Expert deference as a belief revision schema

Joe Roussos
Institute for Futures Studies, Stockholm
joe.roussos@iffs.se

October 28, 2020

Abstract

When an agent learns of an expert's credence in a proposition about which they are an expert, the agent should defer to the expert and adopt that credence as their own. This is a popular thought about how agents ought to respond to (ideal) experts. In a Bayesian framework, it is often modelled by endowing the agent with a set of priors that achieves this result. But this model faces a number of challenges, especially when applied to non-ideal agents (who nevertheless interact with ideal experts). I outline these problems, and use them as desiderata for the development of a new model. Taking inspiration from Richard Jeffrey's development of Jeffrey conditioning, I develop a model in which expert reports are taken as exogenous constraints on the agent's posterior probabilities. I show how this model can handle a much wider class of expert reports (for example reports of conditional probabilities), and can be naturally extended to cover propositions for which the agent has no prior.

1 Introduction

Consider this example of an agent receiving testimony from an expert.

> *Hurricane.* You open your weather app and see, to your complete surprise, that there is a 30% chance that London will be struck by a hurricane on Thursday.

In this essay, I explore a very common idea that philosophers have had about cases like this: the fact that this information comes from an *expert* means that the content of their report, that there is a 30% probability of a hurricane in London on Thursday, ought to *directly* influence your credence in that proposition. Philosophers have typically expressed this thought in the context of ideal agents responding to the testimony of ideal experts (examples of which include chance, or one's future credences), but I am interested in somewhat more realistic cases like *Hurricane*.

I will develop a model of expert deference that works for more realistic agents, who nevertheless find it useful to make the idealisation that the expert is worthy of deference. This partial de-idealisation is of interest when thinking of cases of policymaking in which non-experts take input from various scientists when formulating policies and making decisions. To non-expert policymakers, the testimony of scientists may need to be taken on trust. While there may be complex procedures in place to secure the best expert advice, or to aggregate advice from different experts, when the policymaker is faced with forming their own beliefs or making policy decisions, it seems plausible to represent what they do as deference.

Why do we need an account of partially de-idealised expert deference? As I will show, the standard, idealised Bayesian accounts make unreasonable demands of the deferring agent.

How reasonable the theory is—for real or ideal agents—depends in part on how we interpret its normative demands. At its core, Bayesianism is committed to two norms: that one ought to have probabilistic partial beliefs, and that one ought to update those beliefs by conditioning on one's evidence. Often, they are assumed to be *evaluative norms*: they are features of a good believer. Evaluative norms needn't entail anything about action: a good spring day is cloudless and fresh; these are evaluative standards for assessing days *qua* spring days, but do not directly bear on the actions of any agents. But the Bayesian norms are also sometimes taken to be *action-guiding*. This is especially so in the Bayesian statistics literature where investigations of, for example, expert testimony, include discussions of how real agents might carry out Bayesian processes.[1]

I am interested in policymakers facing actual cases of expert disagreements, and my aim is to contribute to advancing their practice. Therefore, I am interested in guidance for action. There is, of course, a link between evaluative and prescriptive norms. Evaluative norms can give rise to prescriptive ones: rules for baking bread are created with good bread as their target. In the other direction, facts about what one *can do* may constrain standards for evaluating one's goodness. If the prescriptions associated with an evaluative standard are impossible, this may require a revision of that standard.

I begin my discussion not with deference, but with the orthodox Bayesian approach often called "supra-Bayesianism". I argue in section 2 that it is not, in fact, a credible solution by outlining four problems it faces. Two concern the limitations of real agents, and two concern ways in which it fails to make use of the idealness of the experts. (There is a long history of discussions of supra-Bayesianism, so these problems are not new.[2]) In section 3 I turn to expert deference, which is a supplementary principle added

---

[1] For an explicit discussion of these two kinds of norms in epistemology see (Simion et al., 2016, S4.1), and for a similar discussion in decision theory see (Buchak, 2013, Ch 1) and (Thoma, 2019). The Bayesian statistics papers referenced in this section almost all have a prescriptive element, but for a particularly clear example see (French, 1980).

[2] This section, and indeed this paper, is not intended as a complete survey of the literature on Bayesian approaches to expert disagreement and, where I do survey the literature, my review is partial to philosophy. There is a Bayesian statistics literature on the topic of expert testimony covering both supra-Bayesianism and deference, and I engage with it here only partially. Part of the difficulty in using that literature arises from the difference in focus. Statistics papers often *assume* that orthodox Bayesianism is the right norm, while I wish to evaluate that claim. They work through how a real agent might reason in the kinds of cases under consideration, and regularly assume a particular form for the agents' priors and likelihoods (i.e., assuming particular distributions) in order to make progress. While valuable for building understanding of

to Bayesianism to make it appropriately sensitive to expert testimony. I start with the traditional way that deference has been formulated by Bayesians: as a constraint on the priors of an ideal Bayesian agent. I argue that this formulation falls prey to some of the problems I introduced for supra-Bayesianism, and that has two additional problems of its own. First, it arbitrarily distinguishes one kind of expert report as worthy of deference: reports on the unconditional probability of a proposition. Second, it cannot deal with expert *disagreement* (i.e., multiple, contrary testimonies).

In section 4 I present a new development of the expert deference idea, in a manner that is (a) better suited to my non-ideal theory approach, and (b) can resolve most of the problems I introduce for existing accounts. I call the new approach "expert deference as a belief revision schema," and draw on the theory of probabilistic belief revision in order to develop it. Notably, the new approach can handle a much wider class of expert reports than the orthodox model of deference. I show that this new approach does not fall prey to most of the problems that bedevil supra-Bayesianism and expert deference as a constraint on priors. The two most difficult problems are expert disagreement and deferring to propositions for which the agent has no priors. I show in section 5 how my model works naturally with one approach to rational awareness growth, thereby allowing for deference to unfamiliar propositions. In section 6 I reflect on disagreement, and why it cannot be successfully resolved in a deference-type model.

2 Supra-Bayesianism

We begin with simple Bayesian orthodoxy: when you hear the expert report, you update your beliefs by conditioning on what you have learned.

The setup is standard: your probabilities P are defined on an algebra of propositions Ω. P includes a prior for H, the proposition that a hurricane will hit London on Thursday, and when you look at your app you learn the proposition that the weather app says "there is a 30% chance of a hurricane". Such propositions about others' credences are denoted with corner quotes, i.e., $\ulcorner W(H) = 0.3 \urcorner$. When there is no ambiguity, I will use $\ulcorner W \urcorner$ as shorthand for $\ulcorner W(H) = 0.3 \urcorner$. Upon learning that $\ulcorner W \urcorner$, your posterior probability for H is:

$$Q(H) = P(H | \ulcorner W(H) = 0.3 \urcorner) = P(\ulcorner W \urcorner | H) \frac{P(H)}{P(\ulcorner W \urcorner)}$$

So, your posterior for H depends on your prior for H, your prior for hearing this report $\ulcorner W \urcorner$, and your prior likelihood for the hurricane, given the report.

This answer, though perfect Bayesian orthodoxy, faces a number of challenges.[3] I begin with two challenges for real agents.

Bayesianism and its implications, they are rarely directly concerned with my topic here.

[3] The label "supra-Bayesianism" comes from Keeney and Raiffa (1976). It has been much discussed in the Bayesian statistics literature (see Genest and Zidek, 1986), and I do not claim that these problems are without possible responses. In particular, much work has been done on how to make it more tractable in cases where particular symmetries, or known distributions, simplify the updating required. Lindley (1982) notes cases in which it reduces to the very simple expert deference. Others have studied when it reduces to averaging. French (1980) is an early analysis of how thinking through the procedure a real agent might use to enact supra-Bayesianism can generate plausible simplifications.

1. *Cognitive Burden:* There are a great many experts in the world, a myriad of propositions they might report on, and a continuum of reports they could make on each. Experts and their reports may have complex dependencies on one another. Ω must therefore be a very rich algebra indeed, and the range and granularity of the judgements the agent is required to make are breathtaking. Any procedure to enact (even approximately) the supra-Bayesian answer is therefore extremely cognitively demanding for any real agent (this has been extensively discussed; see e.g., the comments on supra-Bayesianism in Genest and Zidek (1986)). It is no failing of rationality not to be able to accomplish this procedure.

This objection concerns what real agents can achieve. It is a problem that targets the status of supra-Bayesianism as a norm, either for action or evaluation that is reflective of the agent's capacities. A Bayesian might quibble: a rational agent simply does behave in a manner that is (representable as) complying with the diktats of Bayesianism. In this case, that involves supra-Bayesian updating on the reports of any experts giving testimony. But it is no more assumed that agents actually perform these calculations than it is assumed that one does mental trigonometry when catching a ball. Catching a ball *amounts to* calculating a trajectory and performing a sequence of movements such that one's hands intersect with that trajectory, but that statement can be true independently of what is going on in the mind of the agent doing the catching.

I acknowledge that in general there is no requirement that agents be consciously attending to the beliefs we are discussing, or consciously perform any calculations. The Bayesian can say that a rational agent has cognitive architecture that accomplishes the belief changes prescribed by Bayesianism *somehow*. Perhaps if we supplement this with a "low cost" analysis of belief such as dispositionalism, then the Bayesian can insist that there are no onerous cognitive demands on agents. As long as they end up with the right dispositions, they're rational. The costliness of performing the calculations to determine the right posterior belief in the representation can completely decouple from the costliness of accomplishing the actual belief change in the agent's cognitive architecture.

This is a reasonable response if one's aim is to construct a theory of rationality that interprets the rational behaviour of real agents. Whether the response works for someone who views the norms of Bayesianism as evaluative will depend on the plausibility of the underlying account of partial belief, an issue I will not address. But if one takes Bayesianism to be action-guiding, then it *is* a flaw if no real agent could reproduce the obvious procedure for determining which are the right actions (i.e., permissible posterior beliefs) and if no alternative procedure for that determination is provided. As I am interested in generating normative guidance for policymakers facing expert disagreement, I view it as such a problem.

2. *Awareness:* In the setup of the example, I stressed your surprise at hearing this report. London doesn't get hurricanes, and so it is quite plausible that you'd never considered *H* before, much less the proposition that the weather app would report precisely 30% as the probability for it!

It is almost constitutive of the notion of expertise that experts will regularly discuss things that novices are unfamiliar with. We should therefore expect that many expert statements will be about propositions that the non-experts were previously unaware of. In such cases, it is simply implausible that the agent has any views on these matters at

4

all—yet the supra-Bayesian view insists that they have priors for them.

When I talk of "awareness" I do not mean that agents must actively reflect on something, or that they pay attention to it. The common English phrase "I've never heard of that before" corresponds well to what I mean by unawareness. An extreme case would be an agent's relation to a proposition involving a concept that they do not have—perhaps, your present relation to some esoteric statement about theoretical particle physics. Between these two extremes—mere inattention and conceptual lack—lie many cases where the agent does not meet common criteria for belief. They may lack any relevant disposition to act, or stand in no relation to the relevant mental representation. If you know nothing about South African politics, then you know nothing about the Democratic Alliance or their electoral hopes in the city of Johannesburg. When confronted for the first time with my claim that the Democratic Alliance has no future in Johannesburg, you are in a state of unawareness.

Nevertheless, supra-Bayesianism requires agents to have attitudes towards all propositions. To stress the implausibility of this, let me point out that in order to apply supra-Bayesianism generally we must require that you have credences for *any* report (i.e., any $x \in [0, 1]$) on *any* proposition made by *any* expert.[4] If, like me, you think that there are many combinations of expert, proposition, and report, toward which even the most rational agents will have no attitudes, then you think Awareness is a problem for supra-Bayesianism.

Now recall the setup that I am interested in: non-ideal agents interacting with (relatively) *ideal experts* (I will refine this description in section 3). In this setup, there are two further concerns with supra-Bayesianism.

3. *Relevance of Priors:* One might reasonably ask: what do you know about hurricanes, anyway? The reliance of $Q(H)$ on $P(H)$ strikes many as problematic: surely it is rational to jettison your ignorant prior in face of reliable expert testimony? Similarly, why should $Q(H)$ depend on how likely you think this expert is to report precisely 30% as their probability for H? What do you know about hurricane prediction, or the methods of this or that forecaster?

4. *Sensitivity to Testimony:* The complementary problem to supra-Bayesianism's over-sensitivity to your priors is that it is *under*-sensitive to the actual content of the expert's report. $Q(H)$ isn't a function of the expert's reported credence! It is instead a function of your priors that the expert will *report*, in this case, 30%. This is because the supra-Bayesian procedure is just the general Bayesian answer to every learning experience. But it seems like there's something different going on in the case of expert testimony: you're receiving information that is *directly* relevant to your credence in H, in a way that is unmediated by your credences in the learned proposition.

To this, the Bayesian can reply the generality of the approach is a strength and that, in this case, that strength manifests as a sensitivity to the particularities of the report. What you condition on is the fact that this particular expert made this precise report in the manner and context that they did. This complex fact is the right thing for the agent

---

[4]One might also worry that this demand, taken literally, means that the simple model above won't work. Experts report probability values, and so these reports are themselves continuous random variables. Strictly speaking your prior for $\ulcorner W(H) = x \urcorner$ should thus be zero, for any x. I won't dwell on this problem, as the issues it raises aren't core criticisms of supra-Bayesianism and I believe that a more complex model could work around it.

to respond to, the Bayesian continues, as it allows their response to depend on what they think about this expert, what they know about the circumstances of the report, and so forth. The precise value reported is just one many features of the learning experience to which the agent should be responsive.

These additional factors are crucial when the agent makes a simultaneous assessment of the reported proposition and the expert, in response to hearing the expert's report. But in many cases, when experts report information within their domain, and when those reports raise no suspicions about their expertise, the impact of the evaluation of the expert is negligible. It is these cases that my real agent/ideal expert model targets. And it is in these cases that Relevance of Priors and Sensitivity to Testimony are pressing concerns. To double down on the flexibility of supra-Bayesianism in such cases is to deny that there is anything special about expert testimony, as opposed to general learning experiences. But there does seem to be something particular about a case of expert testimony, especially when we consider ideal experts: this learning experience contains as one feature of it *a number* that our credence ought to be close to. Without supplementation, supra-Bayesian can't guarantee this sensitivity.

Taken together I regard these problems as damning enough that I want an alternative. As outlined above, I am concerned about real agents; in particular their ability to extract action guidance or a comprehensible normative standard in a variety of cases. I will therefore consider alternatives that are more limited than supra-Bayesianism, but easier to enact.

3   Expert deference as a constraint on priors

One intuitive thought about expert testimony is that, under the right conditions, laypeople should defer to it. Deference means adopting the expert's testimony into your beliefs. For example if you are interested in whether it will rain, you should ask the (ideal) weather forecaster and believe what they tell you: if they say the chance of rain is 40%, you should believe that and thus set your own credence in rain to 40%.

As I am interested in probabilistic opinions, I will think of deference in this way, as taking an expert's probabilities on as your own. Deference is so common as a thought about expert testimony that some have taken it to be the definition of an expert: Gaifman (1988, 193) defines an expert as someone for whom "the mere knowledge of. . . [their] assignment will make the agent adopt it as his subjective probability." This definition is common in Bayesian statistics (for a contemporaneous usage see DeGroot, 1988), and has been adopted in philosophy by, e.g., Joyce (2007) and Elga (2007).[5] This definition is somewhat unhelpful if one is interested in *identifying* which people are experts, but it does highlight the centrality of the deference idea.

Deference may strike the reader as a rather extreme idealisation. Should a real person defer to a real expert? To start with the descriptive: arguably, we do so all the time. The presumption that people speak truly in ordinary conversational contexts results in

_____

[5]There are alternate definitions of expert out there. For example, Easwaran et al. (2016) define experts as reliable witnesses. For them, P^1 is an expert for P, in some domain D, when the following holds: for any $X \in D$, when $P^1(X) > P(X)$, P takes P^1's credence in X as evidence for X and raises their credence. The same applies to lower credences as evidence against.

something like deference: taking propositions to be *true* because someone reports them to be true. When students learn a science they accept the material in textbooks in a way that is close to deference. Professional scientists do this too: experimenters working with radioactive materials do not assess molecular half-lives for themselves, they look them up. More prosaically, I often look at the weather forecast and act accordingly. Later, if someone asks me whether it will rain, I often quote the reported chance of rain.

Regardless of whether we *do* defer, one might reasonably worry whether we ought to. One way to see the idealisation involved in deference is to note its link to calibration. Consider a case where the relevant reports concern something that occurs multiple times.[6] A probabilistic report is called *calibrated* when propositions X assessed to have probability x turn out to be true $x\%$ of the time. Put another way, if we collect all of the predictions that something is $x\%$ probable then, if those predictions are calibrated, the proportion of events that turn out to be true will be x. Calibration is a statement linking the report $\ulcorner W(X) = x \urcorner$ with the actual frequency of occurrence. If we know a predictor is calibrated, then we can project these frequencies into the future—calibrated predictions are *chance signals*. If our agent obeys a principle that known chances should guide her credences, then she ought to set her posterior degrees of belief to match the calibrated prediction.[7]

When we defer to *uncalibrated* experts, the idealisation we're making is to treat them as if they were calibrated. But again, this can be reasonable: experts are often the best stand-in that we have for chance signals, and we should still expect to do better by deferring than we would by holding to our prior credences. While we know that they aren't perfectly calibrated, we don't know the way in which they fail to match the frequencies. Deference is a simple, easy to implement, procedure to improve our own predictive performance. Now, I am interested in *idealised* experts, so why bother with the foregoing defence of deference? I intend it as a defence of making the idealisation that experts are worthy of deference in real cases.

Let us turn to how expert deference is modelled. The classic approach is to model deference as a set of priors that the agent has. Recall *Hurricane*, where your weather app tells you that there is a 30% chance that a hurricane will strike London on Thursday. We can model you as deferring to this expert by equipping you with what I will call a "deference prior", a set of prior credences defined by $P(H|\ulcorner W(H) = x \urcorner) = x$. This covers all experts W, all propositions they may report on (in their domain) H, and all reported values for their credence $x \in [0, 1]$. So when you hear the report in *Hurricane*, and learn the proposition $\ulcorner W \urcorner$, you update by conditioning and get:

$$Q(H) = P(H|\ulcorner W(H) = 0.3 \urcorner) = 0.3$$

We can immediately note some pros to this approach. The idealisation that you

---

[6]We may of course wish to defer to experts on matters which occur only once, in which case this notion of calibration to frequencies isn't useful.

[7]Lindley (1982, 118) notes the connection between calibration and deference in an early discussion of supra-Bayesianism and deference. When the result of supra-Bayesian updating matches the expert's report, Lindley calls the expert "probability calibrated" for that novice. Due to the differences in how we approach the problem, I won't use Lindley's terminology. As DeGroot (1988, 299) says: "it would be unnecessary to use the term 'well calibrated' in this paper because that property is now simply the defining characteristic of an expert."

simply adopt the expert's reported credence has the benefit that the model is much simpler than supra-Bayesianism. It is Sensitive to Testimony by construction, as the answer depends directly on the content of the report. The problem of the Relevance of Priors is alleviated: it doesn't use your prior for H or $\ulcorner W \urcorner$ to arrive at your posterior for H, but instead fixes your prior for H conditional on $\ulcorner W \urcorner$ (thereby constraining your priors for H and $\ulcorner W \urcorner$, but these don't play a direct role).

The problem of Cognitive Burden is reduced, if we think an equation fixing a whole set of priors is less demanding than having free priors for each W, H and x. This doesn't resolve the part of the Cognitive Burden problem that is directed at the granularity and sensitivity of your prior attitudes, however. At least not if these priors are all interpreted as fully-fledged partial beliefs. A defender might respond that priors *aren't* truly beliefs, they are a representation of the agent's evidential standards (see, e.g., Titelbaum, forthcoming, Ch. 4). (Presumably, this view is primarily about prior conditional probabilities.) They tell us how the agent is disposed to respond to various pieces of evidence they may learn. So when we include in our model a constraint like $P(X|\ulcorner W \urcorner) = x$, we aren't imputing any *attitude* to the agent at the time when P is their credence function. Rather we're saying something about how they will respond when they learn that the proposition $\ulcorner W \urcorner$ is true. A very general principle ("defer to experts") can cover a great many propositions without requiring that the agent take any attitudes before receiving that evidence. There is something to this response, and my own proposal has a similar flavour. But it introduces a somewhat arbitrary distinction between which bits of the prior function count as attitudes (unconditional probabilities) and which do not (conditional probabilities). Moreover, the law of total probabilities allows us to decompose unconditional probabilities into sums of conditional probabilities. What does such a sum mean, if one side of the equation involves attitudes and the other involves "standards"?

More difficult to resolve is the the problem of Awareness. The fact that the agent has attitudes to these propositions is implausible and has no part in the requirements of rationality. In light of the previous paragraph, we might grant that the agent has an evidential standard such that they defer to expert reports even on propositions for which they have no prior. So when they learn $\ulcorner W(H) = x \urcorner$ they adopt x as their credence for H. But this only helps for the reported proposition (i.e., that there will be a hurricane on Thursday). Upon becoming aware of proposition H—and assigning it a worryingly high probability—the agent ought to update their other beliefs, e.g., that their car is at risk if parked outside. In order to arrive at a coherent posterior, the agent must have *already had* conditional probabilities connecting proposition H to all others. These conditional probabilities can't be brushed aside as evidential standards, as the agent didn't learn the propositions that they're meant to encode responses to (i.e., H), they learned $\ulcorner W \urcorner$. So no matter what your stance is on priors, this model of expert deference makes unreasonable Awareness demands.

Expert deference as presented also faces two unique problems.

5. Arbitrariness: This model isolates one kind of expert report as worthy of deference above others. Experts reliably report all kinds of things, and yet expert deference is a principle about deferring to reports of unconditional probability. Yet experts can

(and do) report all manner of probabilistic information[8] including conditional probabilities, Bayes factors, comparisons between the probabilities of various events and variables, expected values for variables, functional forms for distributions over variables, and so on. It is unclear how a simple deference principle, which insists on operating through the Bayesian belief revision process, can capture all of these. But what reason could we have for distinguishing only reports of unconditional probability as worthy of deference?

6. Expert Disagreement: It is unclear how to extend this expert deference principle to the case of more than one expert, when they disagree. Supra-Bayesianism's answer to this is the same as ever: conditionalise! One merely updates on the evidence received, making use of the relevant likelihoods and priors for each expert's report. This has all the problems discussed above for the one-expert case, but it *is* an answer. Expert deference doesn't seem to provide much of an answer at all. One cannot simultaneously defer to two experts (as I have set things up, that operation simply isn't defined), and if one defers in sequence then the last report will dominate.

Deference treats all experts equally; as though they were the *same* expert. And if one received two reports from the same expert on the same topic, and deferred to them each time (assuming perhaps that they'd learned relevant new information), then of course the latter report would dominate. (This is desirable in the one-expert case. Suppose I start off thinking the chance of rain is 20%. An expert says it is 40% and so I defer accordingly. The same expert later says it is 30%. Assuming that I maintain my view of them as an expert, I ought to end up with a credence on 0.3 in rain. The current setup gets us this.) But it also reveals that our model has oversimplified by not distinguishing between the following: (a) a report from one expert, and unanimous reports from many, and (b) a sequence of different reports from one expert, and a profile of disagreeing reports from many.

4 Expert deference as a belief revision schema

My own proposal is a development of the expert deference idea, but one which hopes to avoid the issues discussed above. In order to motivate for it, let us examine the source of our continuing problems with expert testimony. At the heart of the matter is the proposition $\ulcorner W(H) = x \urcorner$, "that the expert W reported x as their probability for H," which is what these Bayesian models take the agent to learn when they hear the report. The problems discussed above arise from:

1. having a proposition in the algebra to represent $\ulcorner W \urcorner$ and (all of the) H's—in the discussion above, this was linked to the problems of Awareness, Cognitive Burden, and Relevance of Priors.

2. having priors for (and related to) each value of x for each W and H—Awareness, Cognitive Burden, Relevance of Priors.

---

[8]Experts report non-probabilistic information too, but here I'll neglect such reports. We can perhaps assume, as many probabilists do, that categorical statements (e.g., "it will rain tomorrow") are expressions of high credence ($P(\text{rain tomorrow}) \approx 1$).

3. updating by conditioning on the fact of the report (rather than using its content directly to change your credence)—Arbitrariness, Sensitivity to Testimony.

4. failing to provide a mechanism for delineating and dealing with multiple expert reports—Expert Disagreement.

In developing a better model, I propose to address each in turn. First, I will remove ⌜W⌝ from the model entirely. Second, I will allow for the fact that agents aren't aware of propositions like H, and so don't have priors for them or reports about them. Third, I will use a different belief revision strategy that is sensitive to the content of the report.

In brief, I propose to regard expert deference as a *belief revision schema*. By this, I mean that I will take expert reports as prompts to revise one's credences in a manner that "fits" the content of the report. To make a start, I will set aside the problem of Awareness and develop the model for familiar propositions. In the following section, I will expand the model to cover novel propositions.

Using a belief revision rule other than Bayesian conditioning is a significant departure from orthodoxy, so let me review the motivation that I have thus far presented. We are modelling a realistic agent's response to testimony from a (relatively) ideal expert. Simple orthodoxy, in the form of supra-Bayesianism, faces two problems of particular concern to real agents: Cognitive Burden and Awareness. The assumption that the expert is ideal yields two further problems, or missed opportunities: (In)Sensitivity to Testimony and (Ir)Relevance of Priors. Expert deference allows us to mitigate against some of the cognitive burden on real agents, by making use of the idealisation that the expert is reliable. But modelling deference as a constraint on priors arbitrarily restricts itself to one report-type. This problem, and the remaining Cognitive Burden concerns, are directly linked to the use of Bayesian conditioning as the belief revision rule. Therefore, I propose to examine an alternative; *not* in order to displace Bayesian conditioning as a norm for ideally rational agents, but as a model for real agents in this particular scenario.

In developing this model, I take inspiration from the literature on Jeffrey conditioning. This is the name given to a rule developed by Richard Jeffrey to deal with circumstances like this one.

> The agent inspects a piece of cloth by candlelight, and gets the impression that it is green, although he concedes that it might be blue or even (but very improbably) violet. (Jeffrey, 1983, 165)

Let G, B and V stand for the propositions that the cloth is green, blue or violet. Suppose that these form a partition for the agent, and that after the inspection the agent comes to have posterior degrees of belief 0.7, 0.25, and 0.05 in them respectively.

What has the agent learned here and how does it affect their beliefs? Jeffrey says: if there were a proposition E in the domain of the agent's credences describing the precise quality of this experience, then we would simply say "the agent learned E." The rational response to this learning experience would then be to update their degrees of belief by conditioning on E—this ought to be how they arrived at the posteriors for G, B and V, and how they update all their other beliefs. But, says Jeffrey, there needn't be any such a proposition in their algebra, nor expressible in English. Anything we could

specify would be too vague to convey the precise quality of this uncertain experience, and too vague to support any meaningful ascription of precise conditional probabilities as the Bayesian procedure requires. It is better to admit that there is *nothing* the agent learns for certain.[9]

Instead Jeffrey proposes that we *describe the effects* of the experience on the agent, by stipulating their credences after the experience. We can omit the proposition E altogether, and merely say that they come to have degrees of belief for the partition $\mathbb{A} = \{G, B, V\}$ equal to $(0.7, 0.25, 0.05) = (\pi_A)$. Jeffrey then provides us with a rule for generating a fully-specified, unique, and coherent posterior credence, now called Jeffrey conditioning.

> **Jeffrey conditioning.** Suppose an agent with initial probability function P comes to have new probabilities for a partition $\mathbb{X}$, denoted π_X for each X in $\mathbb{X}$. The agent's new probability function Q is obtained from P by Jeffrey conditioning if and only if, for all $Y \in \Omega$

$$Q(Y) = \sum_{X \in \mathbb{X}} P(Y|X)\pi_X. \tag{1}$$

A point we will come back to: Jeffrey argues that this is the right rule for revising belief whenever the agent's conditional beliefs given $\mathbb{X}$ remain unchanged; that is, for all $X \in \mathbb{X}$, $Q(\cdot|X) = P(\cdot|X)$. This is called the *Rigidity* condition. Indeed, Eq. 1 is just the law of total probability for Q when Rigidity holds.

Here is the point I want to take away from this. An orthodox Bayesian *can* model this situation. They simply insist that there *is* such a proposition E, capturing the precise quality of this experience. (Perhaps it is a "sense data" proposition, inexpressible in any natural language but corresponding to the precise nature of the agent's experience.) They continue to insist that the proposition *is* a part of the agent's algebra, they have priors for it, and when they learn it they update by Bayesian conditioning. Jeffrey's approach represents a different *modelling strategy* for the same problem. He removes the unrealistic data proposition E from the model, and instead takes the experience to provide an exogenous constraint on the agent's posterior credences. He then proves that, for this kind of constraint (unconditional probabilities over a partition) there is a unique "kinematic" update rule that fixes a posterior credence function Q (Jeffrey, 1983, 164–68).

I propose to do the same: take expert testimony to provide exogenous constraints on credences, rather than modelling it endogenously.

4.1   First pass

I'll start by sketching how this would work for an easy case, where the agent is aware of the proposition the expert reports on. Consider the more prosaic *Weather*.

> *Weather.* You open your (super duper) weather app and see that there is a 60% chance of rain in London this evening.

---

[9]This is linked to Jeffrey's rejection of what he calls "dogmatic empiricism", the view under which there is some basic, sense-data proposition capturing exactly what the agent learns.

Here we can assume that you are aware of the possibility of rain (you live in London, after all) and so we assume there is a proposition $R \in \Omega$ representing it. We needn't assume that you have a precise prior for it: we can say your comparative beliefs were incomplete with respect to R, or that you had completely imprecise credences for it, so $\mathbb{P}(R) = [0,1]$. As we will see, the details won't matter for this first pass.

I propose that we model the expert report as providing the constraint that $Q(R) = 0.6$. There is no proposition in the model representing the expert's report, and we don't model the agent (you) as coming to learn any proposition for sure. Instead, your posterior is bound to obey this constraint.

Now typically our beliefs are multiply connected and other beliefs will depend upon this one. If I change my credence in rain, I will also change my credence in having an enjoyable cycle in to work, and my credence in arriving late, and so on. If I fail to do so here, my credences will be incoherent. So some revision is required for the rest of my degrees of belief. Put in terms of our model, the probabilities of various other propositions $(Y, Z, \ldots)$ that are probabilistically dependent on R ought to change when the probability of R changes.

My proposal is to generate the remainder of the posterior credence by Jeffrey conditioning on the partition $\{R, \neg R\}$.[10] So, concerning being late, for example:

$$Q(\text{late}) = P(\text{late}|R)Q(R) + P(\text{late}|\neg R)Q(\neg R)$$

The language I just used is procedural, reflecting my interest in what real agents might do when confronted with expert testimony. First, one *fixes* the credence R. Then, one *generates* the relevant posterior credences by Jeffrey conditioning. This is explicitly envisaged as a process an agent might follow.[11] This assumes, of course, that they are in possession of numerical probabilities for the relevant propositions. That won't normally be the case for prosaic matters like being late to work. But in policy situations these probabilities are often estimated. A policymaker's ability to follow the procedures discussed in this chapter will therefore depend on which probabilities are available to them.

This is our first look at what I call *expert deference as belief revision*.

4.2 Belief revision theory

I will now introduce some ideas from the theory of belief revision to make this more general. Following the formalism of Dietrich et al. (2016), I will think of a belief revision rule as a function, mapping an initial belief state and an experience to a final belief state. Let $\mathscr{P}$ be a set of possible belief states and $\mathscr{I}$ be a set of possible inputs or experiences, so that a belief revision rule maps $\mathscr{P} \times \mathscr{I} \to \mathscr{P}$. Belief states will be probability functions, or sets of probability functions, as before. "Inputs" are taken to be very general, including straightforward observations, noisy signals, and expert reports of various kinds. We therefore specify them *extensionally*, as the set of belief states that are consistent with the experience. A simple example: if I look out the

---

[10]This suggestion is due to Steele (2012).
[11]In this I follow Jeffrey himself (e.g., Jeffrey and Hendrickson, 1989) and much of the Bayesian statistics literature.

window and see that it is cloudy, this input constrains my belief state to include only those in which it is cloudy outside my window. We have just seen this idea at work in Jeffrey updating.

Belief revision rules can be characterised by two conditions: Responsiveness and Conservatism. Loosely, Responsiveness ensures that the final belief state "respects" the input, and Conservatism ensures that the belief revision changes *only* what is "required" by the input. This is captured by a Conservation condition, that specifies which parts of the prior belief state must be conserved by the revision.

Rules which follow these two conditions are called perturbation-propagation rules.[12] First, as the rule is *Responsive* to the experience, the input will directly bring about a change in belief state: the perturbation. Second, the remainder of the belief state is adjusted to reflect the impact of the input; this makes use of the perturbation and the parts of the initial state that are preserved by the *Conservatism* of the revision. Table 1 shows two common examples: Bayesian updating and Jeffrey updating. Note that the propagation step covers what is typically thought of as a "belief revision rule," such as updating by Bayesian conditioning.

Table 1: Two belief revision rules

Rule	Perturbation	Propagation
Bayes	$Q(E) = 1$	$Q(X) = P(X|E)$
Jeffrey	$Q(A) = \pi_A, \forall A \in \mathbb{A}$	$Q(X) = \sum_{A \in \mathbb{A}} P(X|A)Q(A)$

Our first pass followed this pattern. Perturbation: the report sets a constraint, $Q(R) = 0.6$. Propagation: Jeffrey condition on the partition $\{R, \neg R\}$ to restore coherence. But, clearly, expert deference isn't a kind of belief revision, as Bayes and Jeffrey updating are, for in *Weather* it worked as an instance of Jeffrey updating. This is why I call my proposal expert deference as a belief revision schema.

In generalised belief revision theory, a kind of experience is matched with a particular revision rule. Dietrich et al. (2016) characterise the class of Bayesian inputs as those experiences which constrain the agent's belief state to include only probability functions in which the probability of a specific proposition—the one the agent learns during the experience—is 1. We can similarly define the class of Jeffrey inputs (corresponding to Jeffrey updates) or Adams inputs (corresponding to Bradley's Adams updates (Bradley, 2017)). In each case, we can model this with a domain $\mathscr{D}$ that is a subset of the space of possible experiences and initial states: $\mathscr{D} \subseteq \mathscr{P} \times \mathscr{I}$.

Responsiveness consists in ensuring that the final belief state is in the set *I*, i.e., that it is compatible with the experience. Conservatism is harder to spell out. Each domain that Dietrich et al. consider comes with a specification of what those experiences are "silent" on. This notion of silence is used to fill out the norm of Conservatism: put roughly, a belief revision rule should leave unchanged whatever the experience is silent on. Dietrich et al. (2016) then prove characterisation results showing that, for each of these three domains, there is a unique rule respecting Responsiveness and Conservatism and that in each case it is the rule referred to parenthetically above.

---

[12]I take this term from Bradley (2017).

But recall the problem of Arbitrariness: experts may report many kinds of probabilistic information. They might tell you that one event is more likely than another, or that two variables are independent, or they might specify the expected value for some variable. (van Fraassen (1981) says that reporting the expected value for a variable is the most general kind of constraint on your probability function, and that others can all be framed as special instances of it.) We want our theory of expert deference to be able to handle all of these report types.

The problem is that for these more general input domains, no unique belief revision rule is known. Put extensionally, the problem is that once we identify the set of belief states that respect the input, we lack general rules for further refining this set. This may be easier to see by switching to an intensional definition. Let us denote the constraint imposed by the expert report with a formula, ϕ_Q. Responsiveness tells us that the posterior credence function must respect this constraint: we want a Q that respects ϕ_Q. But there are a great many of these! What we want is one which also fits the prior, P, in the right way. What is that way? Conservatism is meant to provide the answer: in the way that preserves as much of P as possible while respecting Q.

In practice, specifying the Conservatism norm for a form of experience is a complex matter. The canonical examples mentioned above have a particularly nice form: each comes with a constraint and a Rigidity condition—a particular realisation of Conservatism. These conditions, summarised in Table 2, are necessary and sufficient conditions for updating according to the associated belief revision rule (see Bradley, 2017, 188-200). ($\mathbb{A}, \mathbb{B}$ are partitions. Adams updating will be introduced below.)

Table 2: Conservatism conditions for Bayes, Jeffrey and Adams updates

	Constraint	Rigidity condition(s)
Bayes	$Q(E) = 1$	$Q(\cdot \mid X) = P(\cdot \mid X), \forall X \in \Omega$
Jeffrey	$Q(A) = \pi_A, \forall A \in \mathbb{A}$	$Q(\cdot \mid A) = P(\cdot \mid A), \forall A \in \mathbb{A}$
Adams	$Q(A \mid B) = \pi_A^B,$	$Q(\cdot \mid AB) = P(\cdot \mid AB), \forall A, B \in \mathbb{A}, \mathbb{B}$
	$\forall A, B \in \mathbb{A}, \mathbb{B}$	$Q(A) = P(A), \forall A \in \mathbb{A}$

In the general cases discussed above, we don't have Conservation conditions that produce unique "kinematic" update rules of this kind. We either need to do more work to identify what is conserved by the experience (or what it is "silent" on, in the language of Dietrich et al. (2016)), in order to formulate a kinematic revision rule, or, lacking that, we need an alternative way of getting from P to the right kind of Q.

4.3 Deference in action

The model outlined above is simple but surprisingly flexible. The problems for supra-Bayesianism were associated with the propositions representing the expert reports, and the content of those reports. I therefore removed the expert reports from the algebra entirely—instead of representing them as propositions they are now externally given constraints on the agent's posterior beliefs. The Bayesian updating procedure has been replaced with expert deference (now realised as the imposition of this external con-

straint) and a belief revision schema, in which Bayesian conditioning is one element. This reduces the Cognitive Burden on agents, as they are not required to have a myriad of prior beliefs. It resolves the Relevance of Priors problem, as the posterior does not depend on the agent's uninformed priors for the propositions in the expert domain (those play no role and aren't required to exist). In deferring, it is properly Sensitive to Testimony.

I'll briefly demonstrate how the model allows agents to defer to a much wider class of expert testimony, thereby resolving the problem of Arbitrariness. First, consider reports of the truth of a proposition. For example, a completely reliable person looks outside and tells you it is raining. You defer to them, and set your credence in this proposition to 1. This is an instance of expert deference that provides a Bayes-type constraint on your credences and so the schema tells you to update your beliefs using the appropriate rule: Bayesian conditioning. You haven't undergone a Bayesian learning event and so this isn't just Bayesian learning—after all, you haven't observed the rain, and the proposition you learned is $\ulcorner R \urcorner$ rather than R itself.

Second, reports of unconditional probabilities. (This is the first-pass case, included for completeness.) If an expert reports probabilities π_X for each X in a partition $\mathbb{X}$, the deferring agent sets their probabilities to $Q(X) = \pi_X$. This constraint is of the Jeffrey type, and so the appropriate propagation is Jeffrey conditioning on the partition $\mathbb{X}$— assuming the appropriate Rigidity condition holds. Does it? It will depend on the case, but it is certainly plausible in many cases. When an expert reports just unconditional probabilities, the agent is given new probabilities over a partition, just as in a Jeffrey experience. Further, what the agent learns from the expert is *entirely* given by the set of new probabilities $\{\pi_X\}_{X \in \mathbb{X}}$, particularly in our ideal expert case where no information about their reliability is indirectly received. Our expert was (literally) silent on other matters, including probabilities conditional on the X's. So Rigidity holds, and Jeffrey updating is the correct propagation.

Third, reports of conditional probabilities. Suppose that an expert reports their conditional probability for X, drawn from a partition $\mathbb{X}$, given some other possibility $Y \in \mathbb{Y}$: π_X^Y. Then the deferring agent should set their conditional probability to $Q(X|Y) = \pi_X^Y$. This constraint is of the Adams type, a name introduced by Richard Bradley (2005). Bradley developed an updating rule for such inputs, assuming that the Rigidity conditions shown in Table 2 hold. Dietrich et al. (2016) prove that this rule, Adams conditioning, is the unique Responsive and Conservative belief revision rule for such cases. Here is the rule: in the situation just described, the new probability function Q is obtained from P by Adams conditioning if and only if, for all $Z \in \Omega$:

$$Q(Z) = \sum_k \left[\sum_j P(Z|X_jY_k)\pi_j^k \right] P(Y_k) \tag{2}$$

Adams conditioning is the right way to update when two conditions hold. The first, Bradley calls *Independence*: $Q(Y_k) = P(Y_k)$ for each k. The second is another Rigidity condition, this time for the probabilities conditional on the cells of the joint partition $\mathbb{X}\mathbb{Y}$: $Q(\cdot|X_jY_k) = P(\cdot|X_jY_k)$. These jointly constitute the Conservation condition for this belief revision rule. The repetition of the familiar Rigidity condition provides some of the motivation for this revision procedure: Adams conditioning can be thought of as

a special case of Jeffrey conditioning. As Bradley puts it, "if Jeffrey conditioning is the correct revision rule for Jeffrey experiences then Adams conditioning is the correct rule for Adams experiences" (Bradley, 2017, 199).

In what kinds of expert deference cases does this rule apply? Consider *Cancer*.

> *Cancer.* Eva has a family history of ovarian cancer. She has been told that, based on this, she has a 3-5% lifetime chance of developing the cancer herself. Recently, she learned that there is a genetic mutation called BRCA2, present in 1/1000 people, which makes ovarian cancer more likely. She meets with a genetic counsellor, who tests her for the mutated gene. Before she receives her results, the doctor tells her: "patients with a mutated BRCA2 gene have a much higher risk of developing ovarian cancer, around 23%." If she tests negative, her risk will be in the range she was previously told, 3-5%.

Let O be the proposition that Eva will develop ovarian cancer at some stage in her life, and B stand for having a mutation on that gene. Assume that coming in, she has $P(O) = 0.04$ and $P(B) = 0.001$. Given the doctor's testimony, it seems reasonable for Eva to set $Q(O|B) = 0.23$ and $Q(O|\neg B) = 0.04$. Her unconditional probability $P(O)$ should change, though it isn't immediately clear how. What does seem clear is that she *shouldn't* adjust the probability that she has the mutated gene $P(B)$—she's received no information relevant to this (yet; she will receive her test results in time).

What Eva needs is a way of updating her beliefs that respects these constraints. This will involve adjusting her joint probabilities across the possibilities: developing cancer, not developing cancer, having the mutated gene, not having it. Realistically, Eva didn't have any views at this fine grain before (it seems implausible that she had a conditional probability of developing cancer given a positive test result), so it is more accurate to say that what she needs to do is to distribute her beliefs over the possibilities of developing ovarian cancer and having the mutated gene in a way that fits the expert testimony she has received. She should defer to the doctor and then update by Adams conditioning.

Applying equation 2 to Eva's case, this means that her probability for a general proposition Z should now be:

$$Q(Z) = [P(Z|OB)Q(O|B) + P(Z|\neg OB)Q(\neg O|B)]P(B)$$
$$+ [P(Z|O\neg B)Q(O|\neg B) + P(Z|\neg O\neg B)Q(\neg O|\neg B)]P(\neg B)$$

If we consider $Z = O$ and ask what her probability in developing ovarian cancer should be, we get:

$$Q(O) = Q(O|B)P(B) + Q(O|\neg B)P(\neg B)$$

This is just the law of total probability, but with $Q(B) = P(B)$ due to the Independence condition. This is the natural way to set $Q(O)$ given what Eva has available to her at this stage in the example. Given the tiny probability of having the mutated gene $P(B)$, we can see that Eva should have $Q(O) \approx Q(O|\neg B) = 0.04$.

What of the more general report-types that I mentioned above: learning that an expert says event A is more probable than B, or that they are independent? The model

can handle them, in a fashion. Whether there are unique kinematic update rules for them, and what the associated Conservation conditions are, are open questions. The resulting procedure is therefore somewhat trivial, and the result very imprecise. The report generates a constraint, ϕ_Q, which says that $Q(A) > Q(B)$, for example. The agent must adopt a posterior which fits this, but in the absence of a conservation condition there are a great many of these. The result is therefore an imprecise posterior which includes all functions that match the prior for any propositions unrelated to A and B, and match the constraint for A and B.

This ends my brief tour, but I hope the message is clear. In this framework you can defer to any kind of expert report. Where there is a unique kinematic belief revision rule associated with the kind of report (as above), the model will recommend its use in arriving at a unique posterior. Where there is no such kinematic update rule, I leave the model open. Just as there is an unresolved problem in general belief revision theory about what to do in such cases, that problem is inherited here.

Those who advocate for updating based on divergence-minimisation can plug those methods in to this model. This involves finding the coherent function which obeys ϕ_Q and is closest to P by some measure of "distance". Such "distance-based" epistemology is increasingly popular. There is a large literature on using divergences between probabilities for epistemological or decision theoretic purposes: such methods are present in foundational work by de Finetti and Savage; van Fraassen (1981) and Diaconis and Zabell (1982) utilise them; and they are present throughout the "accuracy" programme in epistemology including in work by Joyce (1998) and Pettigrew (2016). Recently, Eva et al. (2019) demonstrated how these methods can be used to update belief after learning (a certain kind of) conditional, in a manner that is apt to deployment in my schema. This may well help to close the remaining imprecision for very general kinds of reports: one chooses the closest posterior which matches the constraint. This amounts to integrating my model with the strategy van Fraassen (1981) outlined.

A final note on imprecise probabilities: I have considered different kinds of probabilistic report, in the sense of the expert reporting unconditional probabilities, conditional probabilities, etc. But I assumed that these reports were *precise* in all cases: the expert reported single numbers. However, the model could easily be extended to cover imprecise probabilistic reports, of any kind. The belief revision framework that I use involves determining the "input" that the learning experience provides. There is no reason to suppose that this input must be precise. The expert report could just as easily require that the agent's posterior probability for the relevant proposition lie within a range. More would need to be said about how to conduct the propagation step for such an imprecise input, but there do not seem to be new difficulties here beyond those already found in the imprecise probabilities literature. So, if one is willing to accept imprecise probabilities at all, my approach can be adapted to handle them.

5   Awareness growth

At the outset, I pointed out that one of the distinctive features of expertise is that experts often talk about things that laypeople have never heard of before. In this section I will show how expert deference as a belief revision schema can go together with one model

for rational awareness growth, thereby allowing agents to defer to experts who report on such unfamiliar propositions.

Let us return to our initial case, *Hurricane*: You open your weather app and see, to your complete surprise, that there is a 30% chance that London will be struck by a hurricane on Thursday.

In reviewing this example above, I said that one implausibility of supra-Bayesianism is that the agent had never considered this possibility before. They simply have no attitude toward it. This will be my definition of unawareness: a state in which an agent has no attitude to a proposition. Therefore, in our model, it is natural to represent this with a proposition, e.g., H, that is not in the domain of P. As P's domain is the whole algebra Ω, this means that $H \notin \Omega$

We know where we would like to end up: the agent comes to have an attitude to H, in particular $Q(H) = 0.3$. The problem of awareness growth is to find rational constraints on Q, linking it with the prior belief state P.

Here is how my proposal goes. We can think of the expert's report as having two parts: the proposition it concerns (its *content*), in this case H, and the probabilistic information it conveys about that content (its *value*), in this case 30%. I will separate my treatment of awareness growth into two stages, corresponding to these two parts. (The separation is conceptual, and should not be taken to imply that the agent follows this sequence.) The first stage is purely a matter of awareness: the agent becomes aware of the new proposition(s) and we determine how their old attitudes can be extended to an algebra that contains the new proposition(s). The second stage is a matter of learning: the agent became aware of the proposition(s) via a learning experience, and in particular an experience of expert testimony. This gives them evidence about the new propositions. In the second stage, I will apply my expert deference proposal and show how it fits naturally with one theory of rational awareness growth.

I will start with some discussion of awareness growth in general, to set the stage. Bradley (2017, 256–8) differentiates between two kinds of awareness growth, distinguished by the kind of belief change they require of the agent.[13]

First, an agent may come to realise that the possibilities they considered were too coarse. *Refinement* involves making new distinctions, dividing up the possibilities into finer units. Suppose I am considering the weather, and I initially entertain just two options: rain (R) or shine (S). I later realise that temperature is important too. I distinguish two temperatures, warm (W) and cold (C), and recognise that it can be rainy and warm, rainy and cold, sunny and warm, and sunny and cold. (At this point in the story, I know nothing about the relations between temperature and precipitation.) Note that because this is a refinement, $\{R, S\}$ is still a partition for me ($R \vee S = \top$). What has changed is that R and S are no longer "primitive possibilities," in the sense that they aren't maximally specific: I am now aware of two ways that it might rain, RW and RC. The primitive possibilities are now the four elements of the joint partition $\{RW, RC, SW, SC\}$.

In the second kind of awareness growth, *expansion,* the agent becomes aware of a new primitive possibility. Suppose again that I am considering the weather, and I

---

[13]Bradley takes these two options to be exhaustive; other forms of awareness growth are reducible to combinations of them, perhaps in combination with two corresponding forms of awareness contraction. I do not need this to be true in what follows.

initially entertain just two options: rain (R) or shine (S). Then, upon looking at my app, I consider a third option which I see as mutually exclusive of these two: hurricane (H). Before, I regarded $\{R,S\}$ as a partition: $S = \neg R$ and $R \vee S = \top$. After my awareness grows, $R \vee S$ is merely a contingent proposition, $\neg R = S \vee H$, and $\{R,S,H\}$ is a partition.

Here is how I propose to model these learning experiences. First, we deal with the purely awareness related part of the experience. We will "grow" the algebra Ω to include the content. Then we will extend the agent's prior credence function P to the new, wider algebra $\Omega^{\oplus}$. The thought here is that some aspects of P will be preserved in a process of rational awareness growth, and so we can specify some conditions that any credence function on $\Omega^{\oplus}$ must have in order to "fit" with P. Second, we model the learning experience. Here I bring in my belief revision schema, beginning with a perturbation using the report: the probability for H is set by the report's constraint ϕ_Q. Finally, I "fill in" this Q to "match" the extended prior, in a move analogous to the propagation part of a belief revision process. (All of the terms in scare quotes are loose descriptions that will be made more precise as we go.) For clarity I will separate each stage—awareness change and learning—into two steps.

Step 1: Growing the algebra.

Having outlined what awareness growth is, we can consider how to revise belief in the face of expert testimony on novel propositions. The first step makes use of only the content of the testimony, setting aside its value. We start by forming a new algebra, containing all the propositions the agent was previously aware of and the propositions reported on.

Making this more precise requires slightly more mathematical machinery than we have thus far employed (following Bradley, 2017, 258–9). I have been using a logical framework in which Boolean algebras are lattices of propositions, ordered by an implication relation. In order to make the lattice structure explicit, I will now write $\Omega = \langle \mathscr{X}, \models \rangle$, where $\mathscr{X}$ is a set of propositions and $\models$ is the implication relation that acts as the order for the lattice. The top element of the algebra is typically denoted $\top$, but in the context of multiple algebras it is useful to think of it as merely being the upper bound of the set $\mathscr{X}$: $\bigvee \mathscr{X}$.

In general we can suppose that the agent becomes aware of a set of propositions $\mathscr{U}$, with $U \notin \mathscr{X}$, for all $U \in \mathscr{U}$. We start by forming $\mathscr{Y}$, the closure of $\mathscr{U} \cup \mathscr{X}$ under the Boolean operations. Then $\Omega^{\oplus}_{\mathscr{U}} = \langle \mathscr{Y}, \models \rangle$ is a Boolean algebra, which Bradley calls the extension of Ω by $\mathscr{U}$. Note that $\bigvee \mathscr{X} \in \mathscr{Y}$, and in general $\bigvee \mathscr{X} \neq \bigvee \mathscr{Y}$.[14]
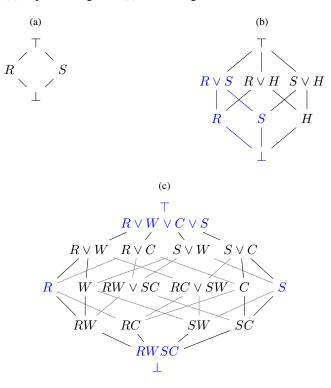
The old algebra is related to the new one via an embedding. A lattice embedding is a one-to-one homomorphism: a function that maps each proposition in the old algebra to a proposition in the new algebra, and which preserves the lattice operations, meet

---

[14]Note a persistent idealisation here: $\models$ is the implication relation which ordered the old algebra, and it also orders the new propositions. So, the agents that we model in this framework are logically omniscient (as is standard) and this omniscience extends to propositions they were previously unaware of. The problem of logical omniscience is a significant one for someone with my non-ideal theory interests. However, treating it is notoriously difficult. I therefore put up with this idealisation, noting that allowing agents to be unaware does mitigate force of the problem of logical omniscience.

and join—which is to say, logical conjunction and disjunction. It does *not* preserve logical complements.

We can see this by considering the weather example again. Simple lattices can be usefully visualised with Hasse diagrams, such as those in figure 1. The lines connect the higher-up propositions with logically stronger propositions below, which entail them. The figure highlights in blue the elements of the new algebra which correspond to the old propositions. Note that the element $R \vee S$ is in $\Omega^{\oplus}$ in both the expansion and refinement case. In the expansion case it is now merely a contingent proposition, as the tautology in $\Omega^{\oplus}$ is $R \vee S \vee H$. In figure 1(b) it is easy to see that the old algebra is a sub-algebra of the new. In the refinement case, $R \vee S$ is also (or, if you prefer, still) the top element of $\Omega^{\oplus}$. Figure 1(c) shows how much more complicated things look with more propositions.

Figure 1: Hasse diagrams showing Bradley's two kinds of awareness growth. (a) Original algebra, (b) expanded algebra, (c) refined algebra.

Step 2: Extending the probability function.

What we have now is an algebra, $\Omega^{\oplus}$, which contains the new propositions. But P is not defined on $\Omega^{\oplus}$, but rather on Ω. Our second task is to extend P to $\Omega^{\oplus}$. "Extension"

is not belief revision. I am interested here merely in what the prior P has to say about $\Omega^{\oplus}$.

There are different proposals for how to extend a probability function to a wider algebra. Two prominent proposals are "Reverse Bayesianism" (due to Karni and Vierø, 2013) and "Rigid Extension" (due to Bradley, 2017). I will employ the latter method.

Bradley (2017, 257) begins by considering simple examples of refinement and expansion, like my weather case. He characterises refinement as coming to realise that the possibilities I previously considered are too coarse-grained. But introducing finer-grained possibilities should not change my attitude to the coarse-grained possibilities; realising that there are two kinds of rainy weather should not change the probability of rain overall.

For expansion this is not the case. Introducing an entirely new kind of weather must change my attitude to at least one of the possibilities I previously considered, since my degrees of belief must sum to one. But, Bradley argues, there should not be any relative change between the old propositions: if I previously thought rain and shine equally likely, I have no reason to alter that relative comparison now that I have discovered that those alternatives do not exhaust the possibilities.

The core idea here is simple, and in line with our discussion above: minimal change. Bradley argues that the Conservation condition for three Bayesian belief revision rules (Bayes, Jeffrey and Adams updating) involves the rigidity of conditional beliefs. So, he concludes, a Conservation condition for extension to a wider algebra should also preserve conditional probabilities. He provides such a condition which captures the intuitions that he defends about refinement and expansion cases.

Specifically, "the agent's new conditional probabilities, given the old domain, for any members of the old domain should equal her old unconditional probabilities for these members." (Bradley, 2017, 258) Or, to use terminology Bradley introduces, the new belief states must be rigid extensions of the old.

> **Rigid Extension.** A probability function $P^{\oplus}$ on $\Omega^{\oplus}$ is called a *rigid extension* of P iff, for all $X \in \Omega$, $P^{\oplus}(X|\top_{\Omega}) = P(X)$.

In general, there will be many rigid extensions of a credence P to a wider algebra $\Omega^{\oplus}$. Rigid Extension concerns only the parts of the new function that involve propositions from the old algebra, so it leaves open many possible assignments of probability to the new propositions, and combinations of new and old. (Rigid Extension does constrain the latter.) In figure 1, these elements of the old algebra are shown in blue, when embedded in the new algebras on the right.

This is a minimal condition, and will therefore typically result in imprecise posterior credences even if one starts with a precise credence P on Ω. I will denote the result of Bradley's procedure $\mathbb{P}^{\oplus}$; it is the set containing all the rigid extensions of P to $\Omega^{\oplus}$.

Step 3: Perturbation.

$\mathbb{P}^{\oplus}$ is an intermediate construct. It represents what the agent's old probabilities have to say about the new space of possibilities the agent is aware of. We can now model the learning experience that the agent undergoes in virtue of *hearing an expert report* about those new possibilities.

As before, the report provides a constraint ϕ_Q on any credence function that is Responsive to the learning experience. Note that in this case a single experience is producing both the growth of the algebra and the perturbation of the credence function. The decomposition into two steps is merely a logical decomposition, helpful in modelling awareness growth—there is no implication that the two are separated in time, or occur separately.[15]

The constraint in example 1 is that $Q(H) = 0.3$, and so we will work with the set of potential functions $\mathbb{Q} = \{Q \text{ is a credence on } \Omega^\oplus : Q(H) = 0.3\}$.

Step 4: Propagation.

We now want to further constrain $\mathbb{Q}$, so that it fits with our extended prior $\mathbb{P}^\oplus$. In our simple belief revision cases above, we accomplished this by using a Conservation condition that told us which parts of the prior should be conserved. We used these conserved quantities, together with ϕ_Q, to fix the posterior (uniquely, in cases of kinematic belief revision).

As discussed above, the Conservation condition depends on the nature of the learning experience. In our example, the report provides us with an unconditional probability for H. When we encountered reports of this kind before, we were able to motivate for the rigidity of conditional probabilities as the Conservation condition, and therefore to use Jeffrey conditioning as our propagation procedure. We *cannot* do that here. The agent was unaware of H before the learning experience, and so does not have prior conditional probabilities concerning H.[16] As they don't exist, they cannot be the subject of any Conservation condition such as rigidity. A similar argument blocks the use of any kinematic rule for any awareness-growing instance of expert testimony. The challenges we face here are therefore akin to those discussed in section 4.2 for general reports.

There will be some cases where we can make progress. If the algebra is simple enough, and the constraint ϕ_Q restrictive enough, we might nevertheless determine a unique posterior Q. I present two simple cases, to illustrate the difficulties and how they can sometimes be avoided.

> *Expansion:* Considering the weather, you initially entertain the possibilities $\{R, S\}$, assigning each credence 0.5. You later become aware of H, which you take to be a distinct possibility, when you hear the weather report, $W(H) = 0.3$.

Let's work through the four steps.

---

[15]Bradley's approach fits naturally with my topic of expert deference. I want the posterior attitudes to the new propositions to come from the expert reports, and therefore it is helpful to maintain the separation between these steps. Karni and Vierø (2013) do not provide anything like this clean separation. In other cases, this methodological separation may be not be desirable: in a forthcoming paper Steele and Stefánsson (forthcoming) argue that this two-step procedure is baseless. In my case I think the value of the separation is clear.

[16]Put another way: $\mathbb{P}^\oplus$ has no constraints on conditional probabilties involving H that aren't just consequences of Rigid Extension.

1. *Grow algebra:* Your initial algebra Ω is generated by the two-element partition $\mathscr{A} = \{R,S\}$. The expanded algebra $\Omega^\oplus$ is generated by the three-element partition $\{R,S,H\}$. It has the structure shown in 1(b).

2. *Extend prior:* Your prior P is fully specified by $P(R) = 0.5$. A function $P^\oplus$ on $\Omega^\oplus$ is a rigid extension of P iff $P^\oplus(X|\vee\mathscr{A}) = P^\oplus(X|R\vee S) = P(X)$. So, your extended prior consists of all such functions, for which $P^\oplus(R|R\vee S) = 0.5 = P^\oplus(S|R\vee S)$.

3. *Perturb:* Here we simply form the set of all Q's on $\Omega^\oplus$ such that $Q(H) = 0.3$.

4. *Propagate:* To make progress we consider the structure of the algebra. This testimony has introduced a third primitive proposition, making $\{R,S,H\}$ a partition. The constraint $Q(H) = 0.3$ determines the probability of $Q(R\vee S) = 1 - Q(H) = 0.7$. By rigid extension we know $Q(R|R\vee S) = 0.5$, and this is sufficient to fix a unique Q, specified by the following probabilities for the atoms: $Q(R) = 0.35 = Q(S), Q(H) = 0.3$.

We are able to make progress here because between the report and the demands of rigid extension, we fix the probabilities of the atoms of the new algebra. This will not always be the case.

> *Refinement:* Considering the weather, you initially entertain the possibilities $\{R,S\}$, assigning each credence 0.5. You later become aware of temperature, $\mathscr{E} = \{W,C\}$, which you know can combine with R, S, when you hear the weather report $W(W) = 0.6$.

Things are harder in the refinement case. We can follow the steps just described, but there is an important difference. Refinement changes the atoms of the algebra: that's part of what it means to fine-grain in this setup. The atoms of the new algebra are the elements of the finest joint partition over all currently known partitions. (These are the maximally specific possibilities.)

If ϕ_Q is specified at the level of the new algebra's atomic propositions —in the example, $\{WR,WS,CR,CS\}$—and determines the probabilities for all atoms, then Q will be uniquely specified. If not, it will be under-specified. There are many ways that a probability function can satisfy $Q(W) = 0.6$: temperature might be independent of precipitation, or warmth may be more likely if there is no rain, or warmth may be more likely if there *is* rain.

If the constraint is at the level of the new coarse-grained partition, rather than the joint partition, the result will be an imprecise posterior $\mathbb{Q}$, containing all the joint distributions meeting the constraint.

We can now see how my proposal for expert deference as a belief revision schema fits naturally with this approach to awareness growth. The fact that the agent is unaware of the content of the expert report means that more orthodox Bayesian methods could not make progress here. It is implausible that the agent has a prior for the reported proposition, or any attitudes about the proposition that the expert make the report they do about the content. In my proposal, this is no problem. Expert reports are not within the model, but rather act as external constraints on posterior credences. It does not

matter that the propositions these reports concern are not in the old algebra, as they only play a role after we have formed a new algebra and extended the agents priors to that new algebra.

My discussion highlights the difficulties of rational awareness growth. When one becomes aware of a new proposition, one needs to consider not only one's attitude to that proposition but also how it relates to all familiar propositions. If the learning experience that goes along with the growth of awareness does not speak to these logical connections, then the agent will not have enough information to form a unique posterior credence. This is not a feature of my proposal for handling expert testimony, however, it is a feature of the problem of awareness growth.

6   Disagreeing experts

We now return to the problem of Expert Disagreement. The belief revision schema I outlined above has not yet said anything about how to deal with the general *n*-expert case. Nor do I have a complete solution to offer, for this is not a question that deference models are suited to.

Let us examine what is difficult about disagreement, by looking at one of the above cases of deference. Consider a report of unconditional probabilities and suppose now that we have two experts each reporting on the same partition $\mathbb{X}$, so that the agent receives π_X^1 and π_X^2. My proposal says to take these as exogenous constraints, and update by Jeffrey conditioning. Now, these successive Jeffrey updates will not commute, which can be easily seen by noting that the agent's probabilities for $\mathbb{X}$ will match whichever report is deferred to last. Indeed, this is a reason that some philosophers dislike Jeffrey conditioning: because of its supposed non-commutativity—i.e., when considering two Jeffrey updates, we get different final probabilities depending on the order of the updates.

This is not actually correct for Jeffrey *experiences*. The non-commutativity arises from simple applications of Jeffrey's rule, without paying proper attention to the experience generating the changes. Let us say an agent "updates on a partition $\mathbb{X}$" to mean that they propagate probabilities for that partition across their probability function using Jeffrey's rule, denoted $P_{\mathbb{X}}$. If we describe an agent updating first on a partition $\mathbb{X}$, then on another $\mathbb{Y}$, the result will in general differ from that derived from first updating on $\mathbb{Y}$ and then $\mathbb{X}$: $P_{\mathbb{X}\mathbb{Y}} \neq P_{\mathbb{Y}\mathbb{X}}$. This is a fact of the mathematics of Jeffrey conditioning. Nevertheless it isn't right to say the result of *sequential Jeffrey experiences* doesn't commute, because this is not the proper representation of sequential Jeffrey experiences. A careful analysis of sequential Jeffrey experiences, such as that provided by Wagner (2002), allows us to model them in a *commutative* manner.[17]

---

[17]Wagner presents his results as an extension of those which stipulate sufficient conditions for Jeffrey conditioning commuting—see the next footnote. For Wagner, these conditions are two identities for the Bayes factors generated by each of two experiences, relative to two different partitions $\mathbb{X}$ and $\mathbb{Y}$. Let $\beta_{1\mathbb{X}}$ represent having the 1st Jeffrey experience occurring relative to partition $\mathbb{X}$. Then Wagner's identities are $\beta_{1\mathbb{X}}(X_i, X_j) = \beta_{2\mathbb{X}}(X_i, X_j), \forall i, j$ and the same for $\mathbb{Y}$. But as Wagner notes, Bayes factors are the right way of representing what is learned in an experience in a prior-free way. So if we stipulate that the experiences are identical, then Wagner's Bayes factor identities hold, and therefore Jeffrey conditioning commutes across the order of the experiences.

However, while successive Jeffrey experiences, properly modelled, will have commutative effects, sequential deference to expert reports of the Jeffrey-type may be *non-commutative*.[18] This is because my procedure *cannot* be understood in terms of sequential Jeffrey experiences. There is no ineffable learning experience going on here; all we are doing is making use of the machinery of Jeffrey conditioning in this particular instance of the belief revision schema of expert deference.

This should not surprise us in the context of expert disagreement. The general lack of commutativity is precisely the problem of expert disagreement, realised in our model. The model treats all experts similarly: as providing a constraint which posteriors must conform to. The idealisation which licences deference—that the expert is completely reliable—also leads to all expert reports being treated identically.

This is a real limitation. But let me make a comment on its scope. In cases where the experts don't disagree very much, we can make some progress. Since the experts are treated identically, we can treat a set of reports simultaneously, as providing a single constraint ϕ_Q that needs to be deferred to (rather than regarding the reports as distinct experiences to be deferred to in some sequence, as I did above). The easy case is one in which the expert reports, while different, are compatible. Consider one expert who says that a probability is above 0.5, while another says that it is below 0.6. They didn't make the same report, and they disagree about the plausibility of certain ranges of probability for this event. But there is a set of probability functions that is compatible with both: those that say the relevant probability is in the interval $(0.5, 0.6)$.

My approach has a natural way of dealing with "easy" cases. The agent can defer to this joint input, taking it as a perturbation of their belief state. Following the recipe above, we can now look for propagation procedures. But we have a choice at this stage. Our first option is to stick with a precise probability model, in which the agent must have a single probability for the proposition, lying within the range $(0.5, 0.6)$. The second option is to allow for imprecise probabilities, and have the agent adopt the range $(0.5, 0.6)$ as their belief state.

If we stick with a precise model, we will need some principle for choosing from the allowed range. There are many possibilities. Perhaps if the agent's prior is within the interval, then they can retain it. For the moment, however, let's suppose that all options are permissible. The agent then picks one and uses an appropriate propagation procedure to bring their credences to coherence. If the content of the report is a familiar proposition, then the agent proceeds as described in section 4. If it is an unfamiliar proposition, they proceed as described in section 5.

If we go with an imprecise model, then we are in the same situation discussed briefly at the end of section 4, where I considered imprecise reports. There is room for disagreement here about how to conduct the propagation step. One natural approach

---

[18]Why only "may be" non-commutative? In the finite setting, Diaconis and Zabell (1982) provide necessary and sufficient conditions for commutativity. Consider two partitions and the sequences of probabilities assigned to them in a Jeffrey update: $\{\mathbb{X}, \langle x_j \rangle\}$ and $\{\mathbb{Y}, \langle y_k \rangle\}$. $\mathbb{X}$ and $\mathbb{Y}$ are *Jeffrey independent* with respect to P, $\langle x_j \rangle$ and $\langle y_k \rangle$, if $P_{\mathbb{X}}(Y_k) = P(Y_k)$ and $P_{\mathbb{Y}}(X_j) = P(X_j)$ holds for all j, k. Then successive Jeffrey updates commute, $P_{\mathbb{XY}} = P_{\mathbb{YX}}$, if and only if $\mathbb{X}$ and $\mathbb{Y}$ are Jeffrey independent with respect to P, $\langle x_j \rangle$ and $\langle y_k \rangle$ (Diaconis and Zabell, 1982, Theorem 3.2). This turns out to be a weaker condition than probabilistic independence, so that if $\mathbb{X}, \mathbb{Y}$ are probabilistically independent with respect to P, successive updates on them will commute for any update probabilities (Diaconis and Zabell, 1982, Theorem 3.3). So, while some sequences of Jeffrey updates will commute, in general we should expect them not to.

is to generate an imprecise posterior state by performing the propagation using every permissible function in the input. Suppose that we are in an unconditional probability case, and the experts have reported on the probability of rain. The perturbation sets your imprecise probability for rain to $(0.5, 0.6)$. You now propagate this through your belief state by Jeffrey conditioning. But instead of doing it once, you do it for every value in the interval $(0.5, 0.6)$. In this way you generate an imprecise posterior credence.

The hard cases involve contradictory reports. If the first expert said the probability was *below* 0.5, and the second that it was *above* 0.6, then there would be no probability functions compatible with both reports. Here, expert deference cannot help us without further supplementation. The way the ideal expert idealisation is realised in this model doesn't make room for disagreement.[19] Perhaps expert disagreement is a sign of some failure, that needs to be dealt with by more realistic tools. Many exist! One might use a form of opinion pooling to arrive at a single report, given a profile of different reports. Alternatively, the experts might engage in a consensus-building process and produce a single report themselves.

7 Conclusion

My proposal was constructed to deal with the problems identified for supra-Bayesianism and expert deference as a constraint on priors. The problems with supra-Bayesianism were: Cognitive Burden, Relevance of Priors, Sensitivity to Testimony, and Awareness. The orthodox model of deference did better on the first three, but introduced two additional problems: Arbitrariness, and Expert Disagreement.

The problems for supra-Bayesianism were associated with the propositions representing the experts' reports and the content of those reports. I therefore removed the expert reports from the algebra entirely; instead of representing them as propositions, they are now externally given constraints on the agent's posterior beliefs. The Bayesian updating procedure has been replaced with expert deference (now realised as the imposition of this external constraint) and a belief revision schema, in which Bayesian conditioning is one element. This reduces the cognitive burden on agents, as they are not required to have a myriad of prior beliefs. It does not depend on the agent's uninformed priors for the propositions in the expert domain, as those play no role (and aren't required to exist). In deferring, it is properly sensitive to the content of the expert's testimony.

I depart from strict Bayesian orthodoxy by considering a wider range of updating rules. This allows me to avoid arbitrarily treating only expert reports of unconditional probability as worthy of deference. My proposal does, therefore, inherit concerns about the forms of heterodox updating I employ, including Jeffrey conditioning and Adams conditioning.

---

[19]In a recent presentation of some work in progress, James Joyce suggested a hierarchy of ideal experts, such that reports from higher-ranked experts trump reports from lower-ranked experts. In his example, learning the truth trumps learning the chances. This proposal could be built into my model, by taking only the report of the highest-ranked expert as a constraint (and perhaps retaining some memory of the source of one's credences, so as not to later have a lower-ranked expert override a higher). This could in principle be extended to cover any disagreeing experts, so long as the agent could order them by reliability, and assuming that the fact of their disagreement did not undermine the grounds for deference to them.

My proposal fits naturally with one proposal for rational awareness growth, Bradley's "rigid extension". Cases involving awareness growth are difficult, often leading to very poorly constrained imprecise credences. But this is not specific to my proposal (about expert testimony), it is a feature of awareness growth as a phenomenon. Indeed deference provides a simple way of setting attitudes on the new algebra that would otherwise be difficult to justify.

However, expert deference lacks the resources to deal with the final problem I highlighted: expert disagreement. (Though I noted that we can model some "easy" cases.) I note this limitation because I developed the model with an eye to offering more realistic guidance to policymakers interacting with experts. For them, disagreement is a present and severe challenge. But while this model provides a more tractable implementation of the deference idea, the way I idealised the experts means that it is limited to cases without disagreement.

Acknowledgements

References

Bradley R (2005) Radical Probabilism and Bayesian Conditioning. Philosophy of Science 72(2):342–364, DOI 10.1086/432427

Bradley R (2017) Decision Theory with a Human Face. Cambridge University Press

Buchak L (2013) Risk and Rationality. Oxford University Press

DeGroot MH (1988) A Bayesian view of assessing uncertainty and comparing expert opinion. Journal of Statistical Planning and Inference 20:295–306

Diaconis P, Zabell SL (1982) Updating Subjective Probability. Journal of the American Statistical Association 77(380):822–30

Dietrich F, List C, Bradley R (2016) Belief revision generalized: A joint characterization of Bayes' and Jeffrey's rules. Journal of Economic Theory 162:352–371, DOI 10.1016/j.jet.2015.11.006

Easwaran K, Fenton-Glynn L, Hitchcock C, Velasco JD (2016) Updating on the Credences of Others: Disagreement, Agreement, and Synergy. Philosopher's Imprint 16(11)

Elga A (2007) Reflection and Disagreement. Nous 41(3):478–502

Eva B, Hartmann S, Rad SR (2019) Learning from Conditionals. Mind pp 1–48, DOI 10.1093/mind/fzz025

French S (1980) Updating of Belief in the Light of Someone Else's Opinion. Journal of the Royal Statistical Society Series A (General) 143(1):43, DOI 10.2307/2981768

Gaifman H (1988) A Theory of Higher Order Probabilities. In: Causation, Chance, and Credence, vol 1, Kluwer, pp 191–220

Genest C, Zidek JV (1986) Combining Probability Distributions: A Critique and an Annotated Bibliography. Statistical Science 1(1):114–135, DOI 10.1214/ss/1177013825

Jeffrey R (1983) The Logic of Decision, 2nd edn. University of Chicago Press

Jeffrey R, Hendrickson M (1989) Probabilizing Pathology. Proceedings of the Aristotelian Society 89(1):211–226, DOI 10.1093/aristotelian/89.1.211

Joyce JM (1998) A Nonpragmatic Vindication of Probabilism. Philosophy of Science 65:575–603

Joyce JM (2007) Epistemic Deference: The Case of Chance. Proceedings of the Aristotelian Society 107:187–206, DOI 10.1111/j.1467-9264.2007.00218.x

Karni E, Vierø ML (2013) "Reverse Bayesianism": A Choice-Based Theory of Growing Awareness. American Economic Review 103(7):2790–2810, DOI 10.1257/aer.103.7.2790

Keeney RL, Raiffa H (1976) Decisions with Multiple Objectives: Preferences and Value Tradeoffs. John Wiley & Sons, New York

Lindley D (1982) The Improvement of Probability Judgements. Journal of the Royal Statistical Society, Series A 145(1)

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

Simion M, Kelp C, Ghijsen H (2016) Norms of Belief. Philosophical Issues 26(1):374–392, DOI 10.1111/phis.12077

Steele K (2012) Testimony as Evidence: More Problems for Linear Pooling. Journal of Philosophical Logic 41(6):983–999, DOI 10.1007/s10992-012-9227-5

Steele K, Stefánsson HO (forthcoming) Belief Revision for Growing Awareness. Mind

Thoma J (2019) Decision Theory. In: Pettigrew R, Weisberg J (eds) The Open Handbook of Formal Epistemology, PhilPapers Foundation

Titelbaum MG (forthcoming) Fundamentals of Bayesian Epistemology. Oxford University Press

van Fraassen B (1981) A Problem for Relative Information Minimizers in Probability Kinematics. The British Journal for the Philosophy of Science 32(4):375–379, DOI 10.1093/bjps/32.4.375

Wagner C (2002) Probability kinematics and commutativity. Philosophy of Science 69(2):266–78