stand offs between Bayes' Rule and explanationism has been the epistemology of individuals. That is, we have considered only the single solitary agent collecting evidence directly from the world and updating on it. But of course we often receive evidence not directly from the world, but indirectly through the opinions of others. I learn how many positive SARS-CoV-2 tests there have been in my area in the past week not my inspecting the test results myself but by listening to the local health authority. In their 2017 paper, 'Inference to the Best Explanation versus Bayes's Rule in a Social Setting', Douven joined with Sylvia Wenmackers to ask how Bayes' Rule and explanationism fare in a context in which some of my evidence comes from the world and some from learning the opinions of others, where those others are also receiving some of their evidence from the world and some from others, and where one of those others from whom they're learning might be me (Douven & Wenmackers, 2017). As for Douven's studies in the individual setting, Douven and Wenmackers conclude in favour of explanationism. Indeed, their conclusion in this case is considerably stronger than in the individual case:

> The upshot will be that if agents not only update their degrees of belief on the basis of evidence, but also take into account the degrees of belief of their epistemic neighbours, then the noted advantage of Bayesian updating [from (Douven, 2013)] evaporates and explanationism does better than Bayes's rule on every reasonable understanding of inaccuracy minimization. (Douven & Wenmackers, 2017, 536-7)

As before, I want to stick up for Bayes' Rule. As in the individual setting, I think this is the update rule we should use in the social setting.

In the individualistic cases we considered above, there's a single urn containing a particular number of violet and green balls. The individual draws and replaces balls one at a time, and updates their credences about the balls in the urn on the basis of those observations. In the social setting case, we assume each individual has an urn, and each of these urns has the same number of violet and green balls in it. So, again, the hypotheses in question are $H_0, \ldots, H_{10}$, where H_i says that every urn contains exactly i violet balls. As before, we assume each individual has the same uniform prior over the hypotheses, and obeys the Principal Principle. Douven and Wenmackers then assume that things proceed as follows:

- STEP (I) Each member draws and replaces a ball from their urn a certain number of times. This produces their worldly evidence for this round.
- STEP (II) Each then updates their credence function on this worldly evidence they've obtained. To do this, each member uses the same updating rule, either Bayes' Rule or a version of explanationism.
- STEP (III) Each then learns the updated credence functions of the others in the group. This produces their social evidence for this round.
- STEP (IV) They then update their own credence function on this social evidence by taking the average of their credence function and the other credence functions in the group that lie within a certain distance of theirs. The set of credence functions that lie within a certain distance of your own, Douven and Wenmackers call your *bounded confidence interval*.

They then repeat this cycle a number of times, and each time an individual begins with the credence function they reached at the end of the previous cycle.

Douven and Wenmackers use simulation techniques to see how this group of individuals perform for different updating rules used in step (II) and different specifications of how close a credence function must lie to yours in order to be included in your bounded confidence interval and thus in the average in step (IV). The updating rules they consider in step (II) are the explanationist's rule for different values of c, the reward that the rule distributes equally among the best explanations of the evidence. That is, for $c = 0$, this update rule is just Bayes' Rule, while for $c > 0$, it gives a little boost to whichever hypothesis best explains the evidence E, where providing the best explanation for a series of coin tosses amounts to making it most likely, and if two hypotheses make the evidence most likely, they split the boost between them. Douven and Wenmackers consider $c = 0, 0.1, \ldots, 0.9, 1$. For each rule, specified by c, they also consider different sizes of bounded confidence intervals. These are specified by the parameter δ. Your bounded confidence interval for δ includes each credence function for which the average difference between the credences it assigns and the credences you assign is at most δ. Thus, $\delta = 0$ is the most exclusive, and includes only your own credence function, while $\delta = 1$ is the most inclusive, and includes all credence functions in the group. Again, Douven and Wenmackers consider $\delta = 0, 0.1, \ldots, 0.9, 1$. Here are two of their main results:

(i) For each bias other than $p = 0.1$ or 0.9, there is an explanationist rule and bounded confidence interval (i.e. $c > 0$ and some specific δ) that gives rise to a lower average inaccuracy at the end of the process than Bayes' Rule with any bounded confidence interval (i.e. $c = 0$ and any δ).

(ii) There is an averaging explanationist rule and bounded confidence interval (i.e. $c > 0$ and $\delta > 0$) such that, for each bias other than $p = 0, 0.1, 0.9, 1$, it gives rise to lower average inaccuracy than Bayes' Rule with any bounded confidence interval (i.e. $c = 0$ and any δ).

Inaccuracy is measured by the Brier score throughout.

Now, you can ask whether these results are enough to tell so strongly in favour of explanationism, but that isn't my concern here. Rather, I want to focus on a more fundamental problem: Douven and Wenmackers' argument doesn't really compare Bayes' Rule with explanationism. Instead, it compares Bayes' Rule-for-worldly-data-plus-Averaging-for-social-data with explanationism-for-worldly-data-plus-Averaging-for-social-data. So their simulation results don't really impugn Bayes' Rule, because the average inaccuracies that they attribute to Bayes' Rule don't arise from it. They arise from using Bayes' Rule in step (II), but something quite different in step (IV). Douven and Wenmackers ask the Bayesian to respond to the social evidence they receive using a non-Bayesian rule, namely, Averaging. And Averaging lies far from Bayes' Rule.[11]

[11]For more on the tension between Averaging and Bayes' Rule, see (Dawid et al., 1995;

Why, then, do Douven and Wenmackers use Averaging rather than Bayes' Rule for step (IV)? Here is their motivation:

> [T]aking a convex combination of the probability functions of the individual agents in a group is the best studied method of forming social probability functions. Authors concerned with social probability functions have mostly considered assigning different weights to the probability functions of the various agents, typically in order to reflect agents' opinions about other agents' expertise or past performance. The averaging part of our update rule is in some regards simpler and in others less simple than those procedures. It is simpler in that we form probability functions from individual probability functions by taking only straight averages of individual probability functions, and it is less simple in that we do not take a straight average of the probability functions of all given agents, but only of those whose probability function is close enough to that of the agent whose probability is being updated. (Douven & Wenmackers, 2017, 552)

In some sense, they're right. Averaging or linear pooling or taking a convex combination of individual credence functions is indeed the best studied method of forming social credence functions. And there are good justifications for it: János Aczél and Carl Wagner and, independently, Kevin J. McConway, give a neat axiomatic characterization (Aczél & Wagner, 1980; McConway, 1981); and indeed Richard Pettigrew has argued that there are accuracy-based reasons to use it in particular cases (Pettigrew, 2019). The problem is that our situation in step (IV) is not the sort of situation in which you should use Averaging. Arguments for Averaging concern those situations in which you have a group of individuals, possibly experts, and each has a credence function over the same set of propositions, and you want to produce a single credence function that could be called the group's collective credence function. Thus, for instance, if I wish to give the SAGE group's collective credence that there will be a safe and effective SARS-CoV-2 vaccine by March 2021, I might take the average of their individual credences. But this is quite a different task from the one that faces me as the first individual when I reach step (IV) of Douven and Wenmackers' process. There, I already have credences in the propositions in question. What's more, I know how the other individuals update and the sort of evidence they will have received, even if I don't know which particular evidence of that sort they have. And that allows me to infer from their credences after the update at step (II) a lot about the evidence they receive. And I have opinions about the propositions in question conditional on the

Bradley, 2018; Dawid & Mortera, 2020).

different evidence my fellow group members received. And so, in this situation, I'm not trying to summarise our individual opinions as a single opinion. Rather, I'm trying to use their opinions as evidence to inform my own. And, in that case, Bayes' Rule is better than Averaging. So, in order to show that explanationism is superior to Bayes' Rule in some respect, it doesn't help to compare Bayes' Rule at step (II) + Averaging at step (IV) with explanationism at (II) + Averaging at (IV). It would be better to compare Bayes' Rule at (II) and (IV) with explanationism at (II) and (IV).

So how do things look if we do that? Well, it turns out that we don't need simulations to answer that question. We can simply appeal to the accuracy arguments we mentioned above: the expected accuracy argument for picking Bayes' Rule on the basis of your prior, and the accuracy dominance argument for picking a prior-rule pair where the rule is Bayes' Rule applied to the prior.

You might respond to this objection by arguing that applying Bayes' Rule at (IV) is all well and good if you are a computer or a robot, but it might require computation that is either not feasible for an ordinary person, or feasible but not worth their while. After all, it might seem to require a great deal of work to extract from those posteriors the evidence that gave rise to them and thus the evidence on which you are going to update using Bayes' Rule at (IV). Suppose each individual has drawn and replaced ten balls at (I). Then the possible evidence an individual might have received falls into eleven groups: those in which they drew and replaced no violet balls, those in which they drew and replaced one violet ball, and so on. Thus, for each of these possibilities, I would have to calculate at (IV) the posterior that they would have mandated. Only then could I compare those with the posteriors that my fellow group members have reported in order to find out what evidence they have. Surely it would be a lot easier to apply Averaging directly to the reported posteriors, even if by doing so I sacrifice some accuracy.

I agree. It would be a chore to extract that evidence. But thankfully this is not the only option. Thanks to a beautiful result due to (Baccelli & Stewart, ms), we can achieve the same effect by using geometric pooling instead of the linear pooling that Douven and Wenmackers use. Given a set of credence functions $p_1, \ldots, p_n$, their straight geometric pool is defined as follows:

$$\mathrm{GP}(p_1, \ldots, p_n)(w) = \frac{\prod_{i=1}^n p_i(w)^{\frac{1}{n}}}{\sum_{w' \in W} \prod_{i=1}^n p_i(w')^{\frac{1}{n}}}$$

That is, instead of taking the arithmetic mean of the credences in each world, we take their geometric mean and normalise. We then have Baccelli and Rush's central result:

Theorem 6 *For any prior p, if $E_1, \ldots, E_n$ are compatible pieces of evidence, then*

$$p(-|\bigcap_{i=1}^{n} E_i) = \mathrm{GP}(p(-|E_1), \ldots, p(-|E_n))$$

Thus, I needn't actually extract the evidence from the reported posteriors. I can simply apply an alternative pooling method at (IV). That will be equivalent to applying Bayes' Rule on the extracted evidence and thus has the same advantages when assessed for accuracy.

4 Choosing your intellectual trajectory

I'd like to finish by taking up a challenge that Douven lays down in passing. He writes:

> [I]n science, we rarely just happen across useful data. Typically, we must actively search for data, in the many areas of science that rely on experimentation even produce our data. Because our time is limited, as is our funding, we constantly have to make decisions as to which instruments (telescopes, microscopes, etc.) to construct, which expeditions to undertake, which experiments to run, and so on. Such decisions will be informed by which hypotheses we deem most promising. Had we deemed hypothesis H promising, and had we wanted to compare that with the hypothesis currently dominant in our field, we might have run a different set of experiments than we actually did, given that in fact we deemed H' more promising than H and were mainly interested in comparing H' with the received doctrine. Which hypothesis or hypotheses we deem most promising, and most worthy of spending our limited resources on, will at least in part depend on how probable they appear to us, compared to their most direct rivals. If (say) a Bayesian update makes H more probable than H', while the opposite will be the case if we update via some non-Bayesian update rule, then our decision to use one of these rules may put us on a very different research path with very different downstream consequences than if we had decided to use the other rule. Which of these paths will eventually lead us to have the more accurate representation of the world will have nothing to do with which of the rules minimizes expected inaccuracy of the piece of evidence now lying before us. (Douven, 2021, 107)

Douven is right, of course. What credences our update rule bestows on us will determine not just how we'll choose when faced with practical decisions, such as whether or not to publicly announce a new medical treatment, but also how we'll choose when faced with an intellectual decision, such as which experiment to run next, which hypothesis to pursue, and so on. So, even if we focus only on our purely epistemic goal of accuracy, we'll want credences that lead us to choose how to gather evidence in a way that maximises the accuracy we obtain after we choose them, perform them, and update on the results. But of course we have a ready-made answer to that challenge. The expected pragmatic argument for Bayes' Rule applies just as well when the options between which you'll choose after updating are whether to pursue one hypothesis or another, or whether to conduct this experiment or that one, and when the utilities that attach to those hypotheses at the different possible worlds are given by the accuracy of the credence functions you'll end up with if you do pursue that intellectual trajectory.

5 Conclusion

Credences play at least two roles in our lives. They guide our actions and they represent the world. When we decide how we'll update our credences in response to evidence, we should pick a rule that leads to credences that play those roles well. Bas van Fraassen argued that there is a particular way in which updating other than by Bayes' Rule leads to credences that guide action poorly. But, as Igor Douven points out, it's a pretty weak argument. Fortunately for Bayes' Rule, there are now much stronger arguments in its favour. Some focus on the pragmatic value of credences, others on the epistemic value. Together, they allow us to answer the sorts of objections to Bayes' Rule that Igor Douven has raised using simulations of the two ways of updating.

This is good news for Bayesians. Is it bad news for those who think that inference to the best explanation is an important and correct rule of inference? I think not. If I have convinced you at all, it can only be that any inference to the best explanation should not require a boost to hypotheses beyond what can be incorporated into a prior credence function and what Bayes' Rule already gives them. But leaves a lot of room for inference to the best explanation to play a role within the confines of probabilism and Bayes' Rule.

References

Aczél, J., & Wagner, C. G. (1980). A characterization of weighted arithmetic means. *SIAM Journal on Algebraic Discrete Methods*, *1*(3), 259–260.

Ahlstrom-Vij, K., & Dunn, J. (Eds.) (2018). *Epistemic Consequentialism*. Oxford: Oxford University Press.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica*, *21*(4), 503–546.

Baccelli, J., & Stewart, R. (ms). Support for Geometric Pooling. Unpublished manuscript.

Bradley, R. (2018). Learning from Others: Conditioning versus Averaging. *Theory and Decision*, *85*(1), 5–20.

Briggs, R. A. (2009). Distorted Reflection. *Philosophical Review*, *118*(1), 59–85.

Briggs, R. A., & Pettigrew, R. (2020). An accuracy-dominance argument for conditionalization. *Noûs*, *54*(1), 162–181.

Brown, P. M. (1976). Conditionalization and expected utility. *Philosophy of Science*, *43*(3), 415–419.

Buchak, L. (2013). *Risk and Rationality*. Oxford, UK: Oxford University Press.

Climenhaga, N. (2017). Inference to the Best Explanation made Incoherent. *The Journal of Philosophy*, *114*(5), 251–73.

Dawid, P., DeGroot, M. H., & Mortera, J. (1995). Coherent Combination of Experts' Opinions. *Test*, *4*(2), 263–313.

Dawid, P., & Mortera, J. (2020). Resolving some contradictions in the Theory of Linear Opinion Pools. *Theory and Decision*, *88*, 453–456.

de Finetti, B. (1937 [1980]). Foresight: Its Logical Laws, Its Subjective Sources. In H. E. Kyburg, & H. E. K. Smokler (Eds.) *Studies in Subjective Probability*. Huntingdon, N. Y.: Robert E. Kreiger Publishing Co.

Douven, I. (1999). Inference to the Best Explanation made coherent. *Philosophy of Science (Proceedings)*, *66*(S424–S435).

Douven, I. (2013). Inference to the Best Explanation, Dutch Books, and Inaccuracy Minimization. *Philosophical Quarterly*, *63*, 428–44.

Douven, I. (2017). Abduction. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, Summer 2017 ed.

Douven, I. (2021). *The Art of Abduction*. Cambridge, Mass.: MIT Press.

Douven, I., & Wenmackers, S. (2017). Inference to the Best Explanation versus Bayes's Rule in a Social Setting. *British Journal for the Philosophy of Science*, *68*(2), 535–570.

Good, I. J. (1967). On the Principle of Total Evidence. *The British Journal for the Philosophy of Science*, *17*, 319–322.

Greaves, H. (2013). Epistemic Decision Theory. *Mind*, *122*(488), 915–952.

Greaves, H., & Wallace, D. (2006). Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility. *Mind*, *115*(459), 607–632.

Horowitz, S. (2014). Immoderately rational. *Philosophical Studies*, *167*, 41–56.

Jenkins, C. I., & Elstein, D. (ta). The Truth Fairy and Epistemic Consequentialism. In N. J. L. L. Pedersen, & P. Graham (Eds.) *Epistemic Entitlement*. Oxford University Press.

Jenkins, C. S. (2007). Entitlement and Rationality. *Synthese*, *157*, 25–45.

Joyce, J. M. (1998). A Nonpragmatic Vindication of Probabilism. *Philosophy of Science*, *65*(4), 575–603.

Joyce, J. M. (2009). Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief. In F. Huber, & C. Schmidt-Petri (Eds.) *Degrees of Belief*. Springer.

Lange, M. (2020). Putting explanation back into "inference to the best explanation". *Noûs*.

Leitgeb, H., & Pettigrew, R. (2010). An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy. *Philosophy of Science*, *77*, 236–272.

Levinstein, B. A. (2015). With All Due Respect: The Macro-Epistemology of Disagreement. *Philosophers' Imprint*, *15*(3), 1–20.

Levinstein, B. A. (2017). A Pragmatist's Guide to Epistemic Utility. *Philosophy of Science*, *84*(4), 613–638.

Lewis, D. (1980). A Subjectivist's Guide to Objective Chance. In R. C. Jeffrey (Ed.) *Studies in Inductive Logic and Probability*, vol. II. Berkeley: University of California Press.

Lewis, D. (1999). Why Conditionalize? In *Papers in Metaphysics and Epistemology*, (pp. 403–407). Cambridge, UK: Cambridge University Press.

Lipton, P. (2004). *Inference to the Best Explanation*. London: Routledge, 2nd ed.

McConway, K. J. (1981). Marginalization and linear opinion pools. *Journal of the American Statistical Association*, *76*, 410–414.

McGrew, T. (2003). Confirmation, heuristics, and explanatory reasoning. *British Journal for the Philosophy of Science*, *54*, 553–67.

Oddie, G. (1997). Conditionalization, Cogency, and Cognitive Value. *British Journal for the Philosophy of Science*, *48*, 533–41.

Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in the History and Philosophy of Science*, *31*, 691–710.

Pettigrew, R. (2013). A New Epistemic Utility Argument for the Principal Principle. *Episteme*, *10*(1), 19–35.

Pettigrew, R. (2016a). *Accuracy and the Laws of Credence*. Oxford: Oxford University Press.

Pettigrew, R. (2016b). Accuracy, Risk, and the Principle of Indifference. *Philosophy and Phenomenological Research*, *92*(1), 35–59.

Pettigrew, R. (2019). On the Accuracy of Group Credences. In T. S. Gendler, & J. Hawthorne (Eds.) *Oxford Studies in Epistemology*, vol. 6. Oxford: Oxford University Press.

Pettigrew, R. (2020). *The Dutch Book Argument*. Elements in Decision Theory and Philosophy. Cambridge, UK: Cambridge University Press.

Quiggin, J. (1993). *Generalized Expected Utility Theory: The Rank-Dependent Model*. Kluwer Academic Publishers.

Ramsey, F. P. (1926 [1931]). Truth and Probability. In R. B. Braithwaite (Ed.) *The Foundations of Mathematics and Other Logical Essays*, chap. VII, (pp. 156–198). London: Kegan, Paul, Trench, Trubner & Co.

Savage, L. J. (1954). *The Foundations of Statistics*. John Wiley & Sons.

Savage, L. J. (1971). Elicitation of Personal Probabilities and Expectations. *Journal of the American Statistical Association*, *66*(336), 783–801.

Schervish, M. J. (1989). A general method for comparing probability assessors. *The Annals of Statistics*, *17*, 1856–1879.

Schoenfield, M. (2016). An Accuracy-Based Approach to Higher-Order Evidence. *Philosophy and Phenomenological Research*, *96*(3), 690–715.

Schoenfield, M. (2019). Permissivism and the Value of Rationality: A Challenge to the Uniqueness Thesis. *Philosophy and Phenomenological Research*, *99*(2), 286–297.

van Fraassen, B. C. (1989). *Laws and Symmetry*. Oxford: Oxford University Press.

Weisberg, J. (2009). Locating IBE in the Bayesian framework. *Synthese*, *167*, 125–143.