Bayes Is Back

Alexander Meehan
University of Wisconsin–Madison

Snow Zhang
University of California, Berkeley

1. Introduction

How should rational agents revise their credences given new information? According to Bayesianism, the answer is Bayesian conditionalization. As a first approximation, the rule of Bayesian conditionalization says: if one begins at t_0 with a rational prior π, and p is the strongest proposition that one will learn between t_0 and t_1, then one should plan at t_0 to revise one's credences by adopting $\pi(.|p)$ at t_1.

Why is Bayesian conditionalization the rational way of revising one's credences? One prominent approach to answering this question appeals to some kind of bridge principle between *accuracy* and rational updating. The guiding thought is that epistemic rationality bears a close connection to *truth*, and this is what sets it apart from, say, prudential rationality. It may be prudentially rational to never change one's

This article benefited tremendously from a lot of feedback and advice. We would like to thank Aaron Bronfman, Jennifer Carr, Cian Dorr, Adam Elga, Hartry Field, Dan Greco, Harvey Lederman, Ginger Schultheis, and three anonymous referees for their extremely helpful and detailed feedback on earlier versions of this article. We would also like to thank Zachary Barnett, Kevin Dorst, Melissa Fusco, Simon Huttegger, Josh Knobe, Matthew Mandelkern, Pablo Zendejas Medina, Richard Pettigrew, Richard Roth, Bernhard Salow, and Patrick Wu for illuminating discussions. Earlier versions of this article were presented at the Berkeley–Stanford Logic Circle (2023), NYU Mind and Language seminar (2023), Central APA Symposium (2023), and Peking University Logic Seminar (2021). We are grateful to the participants for all their helpful comments.

Philosophical Review, Vol. 134, No. 3, 2025
DOI 10.1215/00318108-11873775

beliefs—perhaps doing so would make it much easier to organize one's life—but such a dogmatic policy is not epistemically rational, and part of the explanation of its irrationality has to do with it being non-truth-conducive.

Suppose this is right. A natural conjecture, then, is that Bayesian conditionalization is the rational updating rule because it is most accurate. However, a moment of reflection suggests that this can't be right. For surely, the most accurate updating rule is the truth rule—the rule that tells the agent to be certain of the truth and nothing but the truth. Bayesian conditionalization disagrees with the truth rule in all but the exceptional case where the agent learns all the truths. Why, then, is Bayesian conditionalization the rule of rational credal revision and not the truth rule, even though the latter is always more accurate?

One tempting reply to this challenge is that an updating rule must be *followable* in order to be rational; ought implies can. What does it take for a rule to be followable? A popular answer is that an updating rule is followable just in case it is *evidentially constant*—it recommends the same credence function at any two states where the agent learns the same evidence. Bayesian conditionalization satisfies this condition, whereas the truth rule does not. Moreover, Hilary Greaves and David Wallace (2006) prove that, in any transparent updating problem—case where it is always part of the agent's evidence what is and is not part of their evidence—Bayesian conditionalization has the highest expected accuracy of all updating rules that are evidentially constant in the above sense. This result suggests a promising answer to the question we posed above: Bayesian conditionalization is the rule of rational credal revision because, of all followable updating rules, it is most accurate.

One wrinkle with this answer, as pointed out by Aaron Bronfman (2014) and Miriam Schoenfield (2018), is that it presupposes *evidence internalism*, the thesis that every updating problem is transparent. Evidence internalism is controversial. In particular, it is incompatible with the prima facie plausible hypothesis that, in some cases, what one learns may depend sensitively on facts of one's environment without themselves entailing those facts. For instance, when one looks at a tree from afar, one may not learn the precise height of the tree, even though what one does learn about the height of the tree may vary depending on the precise height of the tree (Williamson 2000). If this is right, then it is possible for one to learn p without learning that one learns p, that is, evidence internalism is false.

Whether the above argument succeeds remains a matter of philosophical dispute, and we won't try to settle it here. Instead, for the purposes of this article, we will assume that evidence internalism is false, or equivalently that *evidence externalism* is true. The explanatory challenge the Bayesian then faces can be stated as follows: Assuming evidence externalism, can we explain why Bayesian conditionalization is the rational updating rule, and the truth rule is not, without giving up on the idea that there is a substantive connection between the epistemic rationality of updating rules and their accuracy?

A result due to Schoenfield (2017) suggests that the answer is No. Schoenfield shows that, in general, the updating rule that has the highest expected accuracy of all evidentially constant updating rules is *meta-conditionalization*, not Bayesian conditionalization. Roughly, meta-conditionalization says: if one begins with a rational prior π at t_0 and learns p between t_0 and t_1 and nothing else, then one should plan at t_0 to revise one's credence at t_1 by conditionalizing on the proposition *that one has learned* p *and nothing else between* t_0 *and* t_1.[1] In particular, meta-conditionalization and Bayesian conditionalization disagree whenever the agent learns p but doesn't learn that she learns p. In such cases, assuming the agent has prior uncertainties about whether she will learn p, Bayesian conditionalization says she shouldn't be certain that she learns p when she does learn p, whereas meta-conditionalization says she should.

What shall we make of this result? The standard reply from many externalists is to give up on the explanatory project.[2] For instance, some

1. Schoenfield (2017) calls this rule *conditionalization**; Matthias Hild (1998) and Pablo Zendejas Medina (2024) call this rule *auto-epistemic conditionalization*; we adopt our terminology from Nilanjan Das (2019). Schoenfield also proves that Bayesian conditionalization does not maximize expected accuracy if evidence is nonfactive. Whether evidence is factive is controversial (see, e.g., Mitova 2018 and references therein). For the purposes of this article, we shall assume factivity of evidence and focus on vindicating Bayesian conditionalization as a rule for updating in factive but possibly nontransparent evidential situations.

2. One exception is Medina (2024). Medina (2024) gives an externalist-friendly justification for Bayesian conditionalization based on *conditional* expected accuracy rather than unconditional expected accuracy. More specifically, he shows that, given certain implementability assumptions about updating plans, for each proposition p, conditionalizing on p is the best conditional updating plan for p. It will take us too far afield to discuss Medina's nuanced proposal in detail, but it is worth noting that, as we will see, there are cases where Bayesian conditionalization can have lower unconditional expected accuracy than retaining one's prior. So, assuming evidence externalism, it is

take Bayesian conditionalization to be a corollary of the primitive norm that rational credence should conform to evidential probability, which is defined as the probability conditional on one's total evidence. Accuracy-based considerations are neither here nor there (see, e.g., Williamson 2000; Das 2022). Other externalists accept the guiding thought that rational updating rules are truth-conducive, and conclude that Bayesian conditionalization is *not* the rule of rational credal revision, though nor is meta-conditionalization or the truth rule. Rather, whether an updating rule is truth-conducive and thereby rational for a given agent depends on how the agent would update if she were to *try to* follow the rule (Gallow 2021; Isaacs and Russell 2023; Schultheis forthcoming). Bayesian conditionalization is only rational for special agents under special circumstances. On this proposal, the explanatory challenge is dissolved, as there is nothing to be explained in the first place.

This article proposes a different response: we argue that the externalist Bayesian can uphold Bayesian conditionalization as the rule of rational credal revision without giving up on the idea that what is rational is also truth-conducive (in a certain sense). The core of our proposal is a new accuracy-rationality bridge principle, which has two main ingredients. First, we introduce a natural strengthening of evidential constancy, which we call *global evidential constancy*. Roughly, an updating rule satisfies global evidential constancy if, for any propositions p and q, if p is the strongest proposition that the agent has learned, then the credence that the rule recommends the agent to assign to q doesn't depend on the state at which the agent learns p or what she could have learned other than p. We think global evidential constancy is a plausible normative constraint on updating rules. In particular, Bayesian conditionalization satisfies global evidential constancy, whereas the truth rule and meta-conditionalization do not.

However, ruling out the truth rule and meta-conditionalization is not enough, for there are updating rules other than Bayesian conditionalization that also satisfy global evidential constancy and have higher expected accuracy than Bayesian conditionalization in certain updat-

possible that the *unconditional* plan (i.e., a set of conditional plans whose conditions form a partition) that maximizes unconditional expected accuracy is *not* the conjunction of the best conditional plans, each of which maximizes conditional expected accuracy. To this extent, it remains a puzzle why Bayesian conditionalization gives the correct prescription for which *unconditional* plan a rational agent should adopt in nontransparent updating problems.

ing problems. In fact, there are cases where retaining one's prior has higher expected accuracy than updating by Bayesian conditionalization. So, insofar as Bayesian conditionalization is rational, it isn't rational in virtue of having the highest expected accuracy (of all updating rules that satisfy global evidential constancy) in every updating problem—it does not.

This brings us to the second important ingredient of our proposal, which is a different way of evaluating the accuracy of an updating rule. Some rules, such as "do not jaywalk," may be good rules because they are best overall, even though following them is suboptimal on some specific occasions. Similarly, an updating rule may be accurate in general, even if updating in accord with the rule yields less accurate credences than alternative rules in some particular updating problems. We propose that this is the sense in which BCOND is truth-conducive. That is, the externalist Bayesian should accept the following accuracy-rationality bridge principle:

> **Highest Total Expected Accuracy (Hi-TEA)**. An updating rule is epistemically rational if and only if it is globally evidentially constant and, of all the updating rules that satisfy global evidential constancy, it is the most accurate in general, in the sense that it has the highest total expected accuracy across updating problems.

We show that, of all updating rules that satisfy global evidential constancy, Bayesian conditionalization has the highest total expected accuracy across updating problems associated with any given prior space. Thus, the proposed bridge principle yields Bayesian conditionalization as the rational rule, even under conditions of evidence externalism. Bayes is back!

Here is a roadmap for the rest of the article. Section 2 lays down the setup for modeling updating problems. Section 3 introduces the accuracy framework and the result of Schoenfield (2017), and discusses the challenges with appeals to followability. Sections 4 and 5 present and defend the positive proposal. Section 6 concludes.

2. Setup

2.1. Updating Problems

We learn information about the world; which information we learn depends on which state of the world obtains. For instance, suppose you toss a die. Maybe you will be told by a reliable informant exactly how it lands. Maybe the informant will only tell you whether the die lands on an even number or an odd number, but nothing more specific. Or maybe you will be able to see the die for yourself, but only through a piece of frosted glass. Intuitively, these different learning situations call for different plans of credal revision. Perhaps you should plan to be certain about exactly how the die lands in the case where the reliable informant will tell you exactly how it lands, but not so if the reliable informant will only tell you whether the die lands even or odd.

We are interested in the question: Prior to your learning experience, how should you plan to revise your credences, given each and every possible piece of information that you might learn? To make this question precise, we'll start with the standard Bayesian modeling assumption that an agent can be associated with a prior probability function π defined over a set of states S. We'll call such a pair (S, π) a **prior space**. Heuristically, one could think of S as a partition of the logical space whose cells settle every salient question in the given context. Unless otherwise specified, we'll assume that S is finite and contains at least two elements, and π is regular, that is, it assigns nonzero probability to every state in S.[3]

Given a prior space (S, π), a **learning situation** relative to (S, π) specifies, at each state s, what the agent thinks she will learn at s. For the purposes of this article, we'll make the idealizing assumption that as a learner sets out to gather information about the world, she is certain that, in each of her possible courses of learning experiences, she will learn a unique, true proposition that is equivalent to a disjunction of states (given her background knowledge). We'll call such a proposition "the strongest proposition learned" at the respective state, or sometimes "the evidence proposition."[4] A learning situation is then a complete conjunc-

3. One may think of S as the set of state descriptions in the sense of Rudolf Carnap (1988), or the atoms of a given agenda (Pettigrew 2016). Importantly, we do not take S to be the space of all metaphysical possibilities. This is partly motivated by the assumption of finiteness, which is necessary given that many accuracy-theoretic results do not generalize to credence functions over infinite domains (Kelley 2023; Pruss 2022; Nielsen 2022; Kelley and Neth 2022). The restriction to regular priors is relatively innocuous given the finiteness assumption and will significantly streamline the presentation. Most existing arguments against imposing regularity as a rational requirement concerns its plausibility in the infinite context (see, e.g., Hájek 2012; Easwaran 2014). See Pettigrew 2016 for an accuracy-based argument for regularity in the finite setting.

4. So we assume that evidence is factive. While we are not wedded to the assumption ourselves, we think it is plausible given evidence externalism. For instance, one influential externalist view of evidence is that evidence is knowledge (Williamson 2000). Since knowledge is factive, it follows that evidence is factive as well. Since our goal is

tion of conditionals that specify the strongest propositions that a learner will learn at each possible state in her information-gathering process. Given this assumption, each learning situation can then be modeled by a function E that maps each state s to a set of states $E(s)$ (a subset of S) that contains s as a member. We'll call a triple of the form (S, π, E) an **updating problem**.

An **updating rule** specifies, for each updating problem (S, π, E), how a learner facing that updating problem should plan to update her prior π over states, that is, what new credence function over S she should plan to have at each state. Formally, we represent an updating rule by a map that takes as input an updating problem (S, π, E), and outputs a function $f_{(S,\pi,E)}$ from states in S to posterior probability functions over S. The intended interpretation is: if a learner faces the updating problem (S, π, E), then she should plan to adopt posterior $f_{(S,\pi,E)}(s)$ over S if s obtains.

Here are some updating rules that will feature prominently in our subsequent discussion:

- Bayesian conditionalization (BCOND): If one faces the updating problem (S, π, E), then one should plan, for each state s, to adopt the credence function over S given by $\pi(.|E(s))$:

$$b_{(S,\pi,E)}(s) = \pi\left(.\middle|E(s)\right).$$

- Meta-conditionalization (META-COND): If one faces the updating problem (S, π, E), then one should plan, for each state s, to adopt the credence function over S given by $\pi(.|E = E(s))$:

$$m_{(S,\pi,E)}(s) = \pi\left(.\middle|E = E(s)\right),$$

  where $[E = E(s)]$ is shorthand for the following set of states: $\{s' \in S : E(s') = E(s)\}$.
- The truth rule: If one faces the updating problem (S, π, E), then one should plan, for each state s, to adopt the credence function over S given by v_s—the credence function that assigns probability 1 to s and 0 to every other state:

$$t_{(S,\pi,E)}(s) = v_s.$$

to explore the implications of externalism, this makes it a natural assumption to adopt here.

- The dogmatic rule: If one faces the updating problem (S, π, E), then one should plan, for each state s, to retain one's prior over S:

$$d_{(S,\pi,E)}(s) = \pi.$$

To conclude this subsection, let us apply this framework to a concrete example.

> *Die.* Xani and Yota are doing an experiment on a fair four-sided die. The die will be tossed in a black box. There are two different sensors in the box. The first sensor simply detects whether the die lands even or odd. The second sensor is more sensitive: if the die lands even, it also only detects that the die lands even, but if the die lands odd, it detects exactly how the die landed. Xani is going to get the information from the first sensor, while Yota will get the information from the second. Both Xani and Yota know everything about this setup before t_0, and neither forgets anything in the course of learning how the die landed.

Here is a simple model of *Die*: Let $S = \{s_1, s_2, s_3, s_4\}$, where s_i represents the state where the die landed on i. Let π be the uniform prior over S. Then Xani faces the updating problem (S, π, E_X), where E_X maps a state s_i to the set $\{s_2, s_4\}$ if i is even and the complement of that set if i is odd. Yota, on the other hand, faces a slightly different updating problem (S, π, E_Y), where E_Y maps a state s_i to the set $\{s_2, s_4\}$ if i is even, and the singleton $\{s_i\}$ if i is odd. The output of the BCOND rule for these two updating problems can be pictured as shown in figure 1.

2.2. *Externalism and Evidence*

Let us zoom in on the case of *Die.* One thing to note about Xani's and Yota's updating problems is that, for any proposition p that these agents might learn, they are certain that if they learn p, then they also learn *that they learn p*. For instance, consider Xani. Suppose Xani learns that the die landed odd. By stipulation, ex ante, Xani is certain of the biconditional: she learns that the die landed odd if and only if the die landed odd. Moreover, she retains that piece of knowledge after she learns that the die landed odd. So, assuming that learning is closed under logical entailment given one's background knowledge, when Xani learns that *the die landed odd,* she also learns that *she learns that the die landed odd.*

Formally speaking, Xani's learning situation satisfies the following property:

$$E(s) \subseteq \left[E = E(s)\right] \quad \text{for all } s \in S. \tag{1}$$

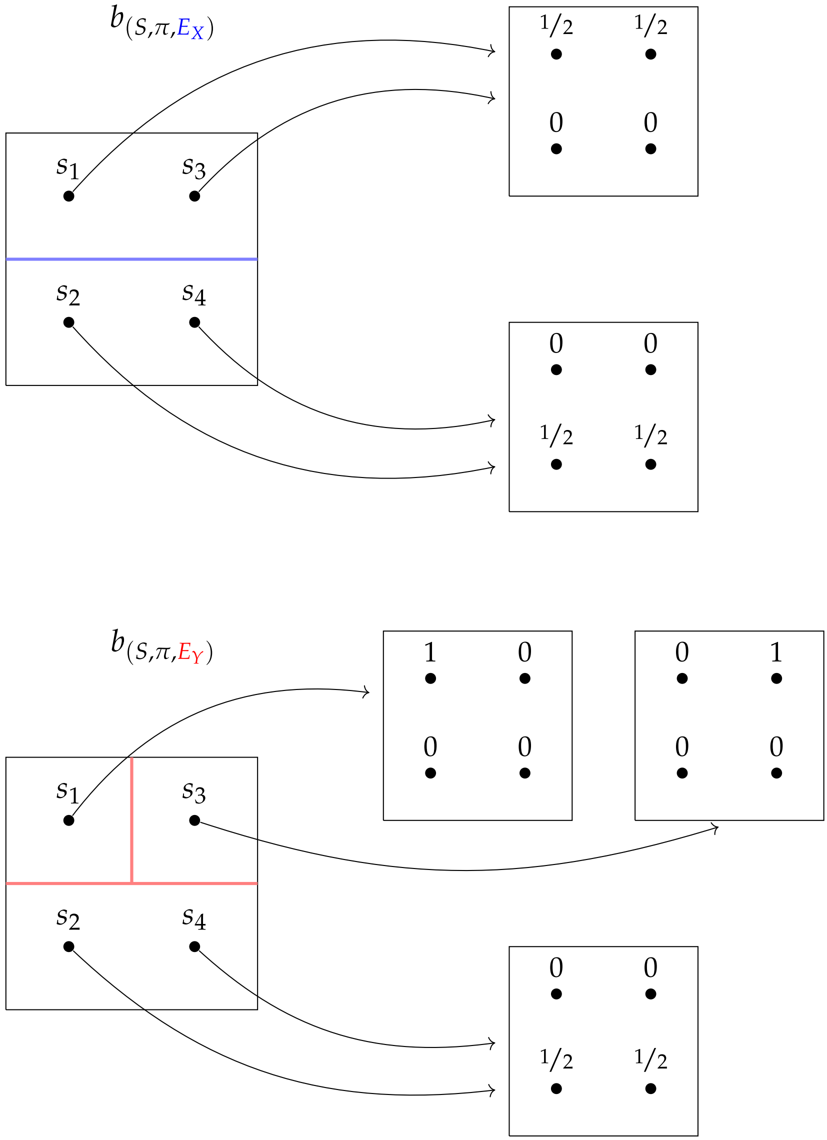

Figure 1. Illustration of BCOND as applied to the *Die* case.

We say an updating problem (S, π, E) is **transparent** if E satisfies (1); an updating problem is **opaque** if it is not transparent.[5] Let **evidence internalism** be the view that all updating problems are transparent, and let **evidence externalism** be the view that some updating problem are opaque.

Is evidence internalism true? Naively, one might think the answer is No. After all, in most cases, p doesn't entail *that one has learned p*. In general, if p doesn't entail q, then there are possible circumstances in which it is rationally permissible to be uncertain whether one will learn p without learning q. It is unclear why p and *that one has learned p* should be an exception.

Apart from this motivating consideration, there are also compelling theoretical considerations against evidence internalism. One such argument is the margin-for-error argument due to Timothy Williamson (2000). The idea is that, in some cases, the strongest propositions that one learns in a given context may depend sensitively on facts of one's environment without themselves entailing those facts. If this is right, then there are cases where it is possible that the agent learns p without learning that she learns p, that is, evidence internalism is false.

Here is a concrete illustration of the argument:[6]

> *Olivia.* Olivia is making cookies. Unfortunately, her scale is broken, so she has to estimate the amount of butter she'll use by hand. She takes out the leftover butter, which she knows weighs no less than 15 oz., and no more than 20 oz. Olivia is quite experienced. In particular, she knows that, if the butter weighs x oz., then she will learn by weighing it in her hand that it weighs $(x \pm 1)$ oz., but nothing more precise.

To see that Olivia's learning situation is opaque, suppose the butter weighs 17 oz. Then by stipulation, after feeling the butter by hand, the strongest proposition that Olivia will learn about the weight of the butter is that the butter weighs between 16 oz. and 18 oz. But if the butter had weighed 18 oz.—a state compatible with what Olivia has learned at the 17 oz. state—then the strongest proposition that Olivia would have learned about the weight of the butter is that the butter weighs between

5. One can check that a (factive) learning situation E is transparent, i.e., not opaque, if and only if it is *partitional*, in the sense that the collection $\{E(s) : s \in S\}$ forms a partition of S.

6. This is a variant of the austere clock from Williamson (2011), which many take to be metaphysically possible; see, e.g., Elga 2013; Gallow 2019b; Horowitz 2014; Ahmed and Salow 2019.

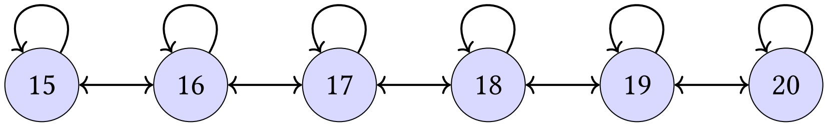

Figure 2. A simple model of Olivia's learning situation ($w \to v$ means $v \in E(w)$).

17 oz. and 19 oz. So in the case where the butter weighs 17 oz., Olivia can't rule out the possibility that she learns something different from what she actually learns; her learning experience is opaque. A simple model of Olivia's learning situation is given in figure 2.

Now, one might object that figure 2 is a misrepresentation of Olivia's learning situation. Perhaps there is no unique, true proposition that she has learned when she weighs the butter by hand (Jeffrey 1965). Or perhaps her possible evidence consists of propositions about her perceptual appearances, such as *the butter weighs like this*, and she has perfect introspective access to her own perceptual appearances. The issue is complicated, and we do not have a settled view on this matter.[7] However, for the purposes of this article, we will assume that cases like *Olivia* are possible for ideal rational agents, and can be approximately represented by Kripke models like the one depicted in figure 2. We will also take for granted that, even in such cases, the rational updating rule is BCOND. Our question is: *Why* (and in what sense) is BCOND the rational updating rule, assuming that it is (in some sense)? As we shall see, a satisfactory answer to this question is harder to obtain than one might have expected.

3. Accuracy, Rationality, and Bayes Lost

Many think that there is a tight connection between accuracy and epistemic rationality. The aim of belief is truth; the aim of partial belief is accuracy.[8] Given this intuitive connection, a natural thought is that an updating rule is rational just in case all of its recommended updating plans maximize expected accuracy.

7. For arguments in favor of nontransparency, see, e.g., Williamson 2000, 2011; Weatherson 2013; Srinivasan 2015. For considerations against nontransparency as conceived here, see Berker 2008; Stalnaker 2015; Salow 2018; Greco 2019. One could also be a pluralist about learning, and think that transparency holds for one notion of learning but not another. See Pettigrew 2022 for related discussion.

8. For a partial list of discussions and precisifications of this intuitive idea, see, e.g., Joyce 1998; Gibbard 2007; Carr 2017; Wedgwood 2002; Horwich 2006.

To make this thought more precise, we need a quantitative measure of accuracy—an **accuracy scoring rule**. Formally, an accuracy scoring rule is a map A that, for each state partition S, associates a pair of a credence function c over S and a state $s \in S$ with a real number, $\mathsf{A}(c, s)$. Intuitively, $\mathsf{A}(c, s)$ represents how accurate c is at s. Throughout this article, we will require that A is **strictly proper**: for any probabilistic credence function c over a state space S, c expects itself to be strictly more accurate than any other credence function c' defined over the same domain.[9] This assumption is widely shared in the accuracy literature (though not uncontroversially), and we won't dispute it here.[10] Note that, if A is strictly proper, then for any state s, the most accurate credence function at s according to A is the omniscient credence function v_s that assigns probability 1 to all true propositions at s and 0 to all false propositions at s.[11]

Now, fix a strictly proper scoring rule A that represents a correct way of measuring accuracy. The **expected accuracy of an updating rule**

9. Formally, $\sum_{s \in S} c(s)\mathsf{A}(c, s) \geq \sum_{s \in S} c(s)\mathsf{A}(c', s)$ with equality if and only if $c = c'$.

10. See, e.g., Joyce 2009; Pettigrew 2016; Greaves and Wallace 2006; Schoenfield 2017; Briggs and Pettigrew 2020; Campbell-Moore and Levinstein 2021. Strict propriety is also widely assumed in the statistical literature on scoring rules for probabilistic forecasts (e.g., Predd et al. 2009; Gneiting and Raftery 2007; Bröcker and Smith 2007). While we will assume strict propriety in this article, we think an externalist could challenge this assumption. In particular, the standard argument for propriety appeals to the assumption that rationality requires immodesty—not expecting oneself to be less accurate than someone else—which seems like an alethic version of the enkratic principle. Like the akratic agent who believes p but also that they shouldn't believe p, an agent with a modest credence function c holds onto c while judging a different credence function to be better accuracy-wise. Many externalists reject the enkratic principle on the grounds that it conflicts with the possibility of self-misleading evidence (e.g., Horowitz 2014; Dorst 2019; Weatherson 2019; Worsnip 2018; Williamson 2011). Given the parallel, we think an externalist BCOND can get around the negative results mentioned below by rejecting strict propriety, though to what extent this is enough to vindicate BCOND as the rational updating rule remains to be seen. There are also other arguments for strict propriety that do not appeal directly to immodesty-type principles (e.g., the calibration argument by Richard Pettigrew (2016); see Levinstein 2017 for objections and Williams and Pettigrew forthcoming for a reply).

11. *Proof.* Suppose A is strictly proper. Then for any s and credence function c not identical to v_s, $\mathsf{A}(v_s, s) = \mathbb{E}_{v_s}[\mathsf{A}(v_s)] > \mathbb{E}_{v_s}[\mathsf{A}(c)] = \mathsf{A}(c, s)$. It's worth noting that, although formally speaking A could be interpreted as any measure of epistemic value, this consequence of strict propriety makes it difficult to interpret A as measuring anything other than accuracy. As a result, strict propriety is incompatible with views that take the "aim" of rational credence to be something different from truths, e.g., objective chance (Hájek 2011; Pettigrew 2012) or evidential probabilities. So proponents of these views could also get around the negative results below by rejecting strict propriety.

f **in an updating problem** (S, π, E) is defined as the weighted average of the accuracy scores of the credence functions that *f* recommends in the updating problem at each state *s*, weighted by the agent's prior probability that *s* obtains:

$$\mathbb{E}_\pi\left[\mathsf{A}(f_{(S,\pi,E)})\right] := \sum_s \pi(s)\mathsf{A}\big(f_{(S,\pi,E)}(s), s\big).$$

With the notion of expected accuracy in hand, we can more precisely state the intuitive connection between epistemic rationality and expected accuracy:

> **Highest Expected Accuracy (HEA)**. An updating rule is epistemically rational if and only if it has the highest expected accuracy in any updating problem.

While the principle may seem intuitively plausible, a moment of reflection suggests that it can't be right. For consider the truth rule—the rule that says, in any updating problem, always assign probability one to the true state. As noted above, in any updating problem, the recommendation of this rule is always maximally accurate. So any prior must expect this rule to have higher expected accuracy than any alternative rule that disagrees with it with positive probability. Therefore, HEA entails that the only rational updating rule is the truth rule. But arguably, this is absurd. Uncertainty seems an inevitable feature of our epistemic lives. If a principle rules that out as epistemically irrational, then so much the worse for that principle.

What, then, is the relationship between accuracy and epistemic rationality? In particular, for the Bayesians: *Why* is BCOND the rational updating rule, despite being less accurate than the truth rule?

One tempting answer to this question is that the truth rule is not even a genuine competitor to BCOND because it is *unfollowable*. The rational updating rule must be followable by an (idealized) agent; ought implies can. Bayesian conditionalization is rational, the thought goes, because and to the extent that it is the most accurate *followable* updating rule.

While we are sympathetic to the idea that followability is relevant to *a* notion of epistemic rationality, we are skeptical that appealing to followability would help explain the sense in which BCOND is rational or most accurate. The rest of this section explains why.

3.1. Bayes Justified

What does it mean for an updating rule to be followable? Let us start with a popular criterion, which we'll call *evidential constancy*. An updating rule is **evidentially constant** if, in any updating problem, the rule doesn't recommend adopting different credence functions at two states where the agent learns the same information. Formally: f is evidentially constant if, given any updating problem (S, π, E), for any $s, s' \in S$,

$$E(s) = E(s') \Rightarrow f_{(S,\pi,E)}(s) = f_{(S,\pi,E)}(s').$$

While evidential constancy has not been explicitly defended as a necessary and sufficient condition for followability, it is often assumed in the background. For instance, Pettigrew (2016: 191) writes, "[U]pdating rules... have to give the same recommendation for any two worlds at which the same element of the partition is true. The reason for this restriction is that updating rules have to be something the agent might follow; and she can only follow a rule if, whenever she can't distinguish two possibilities, the rule doesn't distinguish those possibilities either, and gives the same recommendation for both."[12] It is straightforward to check that the truth rule is not in general evidentially constant, whereas BCOND is. Moreover, Greaves and Wallace (2006) prove that, in any transparent updating problem, BCOND *is* the most accurate updating rule that satisfies evidential constancy.

Fact 1 (Greaves and Wallace 2006). *In any transparent updating problem, BCOND has the highest expected accuracy of all updating rules that satisfy evidential constancy.*

This result suggests a promising answer to our question: the rational updating rule is BCOND rather than the truth rule because (i) the truth rule is unfollowable, as it violates evidential constancy, and (ii) of all followable updating rules, that is, rules that satisfy evidential constancy,

12. Evidential constancy corresponds to a generalization of the notion of "availability" in the sense of Greaves and Wallace (2006: 612), who motivate the constraint on similar grounds: "When [*a* is not available], *a* is an act that the agent is not able to perform given only the experiment *E*, since performing act *a* would require the agent to respond to information that he does not have (hence our refusal to call such acts 'available')." See also the discussions in Barnett 2021; Schoenfield 2017; Das 2022. Similar constraints have been proposed in the economics literature, e.g., Geanakoplos 1989.

BCOND is most accurate by having the highest expected accuracy in every transparent updating problem.

3.2. Bayes Lost

While the above answer might be good enough for evidence internalists, it will not satisfy the externalist Bayesian who thinks that cases like *Olivia* are possible and fall under the scope of BCOND as well. A natural question arises: Can we generalize Greaves and Wallace's (2006) result to opaque updating problems? In particular, does following BCOND have the highest expected accuracy (of all updating rules that satisfy evidential constancy) in cases like *Olivia*?

Unfortunately, the answer is No. To see this, recall that Olivia learns that the butter weighs $x \pm 1$ oz. if and only if the butter weighs x oz. As a result, she *never* learns the same proposition about the weight of the butter at two different states. Since evidential constancy only requires that the updating rule doesn't recommend different credence functions at two states where the agent learns the same proposition, this constraint is vacuously satisfied by any updating rule in Olivia's case. In particular, it is satisfied by the truth rule. As noted above, following the truth rule always has higher expected accuracy than following BCOND (unless they agree). So in *Olivia*, BCOND does not have the highest expected accuracy of all updating rules that satisfy evidential constancy; the truth rule does.

As it turns out, this is a special case of a more general result. Recall the rule of meta-conditionalization, which says that in every updating problem (S, π, E), if the agent's evidence proposition is p, then the agent should update by conditionalizing on the proposition *that p is her evidence proposition*, that is, if $E(s) = p$, then $m_{(S,\pi,E)}(s) = \pi(.|E = p)$. It's straightforward to check that META-COND satisfies evidential constancy: let (S, π, E) be any updating problem. Suppose $E(s) = E(s') = p$. Then

$$m_{(S,\pi,E)}(s) = \pi(.|E = p) = m_{(S,\pi,E)}(s').$$

If the agent always learns what she does and does not learn, that is, $E(s) \subseteq [E = E(s)]$ for every s, then META-COND coincides with BCOND. This is true of Xani and Yota. On the other hand, if the agent learns different propositions at different states, that is, if $E(s) \neq E(s')$ for any distinct states s and s', then META-COND coincides with the truth rule. This is the case for Olivia.

Schoenfield (2017) proves that, of all updating rules that satisfy evidential constancy, updating by META-COND always has the highest expected accuracy.

Fact 2 (Schoenfield 2017). *In any updating problem,* META-COND *has the highest expected accuracy of all updating rules that satisfy evidential constancy.*

At this point, it appears that we have not made much progress after all. While we have ruled out the truth rule by appealing to the constraint of evidential constancy, we have not ruled out META-COND, which is also more accurate than BCOND. So our question remains: *Why* is BCOND the rational updating rule, despite being less accurate than META-COND?

3.3. Stronger Conditions for Followability?

A natural reply to Schoenfield's (2017) result is that evidential constancy is necessary for followability but not sufficient. After all, it seems that Olivia *can't* follow META-COND in her updating problem, for the same reason that Xani or Yota can't follow the truth rule in their updating problems. In both cases, the rules tell them to be maximally certain in some truths that are not entailed by their evidence. Intuitively, this is not something that an agent can do reliably, not even if they have unlimited cognitive power.

Say an updating rule is **conservative** if it doesn't recommend certainties in propositions that are not part of the agent's evidence. Formally: f is conservative if for any updating problem (S, π, E), any $s \in S$, and any proposition $p \subseteq S$, $f_{(S,\pi,E)}(s)(p) = 1$ only if $\pi(p) = 1$ or $E(s) \subseteq p$ (see, e.g., Bronfman 2014; Das 2022). The above thought suggests that we should strengthen the criterion for followability to at least evidential constancy + conservativity. One can check that BCOND is conservative in this sense: if $\pi(p) < 1$, then $\pi(p|E(s)) = 1$ only if $E(s) \subseteq p$. On the other hand, neither META-COND nor the truth rule is conservative, as demonstrated by *Olivia*.

We think there are at least three obstacles to this reply. Let us go through them one by one. First, let's consider a simplified version of *Olivia*:

> *Dana.* Olivia's friend Dana is also making cookies. Like Olivia, Dana's scale is also broken, so she has to weigh everything by hand. Unlike Olivia, Dana knows in advance that the butter she has left weighs

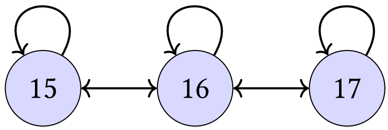

Figure 3. A simple model of Dana's learning situation.

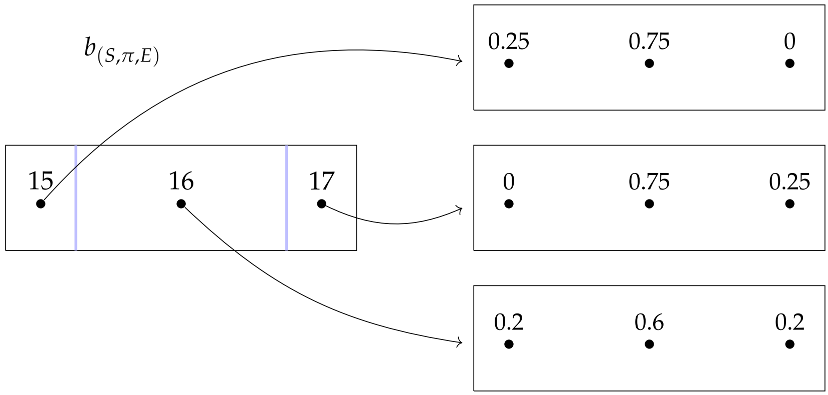

Figure 4. Illustration of BCOND in Dana's updating problem.

> between 15 and 17 oz. She is also relatively confident that the butter weighs 16 oz.—say 60 percent—and splits her remaining credence equally between the other two possibilities. Still, if she feels the butter by her hand, if the butter weighs 15 oz., she can rule out the possibility that it weighs 17 oz., but can't rule out that it weighs 16 oz.; similarly if the butter weighs 17 oz. Lastly, if the butter weighs 16 oz., then she could go either way.

Dana's learning situation is pictured in figure 3, and the recommendation of BCOND for Dana is pictured in figure 4. (Note that if the butter weighs 16 oz., then BCOND recommends that Dana retain her prior.)

Prima facie, if *Olivia* is possible, so is *Dana*. However, it turns out that, if we measure accuracy by the Brier score or the additive logaithmic score, then updating by BCOND has lower expected accuracy than retaining one's prior in *Dana* (proof in the appendix; see also table 1).

Table 1. A comparison of the accuracy of Bayesian conditionalization and the dogmatic rule in Dana's updating problem, according to the Brier score. The first column lists each state s. The second column specifies the evidence proposition that Dana learns at s. The third and fourth columns give the accuracy (at s) of the dogmatic rule and BCOND rule, respectively. The final column gives the difference in accuracy between the two rules.

s	$E(s)$	$\mathsf{A}(d_{(S,\pi,E)}(s), s)$	$\mathsf{A}(b_{(S,\pi,E)}(s), s)$	$\mathsf{A}(b_{(S,\pi,E)}(s), s) - \mathsf{A}(d_{(S,\pi,E)}(s), s)$
15	$\{15, 16\}$	-2.08	-2.25	-0.17
16	$\{15, 16, 17\}$	-0.48	-0.48	0
17	$\{16, 17\}$	-2.08	-2.25	-0.17

Fact 3. *Suppose that accuracy is measured by the Brier score or the additive logarithmic score. Then the dogmatic rule has higher expected accuracy than* BCOND *in Dana's updating problem.*[13]

Presumably, the dogmatic rule is followable if any updating rule is followable; what could be easier than not updating at all? But supposing it is not, this shows that evidential constancy and conservativity are not jointly sufficient for followability either, as the dogmatic rule satisfies both. We would need something stronger still.[14]

Our next result suggests that, even if we can find such a criterion, it can't be strong enough to rule out all updating rules that have higher expected accuracy than BCOND in *Dana*. For a start, let ϵ be some arbitrarily small but positive real number. Consider the updating rule h that

13. One might wonder if this result contradicts Irving John Good's theorem that it is always valuable to learn free information (Good 1967). The answer is No, because the setup for Good's theorem presupposes that the agent's updating problem is transparent. A corollary of fact 3 is that cost-free information is not always epistemically valuable for Bayesians if learning is opaque (assuming that accuracy is measured by the additive Brier score or the additive logarithmic score). See Ahmed and Salow 2019; Das 2023; Neth forthcoming; Geanakoplos 1989 for related results and discussion.

14. Admittedly, the result does not go through for certain choices of accuracy scores, e.g., the simple nonadditive logarithmic score $\mathsf{A}(c, s) = \ln c(s)$. Indeed, under this scoring rule, the expected accuracy of BCOND is higher than that of retaining one's prior, unless the two agree. (*Proof*: Let (S, π, E) be an updating problem. Then $\mathbb{E}_\pi[\mathsf{A}(b_{(S,\pi,E)}) - \mathsf{A}(\pi)] = \sum_s \pi(s)(\ln \pi(s|E(s)) - \ln \pi(s)) = \sum_s \pi(s)(\ln \pi(s) - \ln \pi(E(s)) - \ln \pi(s)) = -\sum_s \pi(s) \ln \pi(E(s)) > 0$.) So one could take this case as motivation to reject the Brier score and the additive logarithmic score as appropriate scoring rules for assessing accuracy. See Bernardo 1979 and Lewis and Fallis 2021 for related discussion. We thank an anonymous referee for noting this connection. It is worth noting that the subsequent argument involving mixtures does not rely on this restriction, and extends to the simple logarithmic score as well.

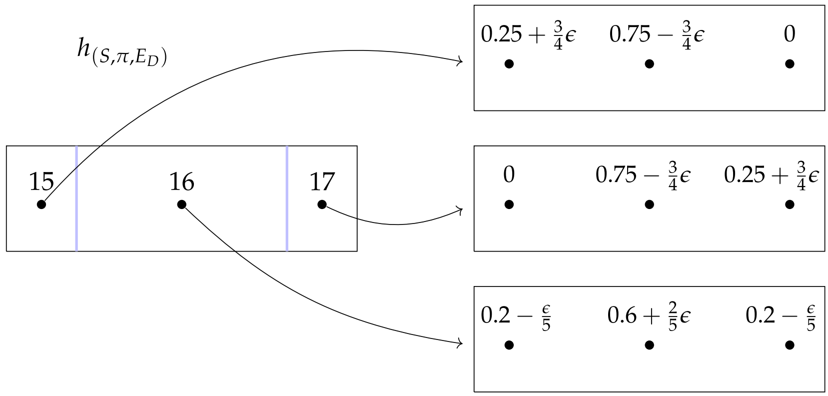

Figure 5. Illustration of h in Dana's updating problem.

agrees with BCOND in every other updating problem except for Dana's, and, in Dana's updating problem, recommends the plan shown in figure 5.

Notice that this rule is evidentially constant and conservative. However, if we measure accuracy by the Brier score, then this updating rule turns out to have higher expected accuracy than BCOND in Dana's updating problem (and trivially the same expected accuracy as BCOND in every other updating problem).

This result generalizes. Say an accuracy scoring rule A is **concave** if, for any state $s \in S$ and credence functions c, c' over S, the accuracy score of any weighted average of c and c' at s is at least as high as the corresponding weighted average of the accuracy scores of c and c' at s:[15]

$$\mathsf{A}\big(\lambda c + (1-\lambda)c', s\big) \geq \lambda \mathsf{A}(c, s) + (1-\lambda)\mathsf{A}(c', s), \quad \forall \lambda \in [0, 1].$$

Most well-known examples of strictly proper scoring rules are concave, including the Brier score, the additive logarithmic score, and the simple logarithmic score. It is controversial whether all legitimate accuracy scoring rules need be concave (Joyce 1998, 2009), but most agree that some legitimate accuracy scoring rules are concave (e.g., the Brier score is one legitimate accuracy scoring rule), which is sufficient for our purposes.

15. If we identify the *in*accuracy score of a credence function c at a state s with its negative accuracy score, then an accuracy scoring rule is concave just in case the associated inaccuracy scoring rule is convex. The latter is more well known in the literature.

Let f and g be two updating rules. We say an updating rule h is a **mixture of** f **and** g if for every updating problem (S, π, E), there is some $0 \leq w \leq 1$ such that

$$h_{(S,\pi,E)}(s) = w \cdot f_{(S,\pi,E)}(s) + (1 - w) \cdot g_{(S,\pi,E)}.$$

One can check that the updating rule h described above is a mixture of BCOND and META-COND, where the weight assigned to BCOND is $1 - \epsilon$ for Dana's updating problem and 1 for any other updating problem. We noted that this rule is evidentially constant and conservative and has higher expected accuracy than BCOND in Dana's updating problem if accuracy is measured by the Brier score. As it turns out, this result holds for any scoring rule that is concave and strictly proper (proof in the appendix.)

Fact 4. *Let* A *be any accuracy scoring rule which is concave in addition to strictly proper. Then any nontrivial mixture of* BCOND *and* META-COND *is evidentially constant, conservative, and has higher expected accuracy (relative to* A*) than* BCOND *in every updating problem where they do not agree.*

Thus, if BCOND is rational to the extent that it is the most accurate followable updating rule, then we must rule out all mixtures of BCOND and META-COND as unfollowable. And there are good reasons to think that this cannot be done. Consider the following plausible platitude about followability.

> **Locality**. An updating rule is followable just in case, in every updating problem, an (ideal) agent facing that updating problem can follow the rule.

To deny locality is to think that, either some updating rule is followable even though there is an updating problem where no ideal agents can follow the rule, or some updating rule is unfollowable even though it can be followed by an ideal agent in every updating problem. Either sounds close to a contradiction.

Suppose locality is true. The question whether an updating rule is followable is then equivalent to the question whether an (ideal) agent can follow the rule in a given updating problem. Now, suppose an agent can follow an updating rule f in an updating problem. Then, intuitively, the agent should also be able to follow any updating rule that only recommends different credence functions when f does, and whose recommendations deviate from those of f by a small enough margin. For instance, in *Dana*, if the agent can follow BCOND, which tells her to adopt the credence function $(0.25, 0.75, 0)$ if and only if the butter in fact weighs 15 oz., then it seems that the agent can also follow a rule that tells her to adopt the credence function $(0.25+\epsilon, 0.75-\epsilon, 0)$ if and only if the butter in fact weighs 15 oz., as long as $\epsilon > 0$ is sufficiently small. It is unclear what the obstacle could be. After all, on pain of simply stipulating that BCOND is the only followable rule, it can't be that BCOND is followable, yet any slight deviation from it is off the table.

Here is one way to make the above thought more precise. Fix an updating rule f and an updating problem (S, π, E). Let δ be a positive real number. We say an updating rule g is **δ-close to** f in (S, π, E) if, at every state s, the Euclidean distance between the credence functions recommended by f and g at s in this updating problem is smaller than δ.[16] We say g is **no more sensitive than** f in (S, π, E) if for any s and s', g recommends different credence functions at s and s' only if f recommends different credence functions at s and s'.

> **Margin**. If an updating rule f is followable in an updating problem (S, π, E), then there is some margin $\delta > 0$ (which possibly depends on f and (S, π, E)) such that any updating rule g that is no more sensitive than f and is δ-close to f in (S, π, E) is also followable in (S, π, E).

Margin is quite plausible. However, it follows from locality and margin that a version of the updating rule h that we constructed above is followable if BCOND is followable (for ϵ smaller than $\delta/2$). And yet, as we saw, that rule has higher expected accuracy than BCOND in Dana's updating problem (and the same expected accuracy in every other updating problem). More generally, given locality and margin, fact 4 entails that, if BCOND is followable, then relative to any strictly proper and concave scoring rule, there are (uncountably many) followable updating rules that have higher expected accuracy than BCOND in every opaque updating problem (and the same expected accuracy as BCOND in every transparent updating problem).

There is one more reason why externalist Bayesians should be skeptical of the idea that BCOND is rational because it is the most accurate followable updating rule (see, e.g., Isaacs and Russell 2023). So far we have assumed that BCOND is followable in all updating problems; the challenge is to say why other updating rules—META-COND, the dog-

16. This argument goes through for other natural measures of distance between credence functions as well, including any p-norm and Bregmann divergence.

matic rule, or the mixture rules—are not similarly followable. However, one might think that this assumption itself is dubious. Take *Dana* again. According to BCOND, if the butter weighs 15 oz., then she should conditionalize on the proposition that the butter weighs 15 oz. or 16 oz., even though her evidence doesn't entail that *this is what* BCOND *tells her to do*. For it is compatible with what she learns at that state that the butter weighs 16 oz., and at that state, Dana shouldn't conditionalize on the proposition that the butter weighs 15 oz. or 16 oz.; she should leave open the possibility that the butter weighs 17 oz. But how could Dana update by conditionalization, if she doesn't know which proposition to conditionalize on?

To make the argument slightly more rigorous: Say an updating rule f is **luminous** in an updating problem (S, π, E) if for every state s, if $f_{(S,\pi,E)}(s) = \rho$, then $E(s)$ entails the proposition *f recommends adopting the credence function* ρ, that is, $E(s) \subseteq \{t \in S : f_{(S,\pi,E)}(t) = \rho\}$. Intuitively, if we take the agent's knowledge to consist of all and only propositions entailed by her evidence (Williamson 2000), then an updating rule is luminous just in case, necessarily, the agent knows which credence function the rule tells her to adopt. The dogmatic rule is always luminous in this sense. BCOND, on the other hand, is only luminous in transparent updating problems. This is because, although BCOND does not require the agent to be certain of what she learns, it does require the agent to be *sensitive* to what she learns: if the agent learns different propositions at s and s', then she should conditionalize on different propositions at s and s'. Since what the agent learns in opaque updating problems may not entail facts about what she learns, it follows that BCOND is not luminous in such cases. In fact, it is easy to check that the only luminous updating rules in *Olivia* and *Dana* are those rules which recommend the same posterior credences at every state, like the dogmatic rule. Therefore, if luminosity is necessary for followability, then it is not the case that BCOND is followable in all updating problems, assuming evidence externalism.

To recap, we started with the observation that BCOND is not the most accurate updating rule; the truth rule is. So why is BCOND the rule of rational credal revision? We have considered one attractive proposal, according to which the rational updating rule is the most accurate updating rule that an agent can follow. The truth rule is not the rational updating rule because it is unfollowable. However, we think this proposal isn't a satisfactory answer to the explanatory challenge for the externalist Bayesian, for the following four reasons:

1. If evidential constancy is sufficient for followability, then META-COND is followable and has higher expected accuracy than BCOND does in every updating problem (unless they agree).
2. There are opaque updating problems where retaining one's prior has higher expected accuracy than updating by BCOND.
3. For any criteria of followability that satisfies locality and margin, there are updating rules that satisfy the criteria and have higher expected accuracy than BCOND in every updating problem (unless they agree).
4. If luminosity is necessary for followability, then BCOND is not followable in any opaque updating problem.

Where does this leave us? The most common reaction amongst externalists is to give up on the explanatory project altogether. Some still accept BCOND as the rule of rational credal revision, but deny the request of explaining its normative status in terms of accuracy. Perhaps it is a primitive epistemic norm that rational agents should update by Bayesian conditionalization; perhaps the norm is derivative from the more basic norm that rational agents should proportion their credences to evidential probabilities or objective chance, both of which evolve by conditionalization. If doing so does not maximize expected accuracy, then so much the worse for the idea that rationality requires maximizing expected accuracy.

More recently, some externalists have gone the other direction of rejecting BCOND as the rule of rational credal revision, on the basis that it doesn't maximize expected accuracy in opaque updating problems. However, nor should rational agents update by META-COND or the truth rule. Instead, the rational updating rule for a given agent is the rule such that, if the agent were to *try* to follow the rule, her credences would be most accurate on expectation. Most agents would fail badly if they were to try to follow the truth rule or META-COND, which is why neither is rational. But the same could be true of BCOND, and in such cases, BCOND is *not* the rule that the agent should try to follow either (Gallow 2021; Isaacs and Russell 2023; Schultheis forthcoming). The explanatory challenge is effectively dissolved, as there is nothing to be explained in the first place.

It is beyond the scope of this article to examine either view in detail. Instead, we'll settle for a more modest aim by arguing that these are not the only options available to the externalist. The externalist *can* have it both ways: they can be Bayesians *and* think that BCOND is rational because it is truth-conducive (in a sense). Our proposal has two key ingredients. The first is a strengthening of evidential constancy, which we call *global evidential constancy*. Unlike evidential constancy, we propose global evidential constancy, not as a followability constraint, but as a normative constraint on updating rules; an updating rule that violates global evidential constancy is irrational, whether or not it is followable. Notably, both META-COND and the truth rule violate this constraint, and are irrational to that extent. The second ingredient of our proposal is a new way of evaluating the accuracy of an updating rule. The rational updating rule, we argue, is the most accurate updating rule *in general*. We propose a precise account for what it means for an updating rule to be most accurate in general, and show that, of all updating rules that satisfy global evidential constancy, BCOND is the unique updating rule that satisfies our criterion. Bayes is back!

4. Global Evidential Constancy

Let us start with global evidential constancy. Formally, it is the conjunction of two constraints: evidential constancy, and a new constraint which we will call *noncontrastiveness*. We think both are necessary conditions that an updating rule needs to satisfy in order to be rational. This is so, independently of whether or not the updating rule is followable. We'll defend this claim in section 4.3. But first, let us introduce and motivate the new constraint of noncontrastiveness.

4.1. Noncontrastiveness

Let's go back to *Die*. Recall that Xani and Yota are getting information about the same four-sided die from two different sensors, and both of them know this. Xani will learn from her sensor whether the die lands odd or even; Yota will learn from her sensor whether the die lands on one, three, or even. Both of them know that it is a fair die, so they share the same uniform prior about how the die will land ex ante.

Suppose the die lands on two, and so Xani and Yota learn from their respective sensors that the die lands even. Question: How should their posterior credences over the four possible outcomes of the die toss compare with one another?

Intuitively, the answer is they should be the same. For one thing, both of their updating problems are transparent, so it follows from fact 1 that updating by conditionalization maximizes expected accuracy in their updating problems. Given that they start out with the same prior distribution over the outcomes of the die toss, conditionalizing on the proposition that the die lands even will bring them to the same posterior distribution over the possible outcomes of the die toss.

However, we think this verdict is plausible even without appealing to accuracy-based considerations. Prima facie, if the agents start out with the same prior over states S and learn the same true strongest proposition about S at s, then they should end up with the same posterior distribution over S at s. The two agents may have very different background beliefs, including beliefs about how they come to learn the relevant proposition. Nevertheless, the thought goes, if that difference doesn't manifest in their priors over S, then it shouldn't manifest in their posteriors over S either. In a slogan: same prior over S, same evidence about S, same posterior over S.

As mentioned, we will assess this intuitive thought in more detail later in the section. But first, let us try to make the thought more precise, and see where it leads. Say an updating rule f is **noncontrastive** if, for any two updating problems (S, π, E) and (S, π, E') involving the same prior space, for any state $s \in S$, if the strongest proposition learned at s is the same, then the rule recommends the same posterior over S at s:[17]

$$E(s) = E'(s) \Rightarrow f_{(S,\pi,E)}(s) = f_{(S,\pi,E')}(s).$$

Whereas evidential constancy requires that a rule recommend the same posteriors over S *at different states* in a given learning situation when the strongest proposition learned is the same at those states, noncontrastiveness requires that a rule recommend the same posteriors over S *in different learning situations* at a given state when the strongest proposition learned is the same at that state; it doesn't matter what one could have learned instead, as long as one learns the same thing. One can check that in addition to satisfying evidential constancy, BCOND is noncontrastive. In particular, let (S, π, E) and (S, π, E') be updating problems and suppose that $E(s) = E'(s)$. Then BCOND recommends the same posterior over S at s:

17. We borrowed this term from J. Dmitri Gallow (2019a), who makes a similar observation about BCOND. Noncontrastiveness is also related to the property of "stability" discussed in Briggs 2010: 27, and symmetry-invariance in the sense of Hughes and Van Fraassen 1984.

$$b_{(S,\pi,E)}(s) = \pi\left(.\middle|E(s)\right) = \pi\left(.\middle|E'(s)\right) = b_{(S,\pi,E')}(s).$$

A quick clarification about the scope of noncontrastiveness: it is important, in the statement of the constraint, that (S, π, E) and (S, π, E') are updating problems involving the same prior over states and the same evidence proposition learned at s. To illustrate the importance of this qualification, consider the following example:[18]

> *Litmus Test.* Albert and Bob are performing a litmus test on a fluid. Prior to the test, they're both 50/50 about whether the fluid is acidic or basic, but they disagree about how to interpret the litmus test. Albert is certain that the litmus paper turns red if and only if liquids are acidic, whereas Bob is certain that the litmus paper turns red if and only if liquids are basic. They then both learn the color of the litmus paper.

Intuitively, Albert and Bob should have different posteriors about the acidity of the fluid after the litmus test, even though they share the same prior about the acidity of the fluid ex ante. Is this a counterexample to noncontrastiveness?

It is not, for the following reason. First, although Albert and Bob share the same prior about the acidity of the fluid, they do not share the same prior about the acidity of the fluid and the color of the litmus test (for instance, while Bob has no confidence in the conjunction *the fluid is acidic and the litmus paper will turn red*, Albert does). Thus, if we model their updating problems using states that specify both pieces of information, then Albert and Bob will not have the same prior over states, and so noncontrastiveness will not apply. On the other hand, suppose we model Albert's and Bob's updating problems using a coarser model $S = \{a, b\}$, where a is the proposition that the fluid is acidic and b is the proposition that b is basic. In this case, although Albert and Bob do share the same prior over S, they cannot be plausibly thought of as learning the same true subset of S (as their strongest proposition learned) at either state. To see why, suppose the fluid is acidic and the litmus paper turns red. What is the strongest proposition about the acidicity of the fluid that Albert and Bob learn at a? The only two propositions representable as subsets of S that are true at a are *the fluid is acidic* ($\{a\}$) and *the fluid is either acidic or basic* ($\{a, b\}$). Presumably, for Albert the answer is not $\{a, b\}$, since intuitively he has learned *something* nontrivial about the acidity of the fluid. And for Bob it is not $\{a\}$, since if the fluid is acidic and the litmus turns

18. We thank an anonymous referee for suggesting this example, which we borrowed almost verbatim.

red, then Bob doesn't learn that the fluid is acidic; he learns that the litmus paper turns red and is disposed to infer that it is basic! So either this coarse-grained model is not adequate to represent Albert or Bob's learning situations in the first place,[19] or Albert and Bob learn different strongest propositions (represented as subsets of S) on the model at a. In either case, noncontrastiveness does not apply. The same argument extends to b, if the fluid is basic and the paper turns blue, and mutatis mutandis (with Albert and Bob switched) if the litmus test behaves as Bob believes. In conclusion, depending on the model we adopt, either Albert or Bob do not have the same prior over states, or learn different evidence propositions at each state; in either case, noncontrastiveness does not apply.[20,21]

19. We are inclined toward this diagnosis of the case, but the adequacy of such formal models is a subtle issue, especially in an externalist setting. See notes 21 and 22 for related discussion.

20. Strictly speaking this is not true on all models; letting r denote the proposition that the litmus paper turns red, one *could* use the state space $S' = \{r, \neg r\}$ to model Albert and Bob's updating problems. In fact, relative to this state space, Albert and Bob's updating problems would be the same: formally, (S', π', E'), where $\pi'(r) = \pi'(\neg r) = 1/2$ and $E'(r) = \{r\}; E'(\neg r) = \{\neg r\}$. Here, noncontrastiveness does apply. But intuitively, Albert and Bob *should* have the same posterior about the color of the litmus paper, so this is not a counterexample to noncontrastiveness either.

21. A similar point extends to some other putative counterexamples to noncontrastiveness. Some theorists, in explaining the Monty Hall problem and its variants (such as the three prisoner paradox), sound close to rejecting noncontrastiveness. For example, Judea Pearl (1988) writes "The lesson [of these cases] is that we cannot assess the impact of new information by considering only propositions implied by the information; we must also consider what information *could have* been reported." Here we should distinguish between the standard transparent versions of these cases and other versions. On the standard version of the Monty Hall problem, when the agent is playing the game, they learn not only, e.g., *there is a goat behind the door*, but also, e.g., *Monty Hall told me there is a goat behind the door*, and the latter may not be equivalent to any proposition solely about what is behind the doors. So the agent's total incremental evidence is not representable as a subset of the state space where the states only specify the possible positions of the car. The standard move is to notice that once we suitably fine-grain S to represent the agent's total incremental evidence, we find that BCOND (which satisfies noncontrastiveness) gives the right answer to the case. Thus noncontrastiveness is still satisfied with respect to this fine-grained S. We interpret Pearl's remark as pointing to the fact that to properly model the epistemic implications of the agent's total incremental evidence, we must take into account the agent's prior over the fine-grained S, including her beliefs about Monty Hall's protocol. This is compatible with noncontrastiveness. On the other hand, we can also imagine opaque versions of the Monty Hall problem where one learns nothing more than *there is a goat behind the door* (Bronfman 2014). In that case, it is less obvious that the standard solution still applies. More generally, in such opaque cases, it will be more controversial whether BCOND gives the right verdicts and thus whether non-

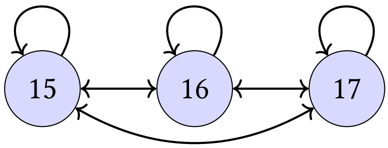

Figure 6. A simple model of Noah's learning situation.

We will be interested in applying noncontrastiveness to the cookie case. To do this, let us introduce a counterpart to Dana, who has the same prior over states as Dana, but is in a different learning situation (see also fig. 6):

> *Noah and Dana.* Noah is making cookies together with Dana. Like Dana, Noah knows that the butter that Dana has left weighs between 15 and 17 oz., and he shares Dana's confidence judgment that it is 0.6 likely that the butter weighs 16 oz., and equally likely that it weighs 15 oz. or 17 oz. However, unlike Dana, Noah is much less perceptive. No matter how much the butter weighs (15 oz., 16 oz., or 17 oz.), he won't be able to learn anything new just by weighing the butter by hand, and he knows this.

Like Xani and Yota, Noah and Dana are in updating problems (S, π, E) and (S, π, E') involving the same prior space but different learning situations.[22] Note that if the butter in fact weighs 16 oz., then Noah and Dana will learn the same proposition, namely, the tautology. So noncontrastiveness demands that Noah and Dana plan to update their credence over the weights the same way at that state. Note that BCOND sat-

contrastiveness holds. On that point, we refer the reader to our arguments below for why we think noncontrastiveness is a plausible normative constraint on updating rules, and BCOND is the rational updating rule. At this stage, our point is just that noncontrastiveness is compatible with the received solution to the standard version of the Monty Hall problem. Thanks to an anonymous referee for helpful insights on this point.

22. It is true that Noah and Dana have different background beliefs; for example, Dana is certain that she will learn that the butter weighs below 17 oz. if and only if the butter weighs 15 oz., while Noah is certain of no such biconditionals about himself. Nevertheless, by stipulation, they share the same prior distribution over the same set of possible states (i.e., $S = \{15\text{oz.}, 16\text{oz.}, 17\text{oz.}\}$), and their respective learning situations can be represented by functions from S to its subsets. One could dispute these stipulations; one might think our models of Noah's and Dana's updating problems are overly simplified. However, the point is that, assuming that the way we model their respective updating problems are formally adequate, nonconstrastiveness says that they should adopt the same posterior at the state where the butter weighs 16 oz.

isfies this constraint; in particular, BCOND recommends that both Noah and Dana retain their prior when they don't learn anything new.

Notably, however, META-COND violates noncontrastiveness in this case: it recommends Noah retain his prior, while Dana become certain that the butter weighs 16 oz. Formally, letting (S, π) denote Dana and Noah's prior space, E denote Dana's learning situation, and E' denote Noah's learning situation, META-COND recommends the following, when s is the 16 oz. state:

$$m_{(S,\pi,E)}(s) = \pi\left(.\middle|E = E(s)\right) = \pi\left(.\middle|\{s\}\right)$$
$$\neq \pi(.|S) = \pi\left(.\middle|E' = E'(s)\right) = m_{(S,\pi,E')}(s)$$

even though $E(s) = E'(s) = S$. This is a violation of noncontrastiveness. In fact, this result is not limited to META-COND: any nontrivial mixture of META-COND and BCOND will violate noncontrastiveness for a similar reason. For instance, consider the mixture rule h described in section 3.3. According to h, if the butter weighs 16 oz., then Noah should be 0.6 confident that the butter weighs 16 oz., whereas Dana should be slightly more than 0.6 confident that the butter weighs 16 oz. ($0.6 + \frac{2}{5}\epsilon$ to be precise). The difference is small, but it is a difference nonetheless, which is at odds with noncontrastiveness.

This is promising. As we argued in the previous section, one challenge for the externalist Bayesian is to explain why the rational updating rule is not some mixture of BCOND and META-COND, even though such rules always have higher expected accuracy than BCOND whenever they disagree. In particular, it is unclear that the externalist Bayesian can rule out such mixture rules on the grounds that they are unfollowable, given the plausibility of locality and margin. The above observation suggests an answer to this challenge: rational updating rules are noncontrastive; the mixture rules are irrational, because and to the extent that they violate noncontrastiveness.

On the other hand, it's worth noting that while the constraint of noncontrastiveness rules out META-COND and the mixture rules, it does not rule out the truth rule. For at any state s, the truth rule does recommend adopting the same credence function over states at s, namely, the omniscient credence function that assigns probability 1 to s. So noncontrastiveness alone is not enough to eliminate all updating rules that always have equal or higher expected accuracy than BCOND; another constraint would be needed to rule out the truth rule as well.

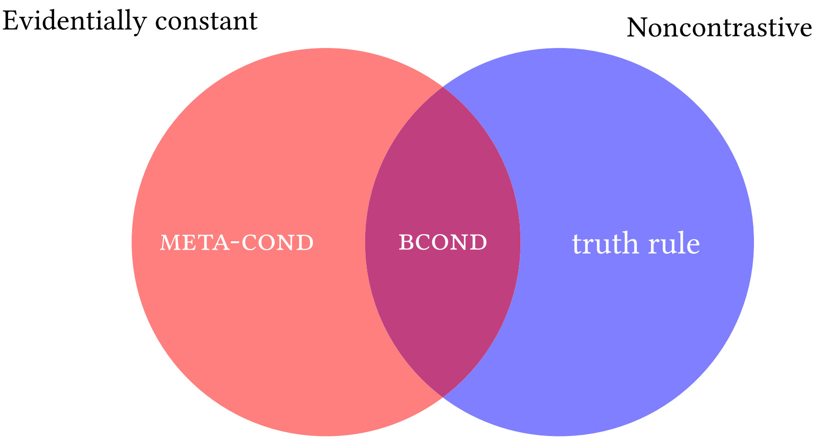

Figure 7. Venn diagram illustrating the relationship between evidentially constancy and noncontrastiveness. While META-COND is evidentially constant but not noncontrastive, the truth rule is noncontrastive but not evidentially constant. BCOND satisfies both.

The constraint we propose to add to nonconstrastiveness is one that we've seen before: evidential constancy. However, in the present context, the constraint is interpreted as a normative constraint, not a followability constraint. That is, on our proposal, the truth rule is not a rational updating rule, not because it is unfollowable, but because it is irrational; it recommends the agent to adopt different credences at two different states, in a given updating problem, despite learning the same proposition at those two states. We think doing so is irrational, whether or not it is possible.

Before proceeding, it's worth highlighting one interesting aspect of our proposal. When faced with fact 2, the intuitive response is that META-COND is very much like the truth rule, and should be ruled out on the same or similar grounds (e.g., on grounds of unfollowability). This is not our proposal. Although we think both rules are irrational and thereby ruled out as competitors to BCOND, they are ruled out for quite different reasons. The truth rule is irrational because it violates evidential constancy, whereas META-COND is irrational because it is contrastive. BCOND, on the other hand, passes the test by being both evidentially constant and noncontrastive (see fig. 7).

As promised, we'll say more about the motivations for imposing evidential constancy and noncontrastiveness as normative constraints in section 4.3. But first, let us give an explicit definition of *global evidential constancy*. As aforementioned, a rule f is globally evidentially constant just in case it satisfies evidential constancy and noncontrastiveness. More explicitly: an updating rule is **globally evidentially constant** if, for any pair of updating problems (S, π, E) and (S, π, E') involving the same set of states and prior over states, for any $s, s' \in S$,[23]

$$E(s) = E'(s') \Rightarrow f_{(S,\pi,E)}(s) = f_{(S,\pi,E')}(s').$$

In a slogan: same prior over S, same evidence about S, same posterior over S regardless of s.

While the fact that global evidential constancy rules out both the truth rule and mixture rules is promising, it's worth noting that imposing global evidential constancy is not going to be a cure-all to the problems raised in section 3. For example, the dogmatic rule, which recommends always retaining one's prior, is globally evidentially constant. Yet we saw that the dogmatic rule has higher expected accuracy than BCOND in *Dana*. So more needs to be said about why BCOND is rational despite this fact. We will have an answer to this question; but we are getting ahead of ourselves. First, let us focus on fleshing out global evidential constancy as a constraint on updating rules. We'll start by giving a helpful alternate characterization of this constraint, and then discuss its status as a normative requirement on updating rules.

4.2. An Alternate Characterization of Global Evidential Constancy

As we defined an updating rule, it takes three inputs—a prior space (S, π), a learning situation E, and a state s in S—and outputs a probability distribution over S, interpreted as the credences over S that an agent facing the updating problem (S, π, E) should plan to adopt at s. However, one might have thought that this definition is too complicated. Instead, an updating rule only needs to take two inputs—a prior space (S, π) and an evidence proposition $p \subseteq S$. It then outputs a probability distribution over S, interpreted as the posterior credences over S that an agent should plan to have if she has prior π over S and learns p and nothing

23. That this condition entails noncontrastiveness and evidential constancy is straightforward. For the converse direction, suppose an updating rule f is evidentially constant and noncontrastive. Let (S, π, E) and (S, π, E') be two updating problems. Suppose $E(s) = E'(s') = p$. Consider the factive learning situation F given by, for each state $t \in S$, $F(t) = p$ if $t \in p$ and $F(t) = S \setminus p$ if $t \notin p$. By factivity, $s, s' \in p$. So $F(s) = F(s') = p = E(s) = E'(s')$. By noncontrastiveness, $f_{(S,\pi,F)}(s) = f_{(S,\pi,E)}(s)$ and $f_{(S,\pi,F)}(s') = f_{(S,\pi,E)}(s')$. By evidential constancy, $f_{(S,\pi,F)}(s) = f_{(S,\pi,F)}(s')$. Therefore, $f_{(S,\pi,E)}(s) = f_{(S,\pi,F)}(s) = f_{(S,\pi,F)}(s') = f_{(S,\pi,E')}(s')$. So f is globally evidentially constant.

else. Indeed, we can think of BCOND this way: if an agent has prior π over S and the strongest proposition she learns is p, which corresponds to a subset of S, then BCOND tells her to revise her credence distribution over S to $\pi(.|p)$. In other words, formally, b is representable by the map b^* that maps a prior space (S, π) and a proposition $p \subseteq S$ to $\pi(.|p)$:

$$b^*_{(S,\pi)}(p) = \pi(.|p).$$

Another updating rule which is globally evidentially constant and can be represented this way is the dogmatic rule. Formally, the rule can be represented by the map $d_{(S,\pi,E)}(s) = \pi$, which in turn corresponds to the map

$$d^*_{(S,\pi)}(p) = \pi.$$

On the other hand, not every updating rule in our sense can be represented by a map from priors over states and propositions to posteriors. In particular, neither the truth rule nor META-COND (nor mixtures of BCOND and META-COND) can be represented this way. To see that META-COND is not representable by a map of this form: let π be Noah's and Dana's common prior over $S = \{15, 16, 17\}$ and let $p = S$. As noted above, the output of META-COND for π and p is underdetermined. If an agent with prior π is in Noah's learning situation and learns p, then META-COND says she should retain her prior and therefore assign a credence of 0.6 to the 16 oz. state; if an agent is in Dana's learning situation and learns p, then META-COND says she should not retain her prior but assign probability 1 to the 16 oz. state.

This is not a coincidence. As it turns out, an updating rule in our sense is representable by such a map if and only if it satisfies global evidential constancy (proof in the appendix).

Fact 5. *An updating rule f is globally evidentially constant if and only if there is a map f^* that takes as input a prior space (S, π) and proposition $p \subseteq S$ and outputs a posterior probability distribution $f^*_{(S,\pi)}(p)$ over S, such that for any updating problem (S, π, E),*

$$f_{(S,\pi,E)}(s) = f^*_{(S,\pi)}\big(E(s)\big).$$

Given this characterization result, in what follows, if a rule f is globally evidentially constant, then we'll abuse notation and use "f" to also denote the map of fewer inputs f^* that represents f. It will often be more convenient to represent globally evidentially constant rules in this alternate way.

4.3. Global Evidential Constancy as a Normative Constraint

As noted above, we take global evidential constancy to be a normative constraint on rational updating rules. To motivate this, let's start with evidential constancy. Recall that an updating rule is evidentially constant just in case it does not recommend different credence functions at states where the agent learns the same proposition. We previously introduced evidential constancy as a popular criterion for followability. But here we are proposing evidential constancy as a normative constraint: an updating rule that violates evidential constancy is irrational, whether or not it is followable.[24]

Why should rational updating rules be evidentially constant? The basic thought is: an agent should adopt the same credence function in response to learning the same proposition: same evidence, same credence. Note that this principle is neutral about one's account of when the agent learns the same proposition, and is therefore compatible with a wide range of views about epistemic rationality. For instance, one might think an agent learns the same proposition at s and s' just in case she knows the same thing at s and s' (Williamson 2000). Then evidential constancy says that an updating rule is rational only if it recommends the same credence function for an agent at any two states where she knows the same thing. Alternatively, one could think that an agent learns the same proposition at s and s' if she is disposed to guess the same state to be true at s and s' (Isaacs and Russell 2023). Then evidential constancy says an updating rule is rational only if it recommends the same credence function for an agent whenever she is disposed to guess the same state to be true at s and s'. So evidential constancy is quite weak, and seems plausible to that extent.

24. This approach is much less amenable to the *accuracy-first* project (Pettigrew 2016), according to which all credal norms are reducible to decision-theoretic norms about promoting accuracy. However, as discussed in section 3, we are skeptical that BCOND can be the vindicated on the grounds that it has the highest expected accuracy out of all followable rules in all updating problems, or on the grounds that trying to follow BCOND has the highest expected accuracy in all updating problems. So for the externalist Bayesian who wants to preserve a substantive connection between accuracy and the epistemic rationality of updating rules, looking beyond followability considerations will be key.

Nevertheless, one might still find evidential constancy too strong, on the following grounds. In our framework, an updating rule is evidentially constant just in case, for every proposition that an agent might learn in an updating problem, there is one unique credence over states that the rule recommends them to adopt in response to learning that proposition.[25] Many find it plausible that there is no unique credence function that an agent should adopt in the state of total ignorance. Presumably, you know absolutely nothing about Xani. How confident are you that she is vegetarian? 0.5? 0.6? Intuitively, both are permissible; neither is irrational. Suppose this is right. Then, the thought goes, there may also be multiple permissible answers to this question—permissible for you, given your prior beliefs—after you acquire some information about Xani (e.g., that she likes the color blue).

Our reply to this objection is quite flat-footed: on its own, evidential constancy is compatible with the idea that there are multiple rationally permissible credence functions that an agent may adopt in response to learning a proposition in an updating problem. The two are incompatible only under the additional assumption that at every state, there is one unique credence function that the agent should adopt at that state. Formally, this corresponds to our stipulation that an updating rule associates each updating problem with a function that maps states to credence functions over states (Pettigrew 2020 calls this assumption *deterministic updating*). One could relax this assumption and adopt a more general notion of updating rules, which associates each updating problem with a function from states to *sets of credence functions*, interpreted as the set of credence functions that it is rationally permissible for the agent to adopt at that state. The notion of evidential constancy still applies in this more general framework, but it no longer entails that, for every updating problem, there is one unique credence function that an agent should adopt in response to learning a given proposition. (On the other hand, as Pettigrew (2020) points out, in this more general framework, it is unclear that BCOND *is* the unique rule of rational credal revision, even if we assume evidence internalism. In particular, depending on how we calculate the expected accuracy of permissive updating rules, BCOND may or may not be the unique updating rule that maximizes expected accuracy. To this extent, Pettigrew's (2020) results present a different

25. In other words, evidential constancy entails a thesis of *intrapersonal uniqueness* in the sense of Kelly 2013.

explanatory challenge for the Bayesians, which we shall not explore in this article.)

So much for evidential constancy for now. What about noncontrastiveness? Prima facie, this constraint looks quite appealing as well. Intuitively, how one should revise one's credence in response to a learning experience doesn't depend on what other learning experience (one thinks) one could have had instead. Maybe you think that you could have seen an UFO when you looked outside of your window, maybe not. Maybe you think that you could have been reading a paper on the feeding behaviors of red pandas (which could have been a better use of your time), maybe not. Either way, those possibilities seem neither here nor there. What you should *actually* believe depends on your initial evidence or beliefs and what you *actually* see and read, not what you could have seen and read.

There is also a theoretical argument for noncontrastiveness. Let's go back to *Die* once again. Suppose Xani learns from her sensor that the die landed even at t_1. She then learns from Yota at t_2 that her sensor says the same thing. There is no surprise here—she already knows this, since she knows—or, if one prefers, is *absolutely* certain—that both her and Yota's sensors are perfectly reliable. What should be Xani's credences in the outcome of the die toss at t_2?

Intuitively, they should be the same as her credences at t_1. Learning again that the die landed even (and nothing else) shouldn't make her more or less confident that the die landed even, or more/less confident that the die landed on a 2 instead of a 4. After all, by stipulation, Xani already knows at t_1 that the die landed even. What could justify a credal shift between t_1 and t_2, if she learns nothing in the interim that she doesn't already know?[26]

As before, let us try to make this thought more precise. Say an updating rule is **idempotent** if it recommends the agent to not plan to update differently when she learns the same proposition twice at a given state as compared to when she learns it once. Formally, f is idempotent

26. Note that this is compatible with the thought that, *if* Xani does have uncertainties about the reliability of her or Yota's sensors, then she may become more confident that the die landed even after learning that Yota's sensor records the same outcome. In particular, in such cases, Xani won't be learning the same proposition twice. Instead, she learns at t_1 the proposition *that her sensor records that the die landed even*, and at t_2 the proposition *that Yota's sensor records that the die landed even*. These two propositions do not tell her the same thing about the die toss, relative to her background knowledge, given her prior uncertainties about the reliability of the sensors.

just in case, for any two updating problems (S, π, E) and (S, π, E') with the same prior space, and any $s \in S$,

$$E(s) = E'(s) \Rightarrow f_{(S,\pi,E)}(s) = f_{(S,\pi,E\wedge E')}(s),$$

where $(E\wedge E')(s) = E(s)\cap E'(s)$.[27] We submit that idempotence is a plausible rational constraint on updating rules. Learning the same proposition twice should have the same effect as learning it once. This constraint seems particularly compelling if we further stipulate that the agent is always certain of what she learns. For then violations of idempotence would involve revising one's credences upon learning a proposition that one is already certain to be true.[28]

The crucial observation is that, for updating rules, idempotence is equivalent to noncontrastiveness (proof in the appendix).

Fact 6. *An updating rule is idempotent if and only if it is noncontrastive.*

To get a feel for this fact, let us consider META-COND. As noted above, META-COND is contrastive, so it is not idempotent. For a concrete illustration, suppose Dana's right hand is sensitive in the way depicted in figure 3, but her left hand is more sensitive: she can tell, by weighing the butter using her left hand, whether or not it weighs 17 oz. (but nothing

27. Note that this intersection is always nonempty given factivity. There is also a different formal regimentation of the idempotence constraint: say an updating rule is **idempotent*** if for any two prior-equivalent updating problems (S, π, E) and (S, π, E'), if $E(s) = E'(s)$, then $f_{(S,\pi,E)}(s) = f_{(S,\pi_s,E')}(s)$, where $\pi_s = f_{(S,\pi,E)}(s)$. It's straightforward to check that the two conditions are equivalent given the following constraint on updating rules:

> **Accumulation**. For any two prior-equivalent updating problems (S, π, E) and (S, π, E'), $f_{(S,\pi_s,E')}(s) = f_{(S,\pi,E\wedge E')}(s)$, where $\pi_s = f_{(S,\pi,E)}(s)$.

Intuitively, accumulation says that learning two (possibly the same) propositions in sequence from two learning experiences has the same rational effect on one's credence as learning their conjunction in one single learning experience. We think accumulation is quite plausible, so we will adopt idempotence in virtue of its formal simplicity, even though idempotence* is perhaps a more faithful formal regimentation of the intuitive idea. Incidentally, META-COND violates both accumulation and idempotence* (see note 30).

28. Admittedly, idempotence may not hold for nonideal agents with cognitive bounds, or in learning situations where there is no strongest proposition that summarizes the agent's learning experience. But neither falls under the scope of BCOND. Instead, it is commonly assumed that BCOND only applies to idealized agents with unlimited cognitive resources and whose learning experiences consist in learning some propositions. It is more difficult to imagine how idempotence could fail in such cases.

more specific). Suppose Dana first weighs the butter using her left hand at t_1, and then weighs it using her right hand at t_2. By stipulation, if the butter weighs 15 oz., then at both times, the strongest proposition that she learns is that the butter weighs 15 oz. or 16 oz. So idempotence says that, at that state, the rational updating rule should tell her to adopt the same credence function at t_1 and t_2.[29] This is not the recommendation of META-COND. Instead, META-COND recommends that Dana be one-fourth confident that the butter weighs 15 oz. at t_1, but certain that it weighs 15 oz. at t_2.[30]

Now one might object: despite initial appearances, it is actually rational for Dana to change her credences between t_1 and t_2 (though perhaps not in the particular manner recommended by META-COND), because there is a crucial difference between her two learning experiences. At t_1, she learns that the butter weighs 15 oz. or 16 oz., and also comes to know that *that* is the strongest proposition that she has learned. At t_2, on the other hand, although she learns the same evidence proposition, she doesn't also come to know that *that* is what she has learned. In other words, even though Dana actually learns the same proposition twice, she doesn't know that herself, and, the thought goes, this higher-order ignorance about whether she has learned the same proposition could potentially justify a credal shift between t_1 and t_2.

29. Strictly speaking, Dana also learns at t_2 that she is weighing the butter using her right hand, and given her prior knowledge about the differential sensitivity of her two hands, she also knows that she learns that the butter weighs 15 oz. or 16 oz. this time if and only if the butter weighs 15 oz. However—and this is crucial—our model assumes that the fact that she is measuring the butter using her right hand does not give her additional information about the weight of the butter itself (just as our model of Xani assumes that the additional information that Xani learns at t_2, namely, that she is learning something from Yota's sensor, does not give her additional information about the die toss). In particular, Dana cannot infer from the biconditional *the butter weighs 15 oz. if and only if I learn that it weighs 15 oz. or 16 oz.*, a biconditional that she knows based on the fact that she is measuring the butter using her right hand, to the conclusion that the butter weighs 15 oz., because although she does learn that the butter weighs 15 oz. or 16 oz., she does not learn that *that* is what she learns.

30. On the other hand, suppose Dana first weighs the butter using her right hand at t_1, and then weighs it using her left hand at t_2. Then META-COND says that, if the butter weighs 15 oz., then Dana should be certain that the butter weighs 15 oz. at t_1 and so retain that certainty at t_2. On the other hand, if she just weighs the butter using her more sensitive left hand, then she should only be one-fourth confident that the butter weighs 15 oz. This illustrates how META-COND violates accumulation and idempotence* (note 27).

This objection is related to a more general consideration against imposing idempotence (equivalently, noncontrastiveness) as a rational constraint on updating rules. Recently, some epistemologists have argued that rational credal revision should be sensitive to one's higher-order evidence about one's first-order evidence. For instance, some think that one should be certain of p if and only if one has (true) higher-order evidence that p is part of one's evidence (Dorst 2019; Dorst et al. 2021). This principle is incompatible with idempotence, as we saw in the previous variation of *Dana*. Idempotence says that, if the butter weighs 15 oz., Dana should adopt the same credence function at t_1 and t_2. In particular, she should be similarly certain (or uncertain) about whether the butter weighs 15 oz. or 16 oz. at those two times, even though she has the higher-order evidence that this disjunction is part of her evidence at t_1, but not at t_2.

One thing to say about this objection is that the status of level-bridging principles like the one mentioned above is quite controversial. Some authors, such as Williamson (2011, 2014), reject such principles, on the grounds that they are violated in cases like *Olivia* and *Dana* (though Williamson's argument assumes that it is rational to update by BCOND in such cases). Moreover, there are weaker level-bridging principles that are compatible with idempotence. For example, Adam Elga (2013) proposes that the correct bridging principle between first- and higher-order uncertainties is the principle of new rational reflection. In our framework, this principle corresponds to the constraint that, if an updating rule f is rational, then for any updating problem (S, π, E) and probability function $\rho \in \Delta(S)$, one's prior conditional probability—conditional on the proposition that ρ is the rational probability function to adopt according to f—should equal ρ conditioning on that same proposition. This principle is compatible with the requirement that rational updating rules are idempotent, since it is satisfied by BCOND, which as we saw is idempotent.

Another thing to note is that, as we defined it, updating rules are first and foremost rules about which *updating plan* an agent should adopt prior to their learning experience. They impose constraints on which credence function the agent should adopt after their learning experience only if we assume that rational agents should be diachronically continent (Paul 2014; Pettigrew 2016), in the sense that they should revise their credences in accord with the plans that they adopt ex ante. However, this assumption seems dubious precisely in cases like *Dana*. As noted before, in *Dana*, the only updating rule that is luminous is a rule that recommends she adopt one particular credence function no matter what. So, if Dana plans to revise her credences in any nontrivial fashion, after weighing the butter using her less sensitive right hand, then when the time comes, she will not necessarily be in a position to tell *what her plan tells her to do* in her situation. Perhaps she planned to be certain that the butter weighs 15 oz. or 16 oz.; perhaps she planned to give some credence to the possibility that the butter weighs 17 oz. as well. Given this uncertainty, it seems rationally permissible for her to not revise her credences in accord with her own plans.

So the remaining question is: Is it rational for Dana to *plan* to not revise her credences at t_2, after learning (again) that the butter weighs 15 oz. or 16 oz. and nothing more specific? We think it is. By her own lights, at t_2, the strongest proposition that she will learn about the weight of the butter is something that she has already learned at t_1. Granted, she won't know that herself at t_2, and to that extent it might be rational for her to change her credences at t_2. But ex ante, it seems that that uncertainty is neither here nor there. Compare: Suppose you are catching a flight the next day. Right now, you think that it is highly probable that the train going to the airport will be delayed based on its track record. However, you also know that you tend not to be able to think clearly when you do not get enough sleep, and so would not be certain that this high confidence is justified when you wake up early the next morning. How confident should you *plan* to be that the train will be delayed next morning? Intuitively, the answer is: highly confident. Yes, you won't be sure if this is too high when you wake up, and it may very well be rational for you to adjust your confidence then in light of this uncertainty. But that is the problem for your future self next morning; right now, you are rationally confident that the train will be delayed, and should plan to retain that confidence when you wake up.[31]

To recap: our argument for global evidential constancy as a rational constraint on updating rules consists of three steps:

31. Side remark: as many theorists have pointed out, the standard Bayesian framework appears unable to model rational belief revision given higher-order defeat (Weisberg 2009; Levinstein n.d.; Christensen 2010; Ye 2023; White 2009). One potential upshot of the discussion above is that perhaps idempotence is one piece of the puzzle. As we saw, BCOND is idempotent, and idempotence amounts to the requirement that how one should update one's credences in first-order propositions (e.g., the weight of the butter) should be insensitive to one's higher-order evidence about what first-order evidence one has.

1. Global evidential constancy is the conjunction of evidential constancy and noncontrastiveness.
2. Rational updating rules should be evidentially constant because facts about which credence function is rational supervenes on facts about the agent's evidence proposition.
3. Rational updating rules should be idempotent and therefore noncontrastive (fact 6).

While we don't take either 2 or 3 to be conclusive, we think they jointly make a strong case for global evidential constancy, which we shall take for granted from now on.

With global evidential constancy in hand, we can now turn back to the externalist Bayesian's explanatory challenge: Why is BCOND and not the truth rule the rule of rational credal revision? We now have a partial answer to that question: even though the truth rule is the most accurate, it is not rational because it violates global evidential constancy. However, we are not done yet, as we still need an account of why BCOND is the rule of rational credal revision. In particular, to address the explanatory challenge, this account must involve some connection between accuracy and epistemic rationality. One natural conjecture, then, is that BCOND is the rational updating rule because, of all updating rules that satisfy global evidential constancy, it is the most accurate.

As we shall see, this conjecture contains a kernel of truth, but making it precise will take some work.

5. Total Expected Accuracy, Rationality, and Bayes Redux

Let us start with a positive result. With the truth rule, META-COND, and mixtures of META-COND and BCOND off the table, there is no remaining rule that strictly outperforms BCOND in terms of accuracy.

Fact 7. *There is no globally evidentially constant updating rule distinct from BCOND that has equal or higher expected accuracy than BCOND in every updating problem.*

It's worth going through the proof here, just to see global evidential constancy in action. Suppose an updating rule f has higher expected accuracy than BCOND in an updating problem (S, π, E). It follows that f must disagree with BCOND about how the agent should plan to update on a proposition p that she might learn in this problem. Now consider the updating problem (S, π, E') where the agent will only learn the truth of p

and nothing else (i.e., $E'(s) = p$ if $s \in p$ and $\neg p$ if $s \notin p$). Since f satisfies global evidential constancy, it must recommend adopting the same credence distribution over states in response to learning p in this updating problem, which by assumption is different from $\pi(.|p)$. But this updating problem is transparent, and in transparent updating problems, BCOND has the uniquely highest expected accuracy amongst all updating rules that satisfy evidential constancy, which includes f. So BCOND has higher expected accuracy than f in this problem.

While this result looks promising, it is not a resounding success for BCOND. For one thing, the result does not establish that BCOND outperforms its competitors in all updating problems. In fact, we know it doesn't, at least assuming the Brier score or additive logarithmic score for measuring accuracy; recall (fact 3) that the dogmatic rule is globally evidentially constant, and yet has higher expected accuracy than BCOND in *Dana*, according to these two accuracy scores.

One might hope that this is an edge case which relies on a particular choice of scoring rule. But it turns out there is a general impossibility result which does not rely on any particular choice of scoring rule. This result says that any globally evidentially constant rule, including BCOND, has lower expected accuracy than some other globally evidentially constant rule in some updating problem.

Fact 8. *Let* A *be any strictly proper scoring rule. Let* f *be a globally evidentially constant updating rule. Then for any prior space* (S, π)*, there is at least one learning situation* E *relative to* (S, π) *such that* f *has lower expected accuracy in* (S, π, E) *than some other globally evidentially constant rule. Moreover, if the accuracy scoring rule is concave in addition to strictly proper, then there is a learning situation* E *in which, for any* $\delta > 0$, f *has lower expected accuracy in* (S, π, E) *than some globally evidentially constant rule that is* δ*-close to* f*, in the sense of section 3.3.*

The proof is in the appendix. To get a feel for this fact, consider *Dana*. Let (S, π, E) be Dana's updating problem and let A be any strictly proper scoring rule (e.g., the simple logarithmic score). Consider the following updating rule—call it the one-shot rule—which agrees with BCOND for any updating problem that does not have the same prior space as Dana's, and for those that do, it gives the following disjunctive instruction:

- If one learns that the butter weighs 15 oz. or 16 oz., be certain that it weighs 15 oz.; if one learns that the butter weighs 16 oz. or 17 oz., be certain that it weighs 17 oz.; if one learns that the butter weighs 15 oz., 16 oz., or 17 oz. (i.e., nothing new), then be certain that it weighs 16 oz.
- Otherwise, update by Bayesian conditionalization.

Formally, the one-shot rule corresponds to the following function f (we use the tuple (x, y, z) to denote the probability function that assigns probability x to the state that the butter weighs 15 oz., and so on):

$$f_{(S,\pi)}(p) = \begin{cases} (1,0,0) & \text{if } p = \{15,16\} \\ (0,1,0) & \text{if } p = \{15,16,17\} \\ (0,0,1) & \text{if } p = \{16,17\} \\ \pi(.|p) & \text{otherwise.} \end{cases}$$

By construction, the one-shot rule f satisfies global evidential constancy (recall fact 5). Moreover, since following the one-shot rule corresponds to following the truth rule (equivalently, META-COND) in Dana's updating problem, following it has higher expected accuracy for Dana than updating by BCOND according to any strictly proper scoring rule.[32]

It's important to note that, like the dogmatic rule, the one-shot rule is not always more accurate than BCOND (as fact 7 shows). In particular, the one-shot rule has lower expected accuracy than BCOND for Noah. For according to the one-shot rule, after weighing the butter by hand, Noah should become certain that the butter weighs 16 oz. no matter what. By strict propriety, Noah would expect this credence function to be less accurate than his prior. Fact 8 shows that this holds quite generally: for *any* updating rule f that satisfies global evidential constancy (BCOND, the dogmatic rule, the one-shot rule, and so on), and any strictly proper scoring rule A, we can find an updating problem where

32. It's also worth noting that, by construction, this rule also has the property ascribed to BCOND in fact 7: there is no globally evidentially constant rule distinct from f such that has equal or higher expected accuracy than f in every updating problem. To see why, note that such a rule would have to give the same recommendations as f in Dana's updating problem, which means that it would only disagree for evidence propositions p such that f recommends $\pi(.|p)$. But then the reasoning of the proof of fact 7 shows that there would be at least some updating problems, namely, transparent problems where one learns either p or $\neg p$, where f has higher expected accuracy than this alternate rule.

f has lower expected accuracy than some other updating rule f' according to A, though that f' would in turn have lower expected accuracy than some other updating rule f'' (possibly f) in some other updating problem.

Here is a slightly different way of seeing what is going on. By Schoenfield's result, in every updating problem, the updating rule that has the highest expected accuracy of all updating rules that satisfy evidential constancy is META-COND. Now, we have argued that META-COND is irrational because it violates global evidential constancy (and in particular noncontrastiveness). However, for every updating problem, one can find one updating rule that satisfies global evidential constancy and "mimics" META-COND for that updating problem; in fact, there are infinitely many such rules. On the one hand, any of these rules will have the highest expected accuracy in that updating problem (of all updating rules that satisfy global evidential constancy). On the other hand, they are going to do very poorly in some other problems. In particular, if they mimic META-COND for an opaque updating problem, then they will have lower expected accuracy than BCOND does for some transparent updating problems. There is no rule that rules them all.

So we can't identify the most accurate updating rule that satisfies global evidential constancy with the rule that maximizes expected accuracy in every updating problem; there is none. What then? Our proposal, in broadstrokes, is to take the perspective of rule evaluation seriously. Some rules, like *no jaywalking*, may be the best rule overall despite being suboptimal on particular occasions. Similarly, we propose that, of all the updating rules that satisfy global evidential constancy, the "most accurate" rule is the one that is "most accurate in general," in a sense to be made precise in the next subsection. Such a rule may not have the highest expected accuracy in some particular updating problem, but this is fine, since—assuming global evidential constancy—this is true of any rational updating rule anyway.

5.1. Highest Total Expected Accuracy

Consider the one-shot rule from the previous section again. As we noted above, although the one-shot rule is most accurate for Dana's updating problem, it won't be very accurate for Noah's updating problem. More concretely, suppose A is the Brier score, that is,

$$\mathsf{A}(c, s) = -\sum_{p \subseteq S} \left(c(p) - v_s(p)\right)^2.$$

A bit of calculation shows that, for Noah, following the one-shot rule, which amounts to being certain that the butter weighs 16 oz. no matter what, has an expected accuracy of −1.6, whereas following BCOND, which corresponds to retaining his prior, has a higher expected accuracy, of −1.12.

The same is true of Dana if she instead weighs the butter using her more sensitive left hand. In that case, according to the one-shot rule, Dana should be certain that the butter weighs 15 oz. if she learns that it weighs 15 oz. or 16 oz.; otherwise she should be certain that it weighs 17 oz. If Dana updates in this way, then the expected accuracy of her future credences would be −2.4. On the other hand, if she updates by BCOND, then the expected accuracy of her future credences would be −0.6.

These results suggest a conjecture: although updating by this one-shot rule is more accurate than updating by BCOND in *Dana*, this is more of an exception than the norm. *In general*, updating by BCOND is more accurate than updating by the one-shot rule.

To make this conjecture more precise, we need a way of evaluating how accurate an updating rule is "in general." One simple idea is to look at the sum of the rule's expected accuracies in different updating problems. However, there is one immediate technical difficulty with this proposal: the class of updating problems does not form a set—since it is at least as large as the class of finite sets—so the sum would not even be well defined.[33] Moreover, even if we restrict attention to updating problems associated with a fixed (finite) state space, there are still infinitely many such updating problems—in fact, uncountably infinitely many (since there are uncountably infinitely priors that one could define over that state space). As a result, while the sum of the expected accuracies of an updating rule is well defined in this case (using the counting

33. Another conceptual difficulty with this approach is that it is not clear that accuracy scores relative to different state spaces are directly comparable with each other. For instance, if we measure accuracy using the additive Brier score, then credence functions over larger state spaces will in general have higher inaccuracy scores than credence functions over smaller state spaces, simply because there are more propositions for the credence functions to accrue inaccuracy scores. While there are technical ways around this problem, each of these ways seems to have its own problems. See Carr 2015; Pettigrew 2018; Talbot 2019 for related discussion.

measure), it will be infinite for almost every updating rule except for the truth rule.[34]

To get around these issues, we propose to evaluate the total expected accuracy of an updating rule with respect to the set of updating problems associated with a given prior space. Formally, let f be an updating rule and let (S, π) be a prior space. Let $\mathscr{E} \subseteq \{E : S \to \mathscr{P}(S)\}$ denote all factive learning situations associated with (S, π). The **total expected accuracy of** f **relative to** (S, π) is given by the quantity

$$\sum_{E \in \mathscr{E}} \sum_{s \in S} \pi(s) \mathsf{A}\big(f_{(S,\pi,E)}(s), s\big).$$

An updating rule is "most accurate in general" if it has the highest total expected accuracy relative to every prior space (S, π).

One quick remark about this definition of total expected accuracy before moving on. If f satisfies global evidential constancy, then its total expected accuracy relative to a prior space (S, π) is proportional to the following quantity:

$$\sum_{s \in S} \pi(s) \sum_{p \subseteq S : s \in p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big).$$

In particular, we have the following.

Fact 9. *Suppose that f is globally evidentially constant. Then for each state space S, there is a positive constant α_S such that, for all π over S, the total expected accuracy of f relative to (S, π) is proportional to the above quantity by a factor of α_S.*

It turns out that this second quantity is often easier to calculate. Since we are only interested in comparing the accuracy of updating rules that satisfy global evidential constancy, it does not matter which quantity we use. So, without loss of generality, we will typically work with the sec-

34. We can circumvent this problem by evaluating the infinite sum using hyperreals, or introducing additional measure-theoretical structures to the set of updating problems and evaluating the sum using certain finite measures. However, both approaches would involve some degrees of arbitrariness that are not clearly well motivated. Moreover, given our main positive result (fact 10), these alternative approaches are likely to converge on a similar verdict.

ond measure, and refer to it as the total expected accuracy of f relative to (S, π) as well.[35]

With an exact definition of "most accurate in general" in hand, we are finally in a position to state our main proposal precisely, which is the following bridge principle between accuracy and rationality:

> **Highest Total Expected Accuracy (Hi-TEA)**. An updating rule is epistemically rational if and only if, of all updating rules that satisfy global evidential constancy, it is the most accurate in general, in the sense that it has the highest total expected accuracy relative to every prior space.

The immediate question is: Of all updating rules that satisfy global evidential constancy, is there an updating rule that is most accurate in general in this sense? If not, then we are back to square one with respect to our earlier impossibility result.

5.2. Bayes Is Back

Fortunately, it turns out the answer is Yes, and that rule is Bayesian conditionalization.

Fact 10. *For every prior space (S, π), BCOND has the highest total expected accuracy with respect to (S, π) of all updating rules that satisfy global evidential constancy.*

Before going through the formal proof, let us illustrate this result with two examples. Earlier we conjectured that although updating by the one-shot rule f has higher expected accuracy than BCOND in Dana's updating problem, it is not more accurate than BCOND in general. We can now show that this is the case, if we measure accuracy in general by total expected accuracy. As above, let (S, π) be the prior space associated

35. Jennifer Rose Carr (2021) makes a proposal that is formally similar to ours. Like us, Carr suggests that we should evaluate updating rules by a weighted average of their accuracies at all possible world-evidence combinations. And while Carr doesn't explicitly discuss the assumption of noncontrastiveness, she takes an updating rule relative to a state space S to be a function from world-evidence pairs to credence functions over S, which is formally equivalent to noncontrastiveness in our framework. However, our proposals differ in terms of the philosophical interpretation of the aggregate quantity. For Carr (2021), the aggregate quantity represents the agent's subjective estimate of how accurate an updating rule is, given her uncertainty about what she might learn at a given state. For us, the total expected accuracy of an updating rule relative to a prior space is part of an objective measure of how accurate the updating rule is in general *for any agent*, irrespective of their uncertainties about their evidential situations.

Table 2. Comparison of the accuracy of BCOND and the one-shot rule f at each state.

s	p	$\mathsf{A}(f_{(S,\pi)}(p), s)$	$\mathsf{A}(b_{(S,\pi)}(p), s)$
15	$\{15\}$	0	0
	$\{15, 16\}$	0	−2.25
	$\{15, 17\}$	−1	−1
	$\{15, 16, 17\}$	−4	−2.08
16	$\{16\}$	0	0
	$\{15, 16\}$	−4	−0.25
	$\{16, 17\}$	−4	−0.25
	$\{15, 16, 17\}$	0	−0.48
17	$\{17\}$	0	0
	$\{16, 17\}$	0	−2.25
	$\{15, 17\}$	−1	−1
	$\{15, 16, 17\}$	−4	−2.08

with Dana and Noah's updating problems. Table 2 shows a comparison of the Brier score of the one-shot rule f and BCOND at each state. To calculate their respective total expected accuracies, we proceed as follows (see fact 9): For each state s, we add the accuracies of the credence functions that the rule recommends the agent with prior π to adopt at s, given each evidence proposition $p \subseteq S$ compatible with s. Then we take the weighted sum of those numbers, weighted by the prior credence $\pi(s)$ in the state. Crunching the numbers, we see that the one-shot rule f has the total expected accuracy

$$0.2(-1-4) + 0.6(-4-4) + 0.2(-1-4) = -6.8,$$

whereas BCOND has total expected accuracy

$$0.2(-2.25-1-2.08) + 0.6(-0.25-0.25-0.48) \\ + 0.2(-2.25-1-2.08) = -2.72,$$

which is strictly higher.

The same is true of the dogmatic rule (table 3). Suppose again that we measure accuracy using the Brier score. As one can see from the table, the fact that updating by the dogmatic rule has higher expected accuracy than updating by BCOND in Dana's updating problem is very much the exception than the rule—for any state s and any proposition p other than the one that Dana will learn at s, conditionalizing on p

Table 3. Comparison of the accuracy of BCOND and the dogmatic rule d at each state.

s	p	$\mathsf{A}(d_{(S,\pi)}(p), s)$	$\mathsf{A}(b_{(S,\pi)}(p), s)$
15	$\{15\}$	-2.08	0
	$\{15, 16\}$	-2.08	-2.25
	$\{15, 17\}$	-2.08	-1
	$\{15, 16, 17\}$	-2.08	-2.08
16	$\{16\}$	-0.48	0
	$\{15, 16\}$	-0.48	-0.25
	$\{16, 17\}$	-0.48	-0.25
	$\{15, 16, 17\}$	-0.48	-0.48
17	$\{17\}$	-2.08	0
	$\{16, 17\}$	-2.08	-2.25
	$\{15, 17\}$	-2.08	-1
	$\{15, 16, 17\}$	-2.08	-2.08

would result in a more accurate credence function at s than retaining one's prior. This is born out by their respective total expected accuracy scores: the dogmatic rule has a total expected accuracy of -4.48, which is significantly lower than that of BCOND (-2.72), though not as low as that of the one-shot rule.

With these examples in hand, let us now work through the proof of fact 10. Fix a prior space (S, π) and let f be a globally evidentially constant updating rule. As noted above, the total expected accuracy of f with respect to (S, π) is proportional to the quantity

$$\sum_{s} \pi(s) \sum_{p \subseteq S : s \in p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big).$$

The crucial observation is that we can reexpress this quantity as[36]

$$\sum_{p \subseteq S} \pi(p) \sum_{s \in S} \pi(s|p) \mathsf{A}\big(f_{(S,\pi)}(p), s\big).$$

Now note that, for a fixed p, the inner sum is the expected accuracy of the credence function $f_{(S,\pi)}(p)$ according to $\pi(.|p)$. Since $\pi(.|p)$

36. To see why, let $\mathsf{A}(f_{(S,\pi)}(p), s) = h(s)$. Note that $\sum_s \pi(s) \sum_{p \subseteq S : s \in p} h(s) = \sum_{(s,p) : s \in p, p \subseteq S} \pi(s) h(s) = \sum_{p \subseteq S} \sum_{s \in p} \pi(s) h(s) = \sum_{p \subseteq S} \sum_{s \in S} \pi(s|p) \pi(p) h(s) = \sum_{p \subseteq S} \pi(p) \sum_{s \in S} \pi(s|p) h(s)$, where the second last equality follows from the fact that, if $s \notin p$, then $\pi(s|p)\pi(p) = 0$, and if $s \in p$, then $\pi(s) = \pi(s|p)\pi(p)$.

is probabilistic, strict propriety entails that this inner sum is maximized if and only if $f_{(S,\pi)}(p)$ equals $\pi(.|p)$. By regularity of π, it follows that the entire sum is maximized if and only if, for all p, $f_{(S,\pi)}(p)$ equals $\pi(.|p)$. This completes the proof.

To recap, we started with the question: *Why is* BCOND*, and not the truth rule, the rule of rational credal revision?* We argued that the externalist Bayesian can't answer this explanatory challenge by appealing to the idea that the rational updating rule is the most accurate followable rule. Even if the truth rule is not followable, it is not obvious that BCOND is always followable in opaque updating problems either, and even if it is, it is unlikely to be the most accurate followable updating rule, given plausible assumptions about followability. Instead, we propose that the externalist Bayesian answers this explanatory challenge by appealing to the principle Hi-TEA: the rational updating rule is the one that satisfies global evidential constancy and is most accurate in general, in the sense that of all updating rules that satisfy global evidential constancy, it has the highest total expected accuracy relative to every prior space. We've shown that this principle entails that BCOND is the unique rational updating rule.

5.3. *Objections*

Before we conclude, let us consider some objections against our proposed principle.

One objection is that the notion of total expected accuracy has no normative weight. Yes, in some sense, the one-shot rule is less accurate than BCOND *in general.* But why should anyone care about that? In particular, why should Dana care about that? After all, by stipulation, Dana knows her learning situation, and updating by the one-shot rule, or even retaining her prior, is more accurate than updating by BCOND in her learning situation. The fact that either rule has lower expected accuracy in other learning situations, such as Noah's, seems neither here nor there, as far as Dana is concerned.[37]

Let us distinguish between two versions of this objection. On the first version, the objection is not so much about the normative signifi-

37. This objection is analogous to the collapse objection against rule consequentialism, according to which the best utilitarian rule is simply the rule that everyone performs the act that maximizes their own expected value (see, e.g., Smart and Williams 1973; Kagan 2000; Hooker 2002, chap. 4). Thanks to an anonymous referee for pressing us on this point.

cance of total expected accuracy as it is about the requirement of global evidential constancy (and noncontrastiveness/idempotence in particular). The thought is that the rational updating rule just is the rule that recommends the most accurate updating plan in every updating problem. If an agent learns the same proposition twice in two different updating problems, and the most accurate updating plans for those two problems recommend different posteriors over states in response to learning that same proposition, then the agent should revise her credences accordingly. This may seem counterintuitive at first glance, but it is quite sensible—the thought goes—once we take the relationship between epistemic rationality and accuracy seriously. Just as the pursuit of accuracy requires planning to adopt different posteriors over states depending on what one learns, it might also require planning to adopt different posteriors over states depending on how many times one learns a proposition, if the situation in which one is going to learn that proposition is changing.

Our response to this version of the objection is relatively flatfooted. Since BCOND satisfies global evidential constancy, to the extent that one is skeptical of global evidential constancy as a constraint that the rational updating rule satisfies, one must also be skeptical of the rationality of BCOND. As noted in the introduction, the objective of this article is not to convince the skeptic of Bayesianism why BCOND is the unique rational updating rule; rather, our proposal is pitched at the externalist Bayesian, and in particular, our aim is to offer an answer to the explanatory challenge: *Why is BCOND the rational updating rule and not the truth rule?*, in a way that (i) does not assume evidence internalism and (ii) respects a connection between accuracy and epistemic rationality. As we acknowledged, an externalist could dissolve this explanatory challenge by rejecting Bayesianism. In particular, they might think that the only thing relevant for determining whether an updating rule is rational is its truth-conduciveness, and different updating rules are truth-conducive for different agents, depending on their individual psychological makeups and learning situations. To this extent, updating by BCOND—or any updating rule that satisfies global evidential constancy—is only most accurate, and thereby rational, for special agents under special circumstances. The main goal of this article is to show that this radical response is not the only option available to the externalist; the externalist Bayesian has a leg to stand on as well.

However, one might question if this goal has really been achieved, namely, is our answer to the explanatory challenge on behalf of the externalist Bayesian a good one? If not, then perhaps the externalist Bayesian does not have a leg to stand on, after all. This brings us to the second version of the objection.

The second version of the objection concedes our claim that rational updating rules need to satisfy global evidential constancy, as well as the idea that rational updating rules are truth-conducive. The concern is about the robustness of the explanation for the truth-conduciveness of BCOND we have offered. In particular, we have been evaluating the truth-conduciveness of updating rules in terms of their total expected accuracy, which is defined as the sum over the rule's expected accuracies in all possible updating problems associated with a given prior space. But this mode of evaluation assumes that all updating problems associated with a given prior space are equally important for judging a rule's truth-conduciveness. One could question this assumption. For instance, one might think that, for Dana, her actual updating problem is clearly more important than some hypothetical updating problems that she doesn't face but could have faced. And even if we abstract away from individual agents, perhaps the updating problem where one learns the truth of some laws of nature is more important than one where one learns the number of stones in Wisconsin. On this line of thought, when determining which updating rule that satisfies global evidential constancy is most accurate in general, we should look at, not the sum of the rule's expected accuracies in different updating problems associated with that prior space, but a *weighted* sum of its expected accuracies, weighted by the relative importance of those updating problems (subjectively or objectively speaking). Intuitively, on this alternative approach, BCOND would not necessarily be the rule that is most accurate in general. If that is right, then it appears that our explanation of the rationality of BCOND is rather thin after all.

Before addressing this worry, let us make it more precise. Let (S, π) be a prior space. Let λ be a function that assigns each learning situation $E : S \rightarrow \mathcal{P}(S)$ a nonnegative real number $\lambda(E)$ (we also assume that $\lambda(E)$ is positive for some E). Intuitively, $\lambda(E)$ represents how epistemically important the updating problem (S, π, E) is. The suggestion, then, is that we should define the total expected accuracy of an updating rule f relative to (S, π) by the following quantity:

$$\sum_{E \in \mathcal{E}} \lambda(E) \sum_{s} \pi(s) \mathsf{A}\big(f_{(S,\pi,E)}(s), s\big).$$

Our proposal assumes that λ is the constant function that assigns 1 to every $E \in \mathscr{E}$. The worry is that this assumption is too restrictive. For our explanation of the rationality of BCOND to be sufficiently robust and interesting, we need to show that BCOND has the highest total expected accuracy relative to a range of "natural" weighting functions.

We think this is a good worry, and exploring its full implications will turn out to be both conceptually and dialectically illuminating. Let us start with the positive part of our reply. As it turns out, our result does generalize to a wide range of weighting functions that satisfy a natural property that we call *extensionality*, which is defined as follows. Fix a prior space (S, π). Let λ be a weighting function. Then for each proposition $p \subseteq S$, λ induces a random variable $\lambda_p : S \to \mathbb{R}$ given by

$$\lambda_p(s) := \sum_{E \in \mathscr{E}: E(s) = p} \lambda(E).$$

Intuitively, one can think of $\lambda_p(s)$ as representing the epistemic importance of updating accurately in response to learning proposition p at state s. We say λ is **extensional** if λ_p is a function of the truth-function of p, that is, for every $s, s' \in S$, if p has the same truth-values at s and s', then $\lambda_p(s) = \lambda_p(s')$. Informally, a weighting function is extensional if how important it is to update accurately in response to learning p at s does not depend on anything other than the truth of p at s. Any positive, constant weighting function is trivially extensional in this sense. For a less trivial example, let $\{E_1, \ldots, E_n\}$ be a set of transparent learning situations. Then the weighting function λ that assigns positive weights to all and only the E_i's is also extensional. Lastly, it's worth noting that the set of extensional weighting functions is closed under linear combinations, that is, if $\lambda_1, \ldots, \lambda_n$ are extensional, then so is $\sum_{k=1}^{n} a_k \cdot \lambda_k$.[38]

So there are (uncountably) many extensional weighting functions. Are there nonextensional weighting functions? Here is an example: Let E be an opaque learning situation. Consider the weighting func-

38. *Proof.* Suppose λ_k is extensional for $k \in \{1, \ldots, n\}$. Let $\lambda = \sum_{k=1}^{n} a_k \cdot \lambda_k$. Then

$$\lambda_p(s) = \sum_{E \in \mathscr{E}: E(s) = p} \lambda(E) = \sum_{E \in \mathscr{E}: E(s) = p} \sum_{k=1}^{n} a_k \cdot \lambda_k(E) = \sum_{k=1}^{n} a_k \cdot (\lambda_k)_p(s).$$

Since $(\lambda_k)_p(s)$ is a function of $v_s(p)$ for each k, their linear combination is a function of $v_s(p)$ as well.

tion λ that assigns 1 to E and 0 to any other learning situation relative to (S, π). Then λ is not extensional in our sense. In particular, let $s, s' \in S$ be such that $s' \in E(s)$ and $E(s) \neq E(s')$—such a pair exists because E is opaque. Let $p = E(s)$. Then $\lambda_p(s) = 1$ whereas $\lambda_p(s') = 0$, even though $v_s(p) = v_{s'}(p) = 1$.

Clearly, if we use the above weighting function, then BCOND would not be the globally evidentially constant updating rule that has the highest total expected accuracy relative to (S, π); instead, it would be any "one-shot"-type updating rule that is extensionally equivalent to META-COND with respect to E. As it turns out, this result generalizes: for BCOND to have the highest weighted total expected accuracy relative to a prior space (S, π), it is necessary and sufficient that the weighting function is extensional in the above sense.

Fact 11. *For every prior space (S, π) and weighting function λ, BCOND has the highest total expected accuracy relative to λ if and only if λ is extensional.*[39]

Fact 11 reveals both the generality and limitation of our explanation of the rationality of BCOND based on Hi-TEA. It is quite general, as it is valid for any way of weighing the expected accuracies of updating rules in different updating problems that is extensional. It is also limitative, to the extent that it is valid only for those weighting functions.

What shall we make of this? At this point, we think we are getting close to the bedrock of normative explanations. The Bayesians could dig in their heels and insist that extensionality is a natural assumption that "cuts epistemology at its joints." Which updating rule is most accurate in general is a matter of which updating rule gives the most accurate recommendations given different evidence, where it is assumed that the importance of the rule giving accurate recommendations given p depends only on the truth of p, and not on other epistemic or pragmatic matters which may vary from agent to agent. It is a nontrivial fact that *this assumption alone* entails that BCOND is the most accurate updating rule in general (of those that satisfy global evidential constancy). The non-Bayesian, on the other hand, might respond with an incredulous stare: *clearly* truth is not the only factor that determines how important it is to update accurately in response to learning p; it also matters *how likely it is for one to learn p at s*, and perhaps also *how epistemically important p is as a propo-*

39. If $\lambda(E) = 0$ for some E, then BCOND may not be the unique updating rule that has the highest expected accuracy, though it is one of such rules when λ is extensional.

sition at s. On this picture, the most accurate updating rule in general would not be Bayesian conditionalization, but that is just a fact that we have to live with rather than explain away. Epistemology is messy, and we shouldn't pretend it to be otherwise.

While we don't have much to say to satisfy the non-Bayesians, one thing to note is that these two perspectives are not necessarily incompatible with one another. Just as deterministic dynamics at the microscopic level is compatible with nontrivial probabilistic patterns at the macroscopic level, perhaps the theory of Bayesianism can be treated as a theory of "high-level" rationality that coexists with a theory of "low-level" rationality. The latter takes into account details of individual psychological dispositions and environmental conditions that the former abstracts away. Epistemology can be both clean and messy, depending on the level of theorizing.[40]

Let us conclude with a side remark. Throughout this article, we have assumed that the explanatory challenge *Why is* BCOND *and not the truth rule the rational updating rule?* is more of a challenge for the externalist Bayesian than it is for the internalist Bayesian. In particular, the internalist Bayesian can justify BCOND on the basis that it is the most accurate updating rule that is followable, thanks to Greaves and Wallace's result (fact 1). The externalist Bayesian cannot. However, fact 11 suggests a different perspective on this issue. As noted above, any weighting function that assigns positive weight to all and only transparent learning situations is extensional. This suggests a strong connection between extensionality on the one hand, and evidence internalism on the other. In fact, we can make this connection precise by the following fact.

Fact 12. *Let* (S, π, E) *be an updating problem. Let* λ *be the weighting function relative to* (S, π) *that assigns* 1 *to* E *and* 0 *to any other learning situation. Then* λ *is extensional if and only if* E *is transparent.*

Thus, from an accuracy-theoretic perspective, Greaves and Wallace's result is really a special case of fact 11; in particular, the result assumes evidence internalism only to the extent that it assumes extensionality. But one does not have to be an internalist to subscribe to extensionality; the principle is compatible with evidence externalism as well. Conversely, the internalist could reject extensionality, perhaps based on

40. Thanks to an anonymous referee for pressing us on this point.

considerations of contingent epistemic importance (Levinstein 2019), and for such internalist Bayesians, the Greaves and Wallace result will be of no avail either. To this extent, insofar as addressing the explanatory challenge is concerned, the important question is not whether evidence internalism is true, but whether the (total) expected accuracy of an updating rule should be evaluated in a way that satisfies extensionality.

6. Conclusion

Why is Bayesian conditionalization the rule of rational credal revision, despite it being less accurate than the truth rule? In this article, we proposed an answer to this challenge that does not assume evidence internalism. The core of our proposal is the new accuracy-rationality principle Hi-TEA: the rational updating rule satisfies global evidential constancy and, of all updating rules that satisfy this property, is the most accurate in general. Bayesian conditionalization is the rule of rational credal revision because it is the unique updating rule that satisfies these two desiderata. As one can infer from the suggestive remarks that we made in the last section, we do not take this to be the final word on this problem, but rather a starting point of further investigation into the relationship between accuracy and epistemic rationality.

Appendix A.

Fact 1 is proved in Greaves and Wallace 2006.

Fact 2 is proved in Schoenfield 2017.

Fact 3. *Suppose that accuracy is measured by the Brier score or the additive logarithmic score. Then the dogmatic rule has higher expected accuracy than* BCOND *in Dana's updating problem.*

Proof. Given a finite state space S and a credence function c defined over all subsets of S, the *Brier score* evaluates $\mathsf{A}(c, s)$ as $-\sum_{p\subseteq S} |c(p) - v_s(p)|^2$, where v_s is the omniscient credence function at s. The *additive logarithmic score* evaluates $\mathsf{A}(c, s)$ as $-\sum_{p\subseteq S} \mathfrak{L}(v_s(p), c(p))$, where $\mathfrak{L}(1, x) = -\ln x$ and $\mathfrak{L}(0, x) = -\ln(1-x)$. In Dana's updating problem, $S = \{15, 16, 17\}$, π is probabilistic with $\pi(16) = 0.6$ and $\pi(15) = \pi(17) = 0.2$, and E is given by

$$E(15) = \{15, 16\}, \qquad E(16) = \{15, 16, 17\}, \qquad E(17) = \{16, 17\}.$$

Since π is probabilistic, $\pi(.|E(16)) = \pi$. So to show that the dogmatic rule has higher expected accuracy than BCOND in Dana's updating problem, it suffices to show that $\mathsf{A}(\pi, 15) > \mathsf{A}(\pi(.|E(15)), 15)$ and $\mathsf{A}(\pi, 17) > \mathsf{A}(\pi(.|E(17)), 17)$. We show that for the Brier score $\mathsf{A}(\pi, 15) > \mathsf{A}(\pi(.|\{15, 16\}), 15)$, the calculations for the other inequality, and for the additive logarithmic score, are similar. We have

$$\mathsf{A}(\pi, 15) = -\big(|0.2-1|^2 + |0.8-1|^2 + |0.4-1|^2 + |0.6|^2 + |0.8|^2 + |0.2|^2\big) = -2.08;$$

$$\mathsf{A}\Big(\pi\big(.\big|\{15, 16\}\big), 15\Big) = -\left(2 \cdot \left|\frac{0.2}{0.8} - 1\right|^2 + 2 \cdot \left|\frac{0.6}{0.8} - 0\right|^2\right) = -2.25. \quad \square$$

Fact 4. *Let* A *be any accuracy scoring rule which is concave in addition to strictly proper. Then any nontrivial mixture of* BCOND *and* META-COND *is evidentially constant, conservative, and has higher expected accuracy (relative to* A*) than* BCOND *in every updating problem where they do not agree.*

Proof. Let A be a strictly proper, concave scoring rule and let f be a nontrivial mixture of BCOND and META-COND. Since BCOND and META-COND are evidentially constant, it's straightforward to check that f is evidentially constant. It's also straightforward to check that f is conservative since BCOND is. Now fix an updating problem (S, π, E). Then there is $0 \leq w \leq 1$ such that

$$f_{(S,\pi,E)}(s) = w \cdot b_{(S,\pi,E)}(s) + (1-w) \cdot m_{(S,\pi,E)}(s).$$

Since A is concave,

$$\mathsf{A}\big(f_{(S,\pi,E)}(s), s\big) \geq w \cdot \mathsf{A}\big(b_{(S,\pi,E)}(s), s\big) + (1-w) \cdot \mathsf{A}\big(m_{(S,\pi,E)}(s), s\big).$$

It follows that

$$\begin{aligned}
&\mathbb{E}_\pi\big[\mathsf{A}(f_{(S,\pi,E)})\big] - \mathbb{E}_\pi\big[\mathsf{A}(b_{(S,\pi,E)})\big] \\
&\quad \geq w \cdot \mathbb{E}_\pi\big[\mathsf{A}(b_{(S,\pi,E)}\big] + (1-w) \cdot \mathbb{E}_\pi\big[\mathsf{A}(m_{(S,\pi,E)})\big] - \mathbb{E}_\pi\big[\mathsf{A}(b_{(S,\pi,E)}\big] \\
&\quad = (1-w) \cdot (\mathbb{E}_\pi\big[\mathsf{A}(m_{(S,\pi,E)})\big] - \mathbb{E}_\pi\big[\mathsf{A}\big(b_{(S,\pi,E)}\big]\big) \\
&\quad \geq 0,
\end{aligned}$$

where for the last step we invoke fact 2. Note that if f and BCOND do not agree in (S, π, E), then $1 - w > 0$ and META-COND and BCOND do not agree in (S, π, E), so the inequality is strict. □

Fact 5. *An updating rule f is globally evidentially constant if and only if there is a map f^* that takes as input a prior space (S, π) and proposition $p \subseteq S$ and outputs a posterior probability distribution $f^*_{(S,\pi)}(p)$ over S, such that for any updating problem (S, π, E),*

$$f_{(S,\pi,E)}(s) = f^*_{(S,\pi)}\big(E(s)\big).$$

Proof. For the "if" direction, let f be the updating rule given by $f_{(S,\pi,E)}(s) := f^*_{(S,\pi)}(E(s))$ for a given updating problem (S, π, E). Then f is trivially representable by f^* and for any updating problems (S, π, E) and (S, π, E'), for any $s, s' \in S$, if $E(s) = E'(s')$, then

$$f_{(S,\pi,E)}(s) = f^*_{(S,\pi)}\big(E(s)\big) = f^*_{(S,\pi)}\big(E'(s')\big) = f_{(S,\pi,E')}(s').$$

So f is globally evidentially constant.

For the "only if" direction, suppose f is a globally evidentially constant updating rule. Given (S, π), associate with each nonempty proposition $p \subseteq S$ a factive learning situation E_p (a function from S to subsets of S such that $s \in E_p(s)$ for all $s \in S$) and a state s_p such that $E_p(s_p) = p$; this is guaranteed to exist since $p \neq \emptyset$. We then specify $f^*_{(S,\pi)}$ by

$$f^*_{(S,\pi)}(p) = f_{(S,\pi,E_p)}(s_p).$$

To check this satisfies the desired property, let (S, π, E) be any updating problem and let $s \in S$ be any state. Let $p := E(s)$. Note $E(s) = E_p(s_p)$ by construction. So, by global evidential constancy of f,

$$f_{(S,\pi,E)}(s) = f_{(S,\pi,E_p)}(s_p) = f^*_{(S,\pi)}\big(E_p(s_p)\big) = f^*_{(S,\pi)}(p).$$

Therefore, f is representable by the two-place function f^*. □

Fact 6. *An updating rule is idempotent if and only if it is noncontrastive.*

Proof. Suppose f is idempotent. Let (S, π, E_1) and (S, π, E_2) be two updating problems involving the same prior space. Suppose $E_1(s) = E_2(s)$. By idempotence,

$$f_{(S,\pi,E_1)}(s) = f_{(S,\pi,E_1 \wedge E_2)}(s) = f_{(S,\pi,E_2)}(s).$$

So f is noncontrastive.

Conversely, suppose f is noncontrastive. Let (S,π,E_1) and (S,π,E_2) be two updating problems involving the same prior space and suppose $E_1(s) = E_2(s)$. Then (S,π,E_1) has the same prior space as $(S,\pi,E_1 \wedge E_2)$ and $E_1(s) = E_1(s) \cap E_2(s)$. So by noncontrastiveness,

$$f_{(S,\pi,E_1)}(s) = f_{(S,\pi,E_1 \wedge E_2)}(s),$$

that is, f is idempotent. □

Fact 7 is proved in the main text.

Fact 8. *Let* A *be any strictly proper scoring rule. Let f be a globally evidentially constant updating rule. Then for any prior space (S,π), there is at least one learning situation E relative to (S,π) such that f has lower expected accuracy in (S,π,E) than some other globally evidentially constant rule. Moreover, if the accuracy scoring rule is concave in addition to strictly proper, then there is a learning situation E in which, for any $\delta > 0$, f has lower expected accuracy in (S,π,E) than some globally evidentially constant rule that is δ-close to f, in the sense of section 3.3.*

Proof. Suppose f is globally evidentially constant. Fix a prior space (S,π). Since META-COND isn't globally evidentially constant, there must be some updating problem (S,π,E) associated with that prior space in which f disagrees with META-COND. It then follows that there is some $q \subseteq S$ such that the credence function recommended by f in response to learning q is not $\pi(.|E = q)$. Now consider the rule g that agrees with f, except that, for all updating problems involving the prior space (S,π), it recommends

$$g_{(S,\pi)}(p) = \begin{cases} \pi(.|E = p) & \text{when } E(s) = p \text{ for some } s, \\ f_{(S,\pi)}(p) & \text{when } E(s) \neq p \text{ for all } s. \end{cases}$$

That is, this rule "mimics" META-COND in the updating problem (S,π,E).

Note that g satisfies global evidential constancy (by fact 5). Since it coincides with META-COND in (S,π,E), it has higher expected accuracy than f in this problem relative to any strictly proper scoring rule A. So f does not maximize expected accuracy in (S,π,E). If A is in addition concave, then we can choose g such that

$$g_{(S,\pi)}(p) = \begin{cases} (1-\epsilon)\cdot f_{(S,\pi)}(p) + \epsilon\cdot \pi(.|E=p) & \text{when } E(s)=p \text{ for some } s, \\ f_{(S,\pi)}(p) & \text{when } E(s)\neq p \text{ for all } s \end{cases}$$

for $0 < \epsilon \leq 1$ that satisfies $\epsilon < \delta/2$. □

Fact 9. *Suppose that f is globally evidentially constant. Then for each state space S, there is a positive constant α_S such that, for all π over S, the total expected accuracy of f relative to (S,π) is proportional to the quantity*

$$\sum_{s\in S}\pi(s)\sum_{p\subseteq S: s\in p}\mathsf{A}\big(f_{(S,\pi)}(p), s\big)$$

by a factor of α_S.

Proof. Fix (S,π). Let $\mathscr{E}$ denote the set of all (factive) learning situations that can be defined with respect to the state space S, that is, the set of all learning situations E such that (S,π,E) is an updating problem. Fix a state $s\in S$. For each proposition $p\subseteq S$ compatible with s, consider the set of learning situations that yield evidence p at s, that is, the set $\{E\in\mathscr{E} : E(s)=p\}$. Each $p\subseteq S$ compatible with s yields a distinct such set, and these sets together partition $\mathscr{E}$. Thus, a rule's total expected accuracy (with respect to (S,π)) can be reexpressed as a sum of its expected accuracy given different possible evidence and learning situations in $\mathscr{E}$ that yield that evidence:

$$\begin{aligned}\sum_{E\in\mathscr{E}}\sum_{s\in S}\pi(s)\mathsf{A}\big(f_{(S,\pi,E)}(s), s\big) &= \sum_{s\in S}\pi(s)\sum_{E\in\mathscr{E}}\mathsf{A}\big(f_{(S,\pi,E)}(s), s\big)\\ &= \sum_{s\in S}\pi(s)\sum_{p\subseteq S: s\in p}\ \sum_{E\in\mathscr{E}: E(s)=p}\mathsf{A}\big(f_{(S,\pi,E)}(s), s\big).\end{aligned}$$

Now, we make two observations. First, continuing the chain of equalities, if f is globally evidentially constant, then

$$\sum_{s \in S} \pi(s) \sum_{p \subseteq S : s \in p} \sum_{E \in \mathscr{E} : E(s) = p} \mathsf{A}\big(f_{(S,\pi,E)}(s), s\big)$$
$$= \sum_{s \in S} \pi(s) \sum_{p \subseteq S : s \in p} \sum_{E \in \mathscr{E} : E(s) = p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big).$$

Since the inner summand is now independent of the specific choice of $E \in \mathscr{E}$, it suffices to calculate the number of $E \in \mathscr{E}$ that yield evidence p at s. This leads to our second observation: Let d be the number of states in S, excluding s. For each of these d states, there are 2^d propositions that can be learned at those states (for each state other than the given state, there are two choices of whether that state is in or out of the evidence proposition). Thus the number of learning situations in $\mathscr{E}$ that yield evidence p at s is equal to the constant $\alpha_S = (2^d)^d$. Thus

$$\sum_{s \in S} \pi(s) \sum_{p \subseteq S : s \in p} \sum_{E \in \mathscr{E} : E(s) = p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big) = \sum_{s \in S} \pi(s) \alpha_S \cdot \sum_{p \subseteq S : s \in p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big)$$
$$= \alpha_S \sum_{s \in S} \pi(s) \sum_{p \subseteq S : s \in p} \mathsf{A}\big(f_{(S,\pi)}(p), s\big),$$

as claimed. (Thanks to an anonymous referee for supplying the key ideas for this streamlined proof.) □

Fact 10 is proved in the main text.

Fact 11. *For every prior space (S, π) and weighting function λ, BCOND has the highest total expected accuracy relative to λ if and only if λ is extensional.*

This fact is a consequence of the following two lemmas.

Lemma 1. *Let (S, π) be a prior space and let λ be a weighting function. For every $p \subseteq S$, define $\pi(.|\lambda_p)$ as the following probability function:*

$$\pi(s|\lambda_p) := \frac{\pi(s)\lambda_p(s)}{\sum_{s' \in S} \pi(s')\lambda_p(s')} = \frac{\pi(s)\lambda_p(s)}{\mathbb{E}_\pi[\lambda_p]}$$

(provided that $\mathbb{E}_\pi[\lambda_p] > 0$). Let f be an updating rule such that $f_{(S,\pi)}(p) = \pi(.|\lambda_p)$ whenever $\mathbb{E}_\pi[\lambda_p] > 0$. Then f has the highest total expected accuracy of all globally evidentially constant updating rules relative to λ and (S, π).

Proof. Fix (S, π) and let f be any globally evidentially constant updating rule. By similar steps as the proof of fact 10,

$$\begin{aligned}\mathbb{E}_\pi\big[\mathsf{A}(f_{(S,\pi)}\big] &= \sum_s \pi(s) \sum_{p \subseteq S} \lambda_p(s) \mathsf{A}\big(f_{(S,\pi)}(p), s\big) \\ &= \sum_{p \subseteq S} \sum_s \pi(s) \lambda_p(s) \mathsf{A}\big(f_{(S,\pi)}(p), s\big) \\ &= \sum_{p \subseteq S} \mathbb{E}_\pi[p_\lambda] \sum_s \frac{\pi(s)\lambda_p(s)}{\mathbb{E}_\pi[p_\lambda]} \mathsf{A}\big(f_{(S,\pi)}(p), s\big).\end{aligned}$$

Note that the inner sum over s (on the final line) is just the expected accuracy of a fixed credence function (namely, $f_{(S,\pi)}(p)$) according to the probability function $\pi(.|\lambda_p)$. By strict propriety, this is maximized precisely by $\pi(.|\lambda_p)$, as claimed. □

Lemma 2. *λ is extensional if and only if, for all (S, π), $\pi(.|\lambda_p) = \pi(.|p)$ for all nonempty $p \subseteq S$.*

Proof. First we note that, given (S, π), if $\pi(.|\lambda_p)$ is well defined, then $\pi(p) > 0$ and so $\pi(.|p)$ is well defined; to see this, note λ_p is a nonnegative variable, so if $\mathbb{E}_p[\lambda_p] > 0$, then $\pi(s') > 0$ for some $s' \in p$, and so $\pi(p) > 0$. Now, for the "only if" direction, fix any (S, π) and $p \subseteq S$. Suppose $\lambda_p(s) = \lambda_p(s') = \alpha_p \geq 0$ for all $s, s' \in p$. Suppose $\pi(.|\lambda_p)$ is well defined. Then $\mathbb{E}_p[\lambda_p] > 0$ and $\pi(p) > 0$ and we have, using the fact that $\lambda_p = \alpha_p 1_p$ where 1_p is the indicator function for p,

$$\pi(s|\lambda_p) = \frac{\pi(s)\lambda_p(s)}{\mathbb{E}_\pi[\lambda_p]} = \frac{\pi(s)\alpha_p 1_p(s)}{\sum_{s' \in S} \pi(s')\alpha_p 1_p(s)} = \begin{cases} \frac{\pi(s)}{\pi(p)} & \text{if } s \in p \\ 0 & \text{otherwise,} \end{cases}$$

which is equal to $\pi(.|p)$, as claimed. For the "if" direction, fix any (S, π) and $p \subseteq S$. Suppose $\pi(.|\lambda_p)$ is well defined, and so $\mathbb{E}_p[\lambda_p] > 0$ and $\pi(p) > 0$, and that for all $s \in S$, $\pi(s|\lambda_p) = \pi(s|p)$. Let $s \in p$. Then

$$\frac{\pi(s)\lambda_p(s)}{\mathbb{E}_\pi[\lambda_p]} = \frac{\pi(s)}{\pi(p)},$$

which implies $\lambda_p(s) = \mathbb{E}_\pi[\lambda_p]/\pi(p)$. Note that this value is the independent of $s \in p$. This argument can be repeated for any (S, π) and $p \subseteq S$, so it follows that λ is extensional. □

Fact 12. *Let (S, π, E) be an updating problem. Let λ be the weighting function relative to (S, π) that assigns 1 to E and 0 to any other learning situation. Then λ is extensional if and only if E is transparent.*

Proof. Suppose λ is extensional. Let $s \in S$. Suppose $p = E(s)$ and $s' \in p$. Then $v_s(p) = v_{s'}(p)$. So $1 = \lambda_p(s) = \lambda_p(s')$, that is, $E(s') = p$. Conversely, suppose E is transparent. Let $p \subseteq S$ and $s, s' \in S$. It suffices to show that, if $v_s(p) = v_{s'}(p)$, then $E(s) = p$ if and only if $E(s') = p$. Suppose $v_s(p) = v_{s'}(p) = 0$. Then $E(s) \neq p$ and $E(s') \neq p$ by factivity. On the other hand, suppose $v_s(p) = v_{s'}(p) = 1$. Suppose $E(s) = p$, then since $s' \in p$, $E(s') = p$ by transparency. By symmetry, $E(s) = p$ if and if $E(s') = p$. □

References

Ahmed, Arif, and Bernhard Salow. 2019. "Don't Look Now." *British Journal for the Philosophy of Science* 70, no. 2: 327–50.

Barnett, Zach. 2021. "Rational Moral Ignorance." *Philosophy and Phenomenological Research* 102, no. 3: 645–64.

Berker, Selim. 2008. "Luminosity Regained." *Philosophers' Imprint* 8, no. 2: 1–22.

Bernardo, José M. 1979. "Expected Information as Expected Utility." *Annals of Statistics* 7, no. 3: 686–90.

Briggs, R. A. 2010. "Putting a Value on Beauty." In *Oxford Studies in Epistemology*, vol. 3, edited by Tamar Szabo Gendler and John Hawthorne. Oxford University Press.

Briggs, R. A., and Richard Pettigrew. 2020. "An Accuracy-Dominance Argument for Conditionalization." *Noûs* 54, no. 1: 162–81.

Bröcker, Jochen, and Leonard A. Smith. 2007. "Scoring Probabilistic Forecasts: The Importance of Being Proper." *Weather and Forecasting* 22, no. 2: 382–88.

Bronfman, Aaron. 2014. "Conditionalization and Not Knowing That One Knows." *Erkenntnis* 79, no. 4: 871–92.

Campbell-Moore, Catrin, and Benjamin A. Levinstein. 2021. "Strict Propriety Is Weak." *Analysis* 81, no. 1: 8–13.

Carnap, Rudolf. 1988. *Meaning and Necessity: A Study in Semantics and Modal Logic*. University of Chicago Press.

Carr, Jennifer Rose. 2015. "Epistemic Expansions." *Res Philosophica* 92, no. 2: 217–36.

Carr, Jennifer Rose. 2017. "Epistemic Utility Theory and the Aim of Belief." *Philosophy and Phenomenological Research* 95, no. 3: 511–34.

Carr, Jennifer Rose. 2021. "A Modesty Proposal." *Synthese* 198, no. 4: 3581–601.

Christensen, David. 2010. "Higher-Order Evidence 1." *Philosophy and Phenomenological Research* 81, no. 1: 185–215.

Das, Nilanjan. 2019. "Accuracy and Ur-Prior Conditionalization." *Review of Symbolic Logic* 12, no. 1: 62–96.

Das, Nilanjan. 2022. "Externalism and Exploitability." *Philosophy and Phenomenological Research* 104, no. 1: 101–28.

Das, Nilanjan. 2023. "The Value of Biased Information." *British Journal for the Philosophy of Science* 74, no. 1: 25–55.

Dorst, Kevin. 2019. "Higher-Order Uncertainty." In *Higher-Order Evidence: New Essays,* edited by Mattias Skipper and Asbjørn Steglich-Petersen. Oxford University Press.

Dorst, Kevin, Benjamin A. Levinstein, Bernhard Salow, Brooke E. Husic, and Branden Fitelson. 2021. "Deference Done Better." *Philosophical Perspectives* 35, no. 1: 99–150.

Easwaran, Kenny. 2014. "Regularity and Hyperreal Credences." *Philosophical Review* 123, no. 1: 1–41.

Elga, Adam. 2013. "The Puzzle of the Unmarked Clock and the New Rational Reflection Principle." *Philosophical Studies* 164, no. 1: 127–39.

Gallow, J. Dmitri. 2019a. "Diachronic Dutch Books and Evidential Import." *Philosophy and Phenomenological Research* 99, no. 1: 49–80.

Gallow, J. Dmitri. 2019b. *Learning and Value Change.* Michigan Publishing, University of Michigan Library.

Gallow, J. Dmitri. 2021. "Updating for Externalists." *Noûs* 55, no. 3: 487–516.

Geanakoplos, John. 1989. "Game Theory Without Partitions, and Applications to Speculation and Consensus." *BE Journal of Theoretical Economics* 21, no. 2: 361–94.

Gibbard, Allan. 2007. "Rational Credence and the Value of Truth." *Oxford Studies in Epistemology* 2: 143–64.

Gneiting, Tilmann, and Adrian E. Raftery. 2007. "Strictly Proper Scoring Rules, Prediction, and Estimation." *Journal of the American Statistical Association* 102, no. 477: 359–78.

Good, Irving John, 1967. "On the Principle of Total Evidence." *British Journal for the Philosophy of Science* 17, no. 4: 319–21.

Greaves, Hilary, and David Wallace. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* 115, no. 459: 607–32.

Greco, Daniel. 2019. "Fragmentation and Higher-Order Evidence." In *Higher-Order Evidence: New Essays,* edited by Mattias Skipper and Asbjørn Steglich-Petersen. Oxford University Press.

Hájek, Alan. 2011. "A Puzzle About Degree of Belief." Unpublished manuscript.

Hájek, Alan. 2012. Is Strict Coherence Coherent? *Dialectica* 66, no. 3: 411–24.

Hild, Matthias. 1998. "Auto-Epistemology and Updating." *Philosophical Studies* 92, no. 3: 321–61.

Hooker, Brad. 2002. *Ideal Code, Real World: A Rule-Consequentialist Theory of Morality*. Oxford University Press.

Horowitz, Sophie. 2014. "Epistemic Akrasia." *Noûs* 48, no. 4: 718–44.

Horwich, Paul. 2006. "The Value of Truth." *Noûs* 40, no. 2: 347–60.

Hughes, RIG, and Bas C. Van Fraassen. 1984. "Symmetry Arguments in Probability Kinematics." In *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, vol. 2. Philosophy of Science Association.

Isaacs, Yoaav, and Jeffrey Sanford Russell. 2023. "Updating Without Evidence." *Noûs* 57, no. 3: 576–99.

Jeffrey, Richard C. 1965. *The Logic of Decision*. University of Chicago Press.

Joyce, James M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65, no. 4: 575–603.

Joyce, James M. 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, vol. 342 of *Synthese Library*, edited by Franz Huber and Christoph Schmidt-Petri. Springer.

Kagan, Shelly. 2000. "Evaluative Focal Points." In *Morality, Rules, and Consequences: A Critical Reader*, edited by Brad Hooker, Elinor Mason, and Dale E. Miller. Edinburgh University Press.

Kelley, Mikayla. 2023. "On Accuracy and Coherence with Infinite Opinion Sets." *Philosophy of Science* 90, no. 1: 92–128.

Kelley, Mikayla, and Sven Neth. 2022. "Accuracy and Infinity: A Dilemma for Subjective Bayesians." *Synthese* 201, no. 12.

Kelly, Thomas. 2013. "Evidence Can Be Permissive." In *Contemporary Debates in Epistemology*, edited by Matthias Steup and John Turri. Blackwell

Levinstein, Benjamin A. 2017. "Accuracy Uncomposed: Against Calibrationism." *Episteme* 14, no. 1: 59–69.

Levinstein, Benjamin A. 2019. "An Objection of Varying Importance to Epistemic Utility Theory." *Philosophical Studies* 176, no. 11: 2919–31.

Levinstein, Benjamin A. n.d. "Higher-Order Evidence as Information Loss." Unpublished manuscript.

Lewis, Peter J., and Don Fallis. 2021. "Accuracy, Conditionalization, and Probabilism." *Synthese* 198, no. 5: 4017–33.

Medina, Pablo Zendejas. 2024. "Just as Planned: Bayesianism, Externalism, and Plan Coherence." *Philosophers' Imprint* 3, no. 28.

Mitova, Veli. 2018. *The Factive Turn in Epistemology*. Cambridge University Press.

Neth, Sven. Forthcoming. "Rational Aversion to Information." *British Journal for the Philosophy of Science*.

Nielsen, Michael. 2022. "Regular Conditional Probabilities and Strictly Proper Loss Functions." *Statistics and Probability Letters* 185, no. 109412.

Paul, Sarah K. 2014. "Diachronic Incontinence Is a Problem in Moral Philosophy." *Inquiry* 57, no. 3: 337–55.

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems*. Morgan Kaufmann.

Pettigrew, Richard. 2012. "Accuracy, Chance, and the Principal Principle." *Philosophical Review* 121, no. 2: 241–75.
Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence*. Oxford University Press.
Pettigrew, Richard. 2018. "The Population Ethics of Belief: In Search of an Epistemic Theory X." *Noûs* 52, no. 2: 336–72.
Pettigrew, Richard. 2020. "What Is Conditionalization, and Why Should We Do It?" *Philosophical Studies* 177, no. 11: 3427–63.
Pettigrew, Richard. 2022. "Radical Epistemology, Structural Explanations, and Epistemic Weaponry." *Philosophical Studies* 179, no. 1: 289–304.
Predd, Joel B., Robert Seiringer, Elliott H Lieb, Daniel N Osherson, H. Vincent Poor, and Sanjeev R. Kulkarni. 2009. "Probabilistic Coherence and Proper Scoring Rules." *IEEE Transactions on Information Theory* 55, no. 10: 4786–92.
Pruss, Alexander R. 2022. "Accuracy, Probabilism and Bayesian Update in Infinite Domains." *Synthese* 200, no. 6: 1–29.
Salow, Bernhard. 2018. "The Externalist's Guide to Fishing for Compliments." *Mind* 127, no. 507: 691–728.
Schoenfield, Miriam. 2017. "Conditionalization Does Not Maximize Expected Accuracy." *Mind* 126, no. 504: 1155–87.
Schoenfield, Miriam. 2018. "An Accuracy Based Approach to Higher Order Evidence." *Philosophy and Phenomenological Research* 96, no. 3: 690–715.
Schultheis, Ginger. Forthcoming. "Accurate Updating." *Philosophy of Science*.
Smart, J. J. C., and Bernard Williams. 1973. "An Outline of a System of Utilitarian Ethics." In *Utilitarianism: For and Against*. Cambridge University Press.
Srinivasan, Amia. 2015. "Normativity Without Cartesian Privilege." *Philosophical Issues* 25, no. 1: 273–99.
Stalnaker, Robert. 2015. "Luminosity and the KK Thesis." In *Externalism, Self-Knowledge, and Skepticism: New Essays*, edited by Sanford Goldberg. Cambridge University Press.
Talbot, Brian. 2019. "Repugnant Accuracy." *Noûs* 53, no. 3: 540–63.
Weatherson, Brian. 2013. "Margins and Errors." *Inquiry* 56, no. 1: 63–76.
Weatherson, Brian. 2019. *Normative Externalism*. Oxford University Press.
Wedgwood, Ralph. 2002. "The Aim of Belief." *Philosophical Perspectives* 16: 267–97.
Weisberg, Jonathan. 2009. "Commutativity or Holism? A Dilemma for Conditionalizers." *British Journal for the Philosophy of Science* 60, no. 4: 793–812.
White, Roger. 2009. "On Treating Oneself and Others as Thermometers." *Episteme* 6, no. 3: 233–50.
Williams, J. Robert G., and Richard Pettigrew. Forthcoming. "Consequences of Calibration." *British Journal of Philosophy of Science*.
Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford University Press.
Williamson, Timothy. 2011. "Improbable Knowing." In *Evidentialism and Its Discontents*, edited by Trent Dougherty. Oxford University Press.

Williamson, Timothy. 2014. "Very Improbable Knowing." *Erkenntnis* 79: 971–99.
Worsnip, Alex. 2018. "The Conflict of Evidence and Coherence." *Philosophy and Phenomenological Research* 96, no. 1: 3–44.
Ye, Ru. 2023. *Higher-Order Evidence and Calibrationism.* Cambridge University Press.