Updating Without Evidence

Yoaav Isaacs Jeffrey Sanford Russell *

March 2022

Abstract

Sometimes you are unreliable at fulfilling your doxastic plans: for example, if you plan to be fully confident in all truths, probably you will end up being fully confident in some falsehoods by mistake. In some cases, there is information that plays the classical role of *evidence*—your beliefs are perfectly discriminating with respect to some possible facts about the world—and there is a standard expected-accuracy-based justification for planning to *conditionalize* on this evidence. This planning-oriented justification extends to some cases where you do not have transparent evidence, in the sense that your beliefs are not perfectly discriminating with respect to any non-trivial facts. In other cases, accuracy considerations do not tell you to plan to conditionalize on any information at all, but rather to plan to follow a different updating rule. Even in the absence of evidence, accuracy considerations can guide your doxastic plan.

I The Nap

A What should I believe?

B The truth!

A Really? I thought the answer would have something to do with what my evidence supports. But before we get to that, I should clarify that I don't really just have "on-off" beliefs. My beliefs come in degrees between zero and one: I have *credences*. So my real question is: what degrees of belief are the best ones for me to have?

*Thanks to Jeremy Goodman, John Hawthorne, and an anonymous referee for comments.

B Well, the best degrees of belief to have are those that maximize epistemic utility. But I suppose that's just a fancy way of saying "best." More substantively, I say degrees of belief are better—have higher epistemic utility—when they are closer to the truth. If p is true, then the credence in p with highest epistemic utility is 1, and if p is false, then the credence in p with highest epistemic utility is 0. Does that fit with your conception of "best"?

A Yes, that fits with my goals. What I want from my beliefs is *accuracy*: closeness to the truth.[1]

B Then that answers your question. The best credences follow the *truth rule*: assign one to every truth, and zero to every falsehood. Like I said: if your goal is accuracy, then what you should believe is the truth.

A That's not a helpful answer! I can't have those degrees of belief.

B What, you have some problem with zeros or ones? Are those degrees of belief beyond your reach?

A No, of course that's not what I mean. I can have credence one just as well as any other number. I suppose it's even within my power to adopt a credence function which *happens* to conform to your truth rule. But if I did that, it would just be by luck. My attitudes just aren't as sensitive to the world as you seem to assume when you recommend the truth rule.

Let me put it another way. Suppose I *plan* to follow the truth rule. I might succeed in following it, by luck, but very likely I won't succeed. In fact, I have a good chance of ending up assigning credence one to a false proposition that way, which is as bad as it gets accuracy-wise. For that reason, the truth rule doesn't seem like a very good plan for me.

B Ah, I see you were really asking a different question then. The question you are interested in is not what degrees of belief are best for you to *have*. Rather, your question is what degrees of belief are best for you to *plan* to have. That's an interesting question, too![2]

1. Arguing for epistemic norms on the basis of accuracy considerations is a dominant theme in recent epistemology: for a small selection, see Joyce (1998), Moss (2011), Pettigrew (2016), Levinstein (2017), Schoenfield (2018), and Horowitz (2019), along with other works we discuss in more depth below.

2. Evaluating *doxastic plans* is another theme in recent epistemology: see for example Schafer (2014), Schoenfield (2018, sec. 6, and references therein), and Pettigrew (2016, ch. 14).

A Yes, that sounds right.

Let me tell you more about a particular situation. (This one should be easy.) My daughter is at the age where she's dropping her nap. She takes an afternoon nap roughly half the time. On the days when she has a good nap, she is usually in a good mood in the evening. On the days when she doesn't have a good nap, she is often really cranky. What I'm wondering is what degree of belief I should have that she will be cranky this evening, *after* I find out whether she naps.

I suppose I could plan to follow the truth rule. That is, I could plan to be *sure* she will be cranky if she really will be cranky, or else be *sure* she won't be cranky if she really won't be. But that plan would be futile. Even after I find out about her nap, I still won't be sure what degrees of belief would conform to the rule I planned to follow. I would just be *guessing* (compare Horowitz 2019; Holguín, forthcoming; Dorst and Mandelkern, forthcoming).

In fact, here's what I think I would really do if I planned to follow the truth rule: if my daughter naps, my best guess would be that I was in the Nap-and-not-Cranky state, so I would put all my credence in that state. If she doesn't nap, I would guess that I was in the no-Nap-and-Cranky state, and put all my credence in that state. (See Figure 1.) That would make my accuracy as good as possible if I really am in one of those two states. The problem is that for all I know now, I might be in a Nap-and-Cranky world or a no-Nap-and-not-Cranky world. That can happen. And in those cases, planning to follow the truth rule would end up making my credences extremely *inaccurate*. It doesn't look like a good trade-off. To be precise, my current *expected* accuracy for planning to follow the truth rule doesn't look all that great.

B That's right. I'll bet you can do better than that—and you won't be surprised how. How about you plan to proportion your belief to the evidence? That is, you could plan to follow this *conditionalization rule*: if she naps, for each proposition p, set your credence in p equal to your current *conditional* credence in p given Nap. Likewise, if she doesn't nap, set your credence in p equal to your current conditional credence in p given no-Nap. What do you think would happen if you planned to follow that rule?

A That's one I think I can handle in this situation. I think I would really have the credences the rule prescribes, whether or not she naps.

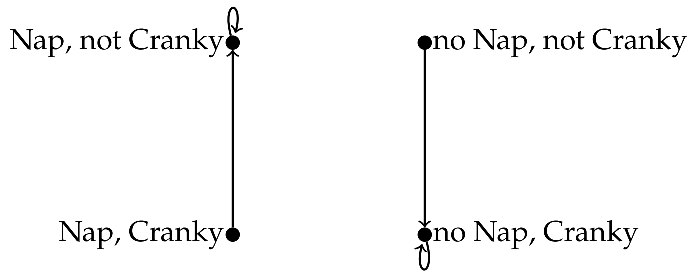

Figure 1: The guess graph for the Nap situation. Nodes in the graph represent states. An arrow from state s to s' means that in state s one guesses state s'.

B Great. So in that case, the expected accuracy of *planning* to follow the conditionalization rule is exactly the same as the expected accuracy of *following* the conditionalization rule. And this is better than what you get from planning to follow the truth rule.

Actually *following* the truth rule would be better (expected accuracy-wise) than conditionalizing. But *planning* to follow the truth rule is worse than *planning* to conditionalize. The basic thing that's going on here is that you are foreseeably imperfect at following the truth rule in this case, so the value of planning to follow that rule comes apart from the value of successfully following it.

A So is conditionalization the *best* rule for me to plan to follow?

B That seems plausible, but to entirely settle the question we'll need to fill in some background assumptions. Let's say a *doxastic rule* (just a *rule* for short) is a function that takes each state of the world to a credence function defined over states of the world.[3] You also have your *current* credences defined over these same states.

In order to evaluate the expected accuracy of planning to follow an *arbitrary* rule, we need to model what credences this will lead you to

3. We are considering "coarse-grained" states, for now, rather than "fine-grained" complete possible worlds. And we are thereby setting aside, for now, questions like what to believe about one's own credences. In our model of the Nap case there are just four states.

have. In this situation, what you said about the truth rule provides a general recipe. Say you plan to follow a rule f. Since all Nap states look the same to you (so to speak), the best you can do is take a guess as to which of the Nap states you are in. If you *guess* you are in state s, your attitude will be $f(s)$.

Your guess will be the same in all Nap states, and it will also be the same in all no-Nap states. So in this case, whatever rule you plan to follow, the attitude you will really end up having is entirely determined by which cell of a *partition* of states you are in: one cell contains all the Nap states, and the other contains all the no-Nap states. The question of which rule to plan to follow comes down to the question of which function from cells in that partition to credence functions maximizes expected accuracy. This is a question we know the answer to: it's conditionalization (Greaves and Wallace 2006; see also Leitgeb and Pettigrew 2010a, 2010b; Easwaran 2013; Pettigrew 2016, chs. 14–15; Briggs and Pettigrew 2018).

A Oh right, I think I knew that.

B We can say something very similar about any situation where the credence that results from planning to follow a certain rule is determined by your *best guess* at which state you are in, and your best guess is itself determined by which state you are *really* in. We can represent a situation like this with a *deterministic guess model*. We can work out the details later when we have some paper (see Appendix B), but here's the basic idea. A deterministic guess model is built up from two bits. First, there's a probability function representing your prior credences about which state you are really in. Second, there's a function g that describes your guessing behavior: "for each state s, if I'm in state s, I will guess I'm in state $g(s)$". Call this your *guess function*. This function induces a partition of states: for any state s' that you might guess, the set of states where you guess s' is a cell of the partition (that is, the set of states s such that $g(s) = s'$). Then planning to follow the rule that says to conditionalize on the cell of that partition given by your guess will be optimal.[4]

A This was a kind of complicated way to get to an unsurprising conclusion.

B Right, no surprises so far.

4. Though not quite uniquely optimal! See Theorem 2 in Appendix B.

II The Clock

A Can I tell you about another problem I have? In the other room, there is "a plain, unmarked circular dial with a single pointer that rotates in imperceptibly short discrete jerks, a modernist clock-face designed with an eye to the appearance rather than the reality of functional efficiency" (Williamson 2011, p. 153; see also Elga 2013; Christensen 2010; Williamson 2014).

At the moment, I think that the pointer is equally likely to be in any of its positions. Soon I'm going to look at that clock and update my opinion about where the pointer is. What should I think?[5]

B Well, as always, the truth rule would be the best rule to follow, if you could pull that off. What do you think would happen if you *planned* to follow the truth rule in this case?

A I'm not sure! My cognitive state isn't perfectly sensitive to the position of the pointer, but of course it also isn't totally insensitive. For example, if the pointer is at 12, I can tell that it isn't at 6. In fact, I'm sure that I wouldn't be off by more than one. If I planned to follow the truth rule, I would have to take a guess at the state of the clock. If the pointer is really at n, I'm not sure what I would guess, but I am sure that my guess would be $n-1$, n, or $n+1$ (modulo 12). And in fact, I assign each of the three possibilities equal credence. (See Figure 2.) So one important difference between this kind of case and the nap situation we were discussing earlier is that my guessing is *indeterministic*, by the lights of my prior credences.[6]

B Well, planning to follow the truth rule doesn't look like a good idea: for example, you might well end up being sure the pointer is at 2 when it's really at 1. But here's a natural idea. The margin for error in your guesses is one step of the clock: you can tell where the pointer is, to within one step. So it's natural to say that your *evidence*, when you look at the clock and you're in state n, is the set of states

$$E(n) = \left\{ n' \;\middle|\; n-1 \leq n' \leq n+1 \right\}$$

5. Gallow (2021) provides a closely related analysis of this case, based on ideas that overlap with the themes developed in this paper. We discuss the relationship between our framework and Gallow's in Section V.

6. While this story is obviously simplified, there is empirical evidence that people do make guesses which are implicitly sampled from an underlying probability distribution (Vul and Pashler 2008).

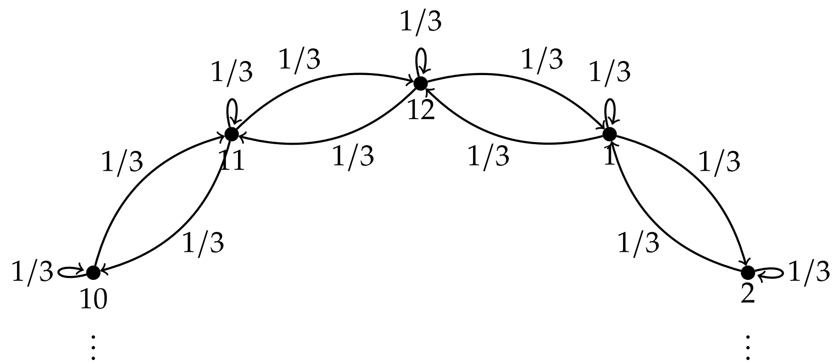

Figure 2: The stochastic guess graph for the clock case. An arrow from s to s' with label p means that in state s one guesses state s' with probability p.

(compare Williamson 2011).

A So should I just plan to proportion my belief to this evidence, like before, by conditionalizing on $E(n)$?

B Let's check. Earlier we mentioned Greaves and Wallace's result: in any case where there is a function E that takes each state s to a set of states $E(s)$, and these sets of states form a partition, then among those doxastic rules which are a function of your evidence, the optimal rule to follow is to conditionalize on $E(s)$ in each state s. But this isn't a case like that: in particular, the sets $E(n)$ are not a partition. For example, the distinct sets of states $E(10) = \{9, 10, 11\}$ and $E(11) = \{10, 11, 12\}$ have states in common.

A I see. That means that if $E(n)$ represents my *evidence* in each state n, then my evidence is not *transparent*: my evidence does not settle what evidence I have. My evidence in state n is $E(n)$, and this set includes states like $n + 1$ in which my evidence is $E(n + 1)$, which is different from $E(n)$. Intuitively speaking, if that's my evidence, then I can't tell what my evidence is.

B That's right.

A So how do things go in this kind of case?

B Schoenfield (2017) shows that in cases like this, conditionalizing on $E(n)$ when you are in state n is *not* the best rule to follow, out of those rules that are a function of your evidence.[7] Rather, out of those rules, the best one to follow is what she calls "conditionalization*": in each state n, conditionalize on the proposition "my evidence is $E(n)$". That is, conditionalization* says to conditionalize your priors on the set of states

$$E^*(n) = \{n' \mid E(n') = E(n)\}$$

A In my clock case, no two states have precisely the same evidence proposition. So for each state n, we have $E^*(n) = \{n\}$. That means that conditionalization* coincides with the truth rule in my case.

B Right.

A Okay, I know that the truth rule is pretty great as far as expected accuracy goes, but it still isn't a rule I can follow! As we discussed before, the expected accuracy of *planning* to follow this rule is dramatically worse than the expected accuracy of actually *following* it.

I see that we have to distinguish two different things. There are rules which are a function of my evidence. And there are rules which are *followable*, in the sense that if I plan to follow the rule, then I can be sure I will succeed in having the attitude the rule prescribes. When evidence is partitional, and I can be sure the evidence determines my attitude, these two notions coincide. But in the clock case, conditionalization* is a function of my evidence, but it is not followable.

Come to think of it, in this situation the ordinary conditionalization rule is the same way: it's a function of my evidence, but I can't follow it any more reliably than the truth rule or conditionalization*. If I plan to follow it, I will probably fail! If I plan to conditionalize on $E(n)$, I am just as likely to conditionalize on $E(n-1)$ or $E(n+1)$ as I am to do what the rule actually tells me to do.

B That's true. But in fact, if your cognitive state is sensitive to the clock in the way you have said, then only *trivial* rules are followable in your sense. Whatever a rule says to do in state n, you might instead do that in the state $n-1$ or $n+1$. So if a rule is to be followable, it must say to do the same thing in all three of those states. And so (by induction) the rule must be trivial, in the sense that it prescribes the very same attitude no matter what state you are in. Restricting your plans to followable

7. That is, those rules f such that for all states s and s', if $E(s) = E(s')$ then $f(s) = f(s')$.

rules does not seem like a good option here: that would keep you from updating at all (compare Williamson 2008; Srinivasan 2015).

A Very well then: if I'm going to make any plans at all, I'd better not exclude rules that aren't followable. I won't rule out the conditionalization rule—and I also won't rule out the truth rule, or conditionalization*, which amounts to the same thing in this case. None of these rules is followable: planning to follow any of these rules is unlikely to turn out precisely the way the rule prescribes. But I can still consider how good each of them is *as a plan*. I just have to take into account the different ways things might go if I plan to follow the rule, whether or not I succeed in following it in the end.[8]

So which rule makes the best plan?

B In this case, like before, we are assuming that if you plan to follow a rule f (which, as we said, is a function from states to credence functions defined on states), then you will end up having the credence $f(s)$, where s is the state that you *guess* that you're in. Unlike before, though, your guess is not simply determined by the state you are *really* in. It's stochastic. But we can use your prior credences about your guessing to calculate the expected accuracy for any rule you might plan to follow. Give me a minute to work this out . . . (See Appendix C for details.)

Okay, I've got it. In this case the rule that uniquely maximizes expected accuracy really is conditionalization on $E(n)$ in each state n.

A Neat! So it looks like *planning* to proportion my belief to the evidence is vindicated! In particular, this is better (in terms of expected accuracy) than planning to proportion my belief to my evidence* along the lines of Schoenfield (2017), or planning to follow the truth rule.

Or maybe this is a better way to put it. What we have vindicated is the idea that, given these assumptions about my powers of discrimination, the set of three states $E(n) = \{n-1,\ n,\ n+1\}$ really specifies something I ought to plan to proportion my belief to. In that sense, we could say this family of propositions plays the *role* of evidence for me. To get to that conclusion, we didn't actually need to assume that the function E played some special role in *constraining* my credences. Rather, it arose from a description of my psychological propensities—the way in which my cognitive state is sensitive to the state of the clock. The set $E(n)$

8. In fact, this is precisely what Schoenfield (2018) recommends in a different context.

includes the states that are "close" to n, in the sense that they are states I could easily mistake for n, if I guess which state I'm in.

B That's how it worked out in this case. But maybe we shouldn't generalize too hastily.

III An Old Friend

A Here's another situation I'm in surprisingly often (Lasonen-Aarnio 2015; Salow 2018). I'm waiting at a cafe for an old friend who I haven't seen for a long time. I think the next person who walks in might be my friend, or they might be a similar-looking stranger. Let's say both possibilities are equally likely. If my friend comes through the door, I feel sure that it's my friend—I'd know that face anywhere. But if a similar-looking stranger comes through the door, I can't tell whether or not it's my friend.

B Why not reason like this? (Salow 2018, pp. 702–707) If you know this is how it goes for you, then if you introspect and recognize that you *don't* feel sure whether it's your friend, you can then deduce that it must really be a stranger. So whether you feel sure or not, you really can tell whether it's your friend.

A That might work for some people, but I've tried it, and when I'm in this situation I'm just not good enough at introspection! When I feel *sure* it's my friend, I can recognize my feeling of confidence just fine: I'm in no danger of mistaking my sure feeling for the more ambiguous feeling. But when it's a stranger and I have an ambiguous feeling, my attempts to identify my own feeling are equally ambiguous. I have the same problem recognizing my own feelings as I have recognizing people!

B And I suppose it's no help to introspect and check whether you feel *sure* that you feel sure . . .

A Right! When it feels ambiguous to me whether it's my friend, it also feels ambiguous to me whether it feels ambiguous, and my attempts to identify that ambiguous ambiguity are *also* ambiguous!

B So how have you dealt with these situations so far?

A Here's something some people have suggested (Lasonen-Aarnio 2015; Williamson 2000). When I see my friend, in fact I *know* it's my friend,

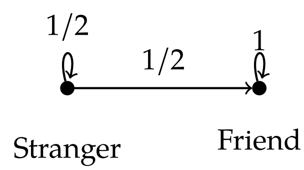

Figure 3: The guess graph for meeting an old friend.

and I have *evidence* that entails it's my friend. When I see a stranger, I don't know whether it's my friend or a stranger, and my evidence leaves both possibilities open. It's natural to think I should plan to proportion my belief to the evidence, like in the other cases we've been talking about. So here's the rule I plan to follow. If it's my friend who comes in, I'll update on my evidence, and thus be sure it's my friend. If it isn't my friend, I don't get any relevant evidence, and so I'll continue to have credence one-half that it's my friend.

B How's that been going?

A Not well! Here's what happens when I plan to follow that rule. In the good case when it *is* my friend, I always wind up having credence one that it's my friend, and that's nice. But in the bad case when it *isn't* my friend, and my phenomenology is ambiguous, I just guess.

B I suppose that you can't *tell* that you're guessing, so you can't use that fact to figure out that it's really a stranger, either.

A Afraid not. So in the bad case, I can go either way. I'm equally likely to end up doing what I planned for the good case when it *is* my friend as I am to do what I planned to do in the bad case when it *isn't* my friend. (See Figure 3.)

With this plan, in the good case, I always end up certain I'm in the good case. But in the bad case, half the time I wind up being wrongly certain I'm in the good case. Now that I'm thinking about optimizing expected accuracy, that seems like it may not be a very good trade-off. What do you think?

B Let me think about it for a minute ... Ok, here's a rule. When it's a stranger, be *certain* that it's a stranger; when it's your friend, have credence 2/3 that it's your friend. Planning to follow *this* rule maximizes expected accuracy.

A That's weird, isn't it? Given my propensities to make mistakes, it would be natural to say that my *evidence* is *stronger* in the good case, when I can easily recognize my friend, than it is in the bad case, when I can't tell what's going on. But while this rule does make different recommendations for the two cases (even though I can't always tell them apart), it recommends *less* confidence when I am in the good case than when I am in the bad case!

There is an asymmetry in my propensity to make mistakes in the good case and the bad case. I won't easily mistake my friend for a stranger, but I might easily mistake a stranger for my friend. We could say that the *mistakability* relation is asymmetric: the good case is not mistakable for the bad case, but the bad case is mistakable for the good case. It was very natural to think that there is a relation of "evidential accessibility" that goes the same way. What I mean is, it is natural to think that when I am in the good case, I should *rule out* the bad case (in the sense of becoming certain that it does not obtain); but in the bad case I should not rule out the good case. But in fact, the optimal rule says exactly the opposite: in the *bad* case, rule out the *good* case, but not vice versa.

B Right, but that's not the only difference. The best plan isn't to *conditionalize* on the set of states that you don't rule out. In the good case, you don't get to rule out any states. But your prior credence that it's your friend is 1/2, while the posterior credence you plan to have in that case is 2/3. The best rule to plan to follow in this case isn't to conditionalize on any set of states. It's a different kind of thing.

A Can you explain why that's the best plan?

B Sure. Here's the reasoning.[9] There are three relevant possibilities:

1. You see your friend, and you guess it's your friend.
2. You see a stranger, and you guess it's your friend.
3. You see a stranger, and you guess it's a stranger.

You assign case 1 prior credence 1/2 and cases 2 and 3 prior credence 1/4 each. If you plan to follow a certain rule f, then you will end up having the credences $f(\text{Friend})$ in cases 1 and 2, and the credences $f(\text{Stranger})$ in case 3. There are a few fiddly details here about scoring rules (see Appendix C) but things work out basically the way you would

9. See Theorem 3 in Appendix C for a generalization.

expect: since case 1, where you see your friend, is twice as probable as case 2, where you don't, the optimal credence function to have for those cases gives twice as much probability to seeing your friend as to seeing a stranger. Meanwhile, in case 3 it's best to be sure it's a stranger, since that's the only way case 3 can arise.

A So this case goes differently from the clock case we talked about before. In that case, the best rule for me to plan to follow involved becoming sure of something: that the clock pointer is at one of those positions that could easily be mistaken for the actual position. Furthermore, my best plan was to conditionalize on this proposition. Here, though, while my best plan sometimes involves becoming sure of something—that it's a stranger—this certainty doesn't line up with the strength of my epistemic position in the way I expected. Furthermore, the rule that is best for me to plan to follow sometimes prescribes changing my credences without becoming certain in anything new. So in some cases, this optimal rule does not say to conditionalize on any proposition.

IV The Music Competition

A Here's another situation I'm in. It feels quite different to me, and I'm curious what you'll make of it.

My little brother has his heart set on attending Julliard. He plays the oboe, and is currently one of two finalists in a music competition. I've been told that applicants like him who win this sort of competition are admitted to Julliard around 90% of the time, and that applicants like him who haven't won this sort of competition are admitted around 10% of the time. My brother says that he and his competitor are very evenly matched, so it's about fifty-fifty whether he wins. As you'd expect, right now my credence that he'll get into Julliard is 0.5. But I have a hard time figuring out how to react when my brother lets me know whether he won the competition.

B Why are you finding it hard? This one seems pretty straightforward: structurally, it sounds just like the nap case from earlier.

A Here's the thing: I'm very emotionally invested in my brother's academic wellbeing, and that makes my thoughts about him less stable than my other thoughts. Whatever I try to do, I'm always either *overconfident*, becoming certain that my brother will get in or certain that he won't, or else I'm *underconfident*, staying at my original credence of 0.5. I can't seem to sustain any other degree of belief on this question besides 0, 0.5, or 1. I hear that this kind of thing is not that uncommon.[10]

B What determines which state of belief you're in?

A It's like this. If I aim to be confident that my brother will be admitted, and go for credence 0.8 or above, I wind up certain that he'll get in. If I aim to be confident that my brother won't get admitted and go for credence 0.2 or below, I wind up certain that he won't get in. And if I aim to have middling credence between 0.2 and 0.8 then I wind up at 0.5. As you might imagine, my credence about my brother's academic trajectory has jumped around dramatically! I'm like Plato's charioteer, trying to guide winged horses that are irrationally pulling me off in different directions.

B So you'd better make a plan that compensates for your dispositions toward overconfidence or underconfidence. Here's how it's going to go. Your dispositions are coarse-grained enough that it doesn't really matter what precise credence you aim for—in most cases you won't hit it anyway. All you get to choose is, in each situation, whether to go High, where you end up at 1, Middling, ending up at 0.5, or Low, ending up at 0. So consider the trade-off between these two options:

1. A 90% chance of credence 1 in a truth, and a 10% chance of credence 1 in a falsehood.
2. A 50% chance of credence 0.5 in a truth, and a 50% chance of credence 0.5 in a falsehood.

A I'm not sure which one is better. Which option has higher expected accuracy?

B That depends! There are many different *scoring rules*, which are different ways of measuring closeness to the truth. For the other problems we've talked about so far, it hasn't mattered which scoring rule you use. In all of those cases, all (proper) scoring rules have given the same verdicts about accuracy-optimizing plans.[11] But in this case, things are messier. Different scoring rules give different answers.

10. For more realistic versions of such overconfidence effects and conservatism bias, see Fischhoff, Slovic, and Lichtenstein (1977) and Edwards (1982).

11. "Proper scoring rules" (see Appendix A) include all of the accuracy measures that are widely defended in the literature, such as the Brier score or logarithmic score (though technically this is not a real-valued scoring rule, since its value can be negative infinity).

For example, the *logarithmic score* says that it's *really* bad to be certain of a falsehood.[12] If that's your scoring rule, then you should avoid the risk of extremely costly overconfidence, and just stick to some Middling credence, no matter what you find out about your brother's competition—it isn't worth the risk. For a different example, the *Brier score* says that the *inaccuracy* of having credence p in a falsehood is p^2, and the inaccuracy of having credence p in a truth is $(1-p)^2$. (A credence is more *accurate* when its inaccuracy is closer to zero.) According to this scoring rule, having a Middling credence has inaccuracy 0.25 whatever happens to your brother. Meanwhile, if your brother wins then a High credence has a 10% chance of inaccuracy 1 and a 90% chance of inaccuracy 0, so the expected inaccuracy is 0.1. So you'd be more accurate on balance by going High and taking the risk.

Your situation is pretty much the way William James famously put it (1896):

> [B]y choosing between [different scoring rules] we may end by coloring differently our whole intellectual life. We may regard the chase for truth as paramount, and the avoidance of error as secondary; or we may, on the other hand, treat the avoidance of error as more imperative, and let truth take its chance.

But the other cases we discussed before weren't "Jamesian" in this way: all proper scoring rules agreed on what credal plan they recommended.

Still, things aren't completely up in the air even now. Every reasonable scoring rule is going to give rise to the same *kind* of plan for cases like this one. There will be some pair of threshold probabilities t_0 and t_1. What the thresholds are will be determined by your scoring rule. (Note these won't generally be the same as the thresholds that figured in your dispositions, which were 0.2 and 0.8.) The optimal rule will say something of this form: if you are in a state s such that the conditional probability of p given $E(s)$ is above t_1, go High; if it is below t_0, go Low; and otherwise go Middling.

Even more complicated situations can lead to even greater divergence in the recommendations of apparently sensible scoring rules. But so it goes.

12. The logarithmic inaccuracy of credence p in a truth is $-\log p$, and for a falsehood the inaccuracy is $-\log(1-p)$. This diverges to infinity for credence 0 in a truth or credence 1 in a falsehood.

V Taking Stock

Let's draw some lessons from the preceding scenes.

We've been thinking about *updating*: changing your degrees of belief in response to changes in your situation. Which ways of updating are epistemically upstanding? We have taken up two themes from recent epistemology. First, our standard of evaluation is *accuracy*—closeness to the truth. Second, our object of evaluation is a *doxastic rule*: a prescription for what credences to adopt in various circumstances. But (as others have noted in other contexts) *planning* to follow a rule and successfully *following* it are two different concerns. Sometimes it is better to aim for a lower target than to aim high and risk missing badly. So the central ideas we have been exploring are that methods for epistemic updating are described by doxastic rules, and that such a method is better insofar as planning to follow the rule is more accurate in expectation, taking into account one's fallibility at following plans.

Gallow (2021) pursues some closely related ideas. Where we consider doxastic rules, Gallow considers "learning dispositions": these are modeled as functions that take each proposition which might be your *total evidence* to a credal state. Where we consider the possibility of imperfectly following one's planned rule, Gallow considers the possibility that one's "learning disposition may misfire". There is a stochastic relationship between which proposition actually *is* your total evidence and which proposition is *updated* on (see Gallow's sec. 2 and sec. 4). Gallow then considers the question of which dispositions optimize expected accuracy under these conditions—this is directly analogous to our main line of inquiry, cast in his alternative framework. Unsurprisingly, given these shared starting points, he recommends an updating rule that coincides with the ones we discuss in many cases. (His central case is a version of the unmarked clock we discuss in Section II. Compare, in particular, our Theorem 3 with Gallow's proposition 1 in appendix B.)

Still, there are many differences between Gallow's framework and ours. First, Gallow's "misfiring" always consists in updating in the right way on the wrong evidence. The credences you end up with are the ones your "learning disposition" specifies for *some* possible evidence proposition—it just may not be the evidence you really have. But as we considered in Section IV, your credences might be imperfectly sensitive to the state of the world in many different ways besides this.

More generally, the most important point of contrast concerns the role of *evidence*. Gallow treats evidence as an "input": the propositions that one might receive as evidence are taken as given, and they provide the key structure that guides his analysis. But a very different picture emerges from the preceding scenes. When "evidence" appears in our analysis at all, it is not as an input, but as an output.

It is standardly thought that evidence constrains updating. What does this amount to? Here was the picture from Greaves and Wallace (2006) and Schoenfield (2017). First, as a simplifying assumption, we factor out *states* from *plans*. This is to suppose that the subject matter under consideration is not so fine-grained that it settles which doxastic rule you plan to follow. (Otherwise the prior probability of being in one state or another would vary with your plan, rather than being held fixed as Greaves and Wallace assume.) Second, there is some feature of your state that, together with your planned rule, determines your credences. We can represent this feature by a partition $\mathscr{E}$ of the set of states. The only doxastic rules that are "available" (to use Greaves and Wallace's term) are those which are determined by $\mathscr{E}$—that is, the available rules are functions from states to credence functions which are constant on each $\mathscr{E}$-cell. This has the upshot that, among the available rules, the one that maximizes expected accuracy is *conditionalizing* on whichever $\mathscr{E}$-cell is true.

The partition $\mathscr{E}$ plays the role of *evidence* in two ways. First, it circumscribes which updating rules count as "available": those which are a function of the evidence. Second, among these available rules, the one that maximizes expected accuracy tells you to be certain of a proposition in this partition.[13]

We have relaxed the assumptions of Greaves and Wallace's framework. They only considered certain kinds of change of mind to be available as plans at all—those which are determined by an antecedently given partition $\mathscr{E}$. So what evidence you might have imposes a constraint on what doxastic plans you can have. For example, you can't take up the truth rule as your updating policy—that rule is not "available". Meanwhile, if you adopt one of the *available* rules, the framework takes for granted that you will succeed in conforming to it: the possibility of ending up with credences other than those you planned to have is not even considered.

Cases like the unmarked clock suggest that this is not a realistic picture. In such cases your change of mind is predictably sensitive to features of the world, but there is no feature of the world to which it is *perfectly* sensitive.

13. Note, though, that calling the true $\mathscr{E}$-cell (as we have characterized it) your "evidence" conflicts with how Schoenfield (2017) describes things: this proposition corresponds to what she calls E^*, rather than E. Your $\mathscr{E}$-cell, which determines your posterior credences, is not your evidence proposition itself, but rather the proposition "my evidence is E."

Then there may be no non-trivial rule that you can plan to follow with assurance of perfect success. Anti-luminosity considerations more generally suggest that such predicaments are pervasive: our access to the world, and our ability to guide our own minds in response to our environment, is inexact.

The picture we have explored instead is that there is *no* prior constraint on which doxastic plans you are allowed to have. You can plan to follow the truth rule, or whatever—but there is no guarantee that adopting such a plan will have the intended results. In this sense, then, this is a picture of updating without evidence. There are no criteria that rule in some "available" plans which are sensitive only to "accessible" features of the world, and rules out others that appeal to "inaccessible" features. Rather, there are many ways we can come close to fulfilling our doxastic plans or departing from them—many varieties of doxastic triumph or tragedy. There also need not be anything with the second feature characteristic of evidence in the setting of classical conditionalization: the rule that maximizes expected accuracy as a plan may not recommend certainty in any non-trivial set of states. But even in these conditions, where there is no natural sense in which you are gaining classical "evidence," we can still evaluate how well an updating rule conduces to forming accurate beliefs—as a fallible plan.

In order to carry such evaluations out in any detail, we begin—as we did in the preceding scenes—by describing prior probabilities about what credences an agent will end up having in various circumstances if they adopt a certain doxastic rule as their plan. In Sections I to III, we considered an agent whose psychological propensities had a simple structure. What determined their credence was a *guess* about the state of the world, which was imperfectly correlated with the actual state of world. If the agent guesses they are in state s, they adopt those credences that they planned to adopt in s. In Section I, the agent's guess was determined by the actual state of the world, while in Sections II and III the agent's guess was only stochastically related to it. We can represent situations like these using what we call a *guess model*. Such a model is given by a probability distribution over pairs of states, one component representing the actual state, and the other the agent's guess. For such simple cases, it is possible to exactly characterize the rules which maximize expected accuracy as plans: this is worked out in Appendices B and C.

In situations apt to be represented by these simple guess models, we can say precisely when a certain partition of states $\mathscr{E}$ *does* play the classical evidence role—that is, when the best plan is to conditionalize on whichever cell in $\mathscr{E}$ is true. (This is made precise in Appendix D.) This holds if and only if the following two conditions hold. First, your beliefs are perfectly sensitive to this subject matter: your guess state is certain to be in the same $\mathscr{E}$-cell as the actual state. Second, your beliefs are perfectly insensitive to anything else: the probability of guessing any state, given that you are actually in a certain state s, is the same for each state s within the same $\mathscr{E}$-cell. These conditions correspond to two idealizations that commonly underwrite traditional notions of evidence. But these conditions are not realistic in general; when they fail, nothing plays this classical role.

Still, one might suspect that the evidence has simply been hidden elsewhere in the model. In simple guess models, we spoke as if your *guess* plus your plan determines your credences. And the optimal rule to plan to follow effectively says to conditionalize on "my guess is s" when you are in state s. (We discussed this informally in Section III, and we spell it out carefully in Appendix C.) So isn't the fact about what your guess is playing the standard evidence role? In particular, doesn't this have the upshot that you should plan to be certain of what your guess is?

First, that isn't quite what a guess model actually says. A doxastic rule only represents your credences about what *state* you are in, a guess model only says anything about the evolution of these credences, and the accuracy-based evaluations we glean from the model only apply to rules for updating these credences. But the state you are in does not include your guess.[14] So in fact, instead of prescribing certainty about what state you have guessed that you are in, the accuracy evaluations we can derive from a guess model are simply silent on this question.

Indeed, we need not think of the *guess* in the model as representing anything psychologically real at all, which you might have some opinion about. Rather, it can be thought of as a device for taking a plan and a prior and producing a probability distribution over pairs (s, P) of a state and credence function defined on states.[15] In the "guess anti-realist" picture,

14. It is important that we are still supposing, with Greaves and Wallace, that the "state" is not a fine-grained possible world that settles every question; it is coarse-grained, and in particular it does not settle your state of mind. But in general, you can of course have credences in propositions about your own credences; and of course, what you plan to believe about your own beliefs is not independent of what you believe. This makes things hard in ways we do not address here (see Greaves 2013). For example, it might look like it's worth taking a small accuracy hit by having positive credence that $2 + 2 = 5$, if that lets you, for example, improve your accuracy score for your credence in the proposition *someone has positive credence that* $2 + 2 = 5$.

15. The resulting picture is similar to Schwarz 2018's "imaginary foundations": one conditionalizes on some "imaginary" proposition, which does not represent some fact about one's environment or about one's psychology, but rather serves as a mere index.

you have genuine prior credences about how your doxastic plan might turn out—how your possible beliefs are correlated with your possible state—and this ends up leading to the recommendation that you should plan to update your credences *as if* you were conditionalizing on some further subject matter. But that subject matter—your "guess"—isn't one you really had any prior opinion on, and it isn't one you need to have any posterior opinion on, either. It's a mere calculation aid.

Guess models capture two idealizing assumptions. (1) You are sure to do something that you planned to do. That is, it is certain that your credal state will be one that fits with your plan for *some* possible state, even if not the state you are actually in. (2) The kinds of mistakes you are liable to make do not depend on which doxastic rule you plan to follow. Condition (1) guarantees that for each rule, the probability of having certain credences in a given state when you plan to follow that rule can be generated by an *error model*, which tells you how likely you are to follow any given state t's recommendations when you are actually in some state s. What condition (2) says is that there is a single error model that generates these probabilities for every rule. But these conditions on what kinds of credences you anticipate having given various doxastic plans do not require that you are really going through some psychological process of guessing a state at all.

The idealizing assumptions incorporated in guess models do not always hold. There are other ways in which one's beliefs might depend on one's doxastic plan without being perfectly reliable. The music competition example in Section IV is like this: in particular, there the agent might end up with credences they did not plan to have in any circumstances. The model we described for a case like this did not include a "guess": instead, we directly described a probabilistic relationship between an agent's planned update rule, the state of the world, and their eventual credences. In Appendix A we call representations of these more general relationships *planning models*. Planning models without any guess structure do not make it tempting to think that there is really some hidden fact playing the evidence role—neither an input parameter that is fed to the doxastic plan, nor some proposition the optimal plan tells you to be sure of.

In these more general planning models, the doxastic rule itself may not really play a crucial role at all: the key thing that we end up evaluating is just what dispositions to form beliefs result from planning to follow that rule. What really matters is not what the rule *says* to believe, but how planning to follow the rule might affect your beliefs, which may in principle come very far apart from what the rule officially says. The role the rule still plays is as a *parameter* for a comparison class of potential doxastic dispositions.

In this way of thinking about it, epistemic updating comes down to a question of robot design.[16] The picture is that you have some stock of cognitive widgets that can be assembled to produce various doxastic dispositions, and the question is just how to put them together to build the best believing machine. What we have explored here is how this kind of project might go, by examining a few different kits of simple, tractable widgets—in particular, noting some situations in which the best robots we can build are conditionalizers, and others where they are not. This project is in the broad spirit of Quine's naturalized epistemology: as he put it, "For me normative epistemology is a branch of engineering. It is the technology of truth-seeking" (Quine 1986, pp. 663-665).[17]

Once upon a time, Bayesians thought we had a "cognitive home" (Williamson 1996): a domain of transparently recognizable and unmistakable facts to which our beliefs were sensitive—perhaps facts about our own sense data. But there are good reasons to think that we have no such cognitive home (among others, see Jeffrey 1983; Williamson 2000). We can make mistakes about our own experiences, and even about how our own experiences seem to us. Here we have been exploring cognitive homelessness.

Some philosophers have turned to other theories of evidence: for example, Jeffrey (1983) conceived of evidence as imposing "arational" doxastic constraints. Experience knocks certain degrees of belief around; your job is to do the best you can to keep everything coherent (compare also van Fraassen 1989, ch. 13). But that's not what's happening here, either. We impose no constraints on what doxastic plans you may form. You are not required to be certain of anything in particular; no more are you required to be uncertain to any particular degree in anything in particular. The contingent facts your beliefs are responding to are all the facts in the world, though the degree to which you can successfully hope to respond to them varies quite a bit—you'll do better if you try to follow traffic signs than if

16. Compare Carnap (in Carnap and Jeffrey 1971, p. 17):

> Thinking about the design of a robot might help us in finding rules of rationality. Once found, these rules can be applied not only in the construction of a robot but also in advising human beings in their effort to make their decisions as rational as their limited abilities permit.

Carnap quaintly thought of his "robot" as an ideal reasoner, serving as a foil for humans' "limited abilities". In contrast, our "robots" are also very limited, but in predictable ways that we can study and accommodate.

17. Compare Pollock (1986)'s notion of *procedural justification*, as well as approaches that give a central place to cognitive *dispositions* (Lasonen-Aarnio, n.d.) or *habits* (Hawthorne and Srinivasan 2013).

you try to follow fluctuations in cosmic background radiation. The overall story here is not about optimizing your belief state under certain constraints imposed by the "evidence" (whatever that may be)—but about optimizing a belief-plan under the psychological constraints given by your limited ability to follow belief-plans. Instead of trying to come up with a new theory of what evidence is, we have advanced a theory that gives evidence no central role at all.

A Preliminaries

Planning to follow an updating rule does not generally guarantee that you will successfully conform to it, but there is *some* connection between your planned rule and your future credences. In Sections I to IV we considered various ways this connection might go, proceeding from stronger to weaker constraints. In Section I, an agent's credences were determined by their *guess*, which was deterministically settled by the state of the world. In Sections II and III, we generalized this to consider indeterministic connections between the state of the world and an agent's guess. In Section IV, we dispensed with guesses and associated each plan with predicted credences more directly. These appendices will make each of these approaches more precise, and spell out and justify the main claims in the text about expected-accuracy-maximizing plans.

For a set X, a *discrete probability distribution* on X is a function from X to $[0, 1]$ whose values sum to one. (We will focus on the discrete case just to keep the math simple, avoiding integrals.) Let $\mathbb{P}X$ be the set of all discrete probability distributions on X.

Let S be a set of *states*. For short we'll call elements of $\mathbb{P}S$ *opinions*. An *accuracy scoring rule* is a function $A : S \times \mathbb{P}S \to \mathbb{R}$ that takes each pair of a state and an opinion to a real number—intuitively, the accuracy of credal state Q if s turns out to be actual. Greater numbers represent more accurate opinions.[18]

For an opinion $Q \in \mathbb{P}S$, let $EA(P, Q)$ be the *expected* accuracy score

$$EA(P, Q) := \sum_{s \in S} P(s)\, A(s, Q)$$

An opinion $P \in \mathbb{P}S$ is *self-recommending* with respect to a scoring rule A iff

$$EA(P, P) > EA(P, Q) \quad \text{for all } Q \neq P \in \mathbb{P}S$$

18. Greaves and Wallace (2006) call these *epistemic utility functions.* In other contexts sometimes the sign is reversed, and one instead considers *inaccuracy* scores.

A scoring rule A is *strictly proper* iff every opinion in $\mathbb{P}S$ is self-recommending.[19]

A *(doxastic) rule* is a function $f : S \to \mathbb{P}S$ from states to opinions.

Our most general way of describing the likely consequences of planning to follow a rule is a *planning model*. A planning model is a function $M_{(-)}$ that takes each doxastic rule f to a joint probability distribution $M_f \in \mathbb{P}(S \times \mathbb{P}S)$. Intuitively $M_f(s, Q)$ represents the probability of being in state s and having opinion Q, if you plan to follow the doxastic rule f.[20]

For such a joint distribution $P \in \mathbb{P}(S \times \mathbb{P}S)$, the *expected accuracy* of P is

$$EA(P) \coloneqq \sum_{s \in S, Q \in \mathbb{P}S} P(s, Q)\, A(s, Q)$$

(This definition generalizes the definition of relative expected accuracy for two opinions given above: for $P, Q \in \mathbb{P}S$, $EA(P, Q) = EA(R)$ where $R(s, Q') = P(s)$ for $Q' = Q$ and 0 otherwise.) A rule f is *optimal* with respect to a planning model $M_{(-)}$ iff for any rule f',

$$EA(M_f) \geq EA(M_{f'})$$

The main results below are applications of Greaves and Wallace's Theorem (2006); so we begin by restating that result in the terms of this appendix.

Besides the value of *planning* to follow a doxastic rule, we can also consider the value of actually *following* a rule. For a rule $f : S \to \mathbb{P}S$ and a probability distribution $P \in \mathbb{P}S$, the expected accuracy of following f is

$$EA(P, f) \coloneqq \sum_{s \in S} P(s)\, A(s, f(s))$$

(This definition is also a special case of the definition of expected accuracy for a joint distribution given above: $EA(P, f) = EA(R)$ where $R(s, Q) = P(s)$ if $f(s) = Q$ and 0 otherwise.)

Let E be a partition of S. For any state $s \in S$, let $E(s)$ be the unique E-cell that contains s. A rule $f : S \to \mathbb{P}S$ is *E-determined* iff for all $s, s' \in S$, $f(s) = f(s')$ whenever $E(s) = E(s')$.

For a probability distribution $P \in \mathbb{P}S$ and a partition E, the *conditionalization rule* $\text{cond}_{P,E} : S \to \mathbb{P}S$ is the function that takes each state s to the distribution

$$\text{cond}_{P,E}(s) = P\big(- \bigm| E(s)\big)$$

19. In Greaves and Wallace's terminology, *everywhere strongly stable*.

20. Compare Pettigrew's (2020) closely related idea of a *stochastic update rule*.

Theorem 1 (Greaves and Wallace 2006). *Let $P \in \mathbb{P}S$ be a probability distribution, let E be a partition, and let $A : S \times \mathbb{P}S \to \mathbb{R}$ be a strictly proper scoring rule. Suppose that* $\mathrm{cond}_{P,E}(s)$ *is self-recommending for each $s \in S$. Then*

$$EA(P, \mathrm{cond}_{P,E}) > EA(P, f)$$

for every E-determined rule f distinct from $\mathrm{cond}_{P,E}$.

It's worth re-emphasizing that this theorem is about the optimality of actually *following* a certain rule, rather than the optimality of *planning* to follow a rule when one is not sure to succeed.

B Deterministic guess models

In principle, a planning model can represent an arbitary relationship between a rule and what one expects to ensue from planning to follow it. To make progress, we considered a tractable class of planning models that we called *guess* models. In Section I we began with *deterministic* guess models, where the actual state determines which state you guess you are in. The intuitive picture is that when you plan to follow a rule f, the following happens. First, you guess which state you are actually in: if your actual state is s, call your guess $g(s)$. Second, you adopt the attitude $f(g(s))$.

Definition 1. A *deterministic guess model* is a planning model $M_{(-)}$ for which there exists a *prior* probability distribution $P \in \mathbb{P}S$, and a function $g : S \to S$ that we call a *guess function*, such that for each rule $f : S \to \mathbb{P}S$,

$$M_f(s, Q) = \begin{cases} P(s) & \text{if } f(g(s)) = Q \\ 0 & \text{otherwise} \end{cases}$$

For simplicity, we will only consider *regular* prior probability distributions such that $P(s) > 0$ for each state $s \in S$.

Theorem 2. *A doxastic rule f is optimal for the deterministic guess model corresponding to (regular) prior P and guess function g iff, for each state s in the range of g,*

$$f(s) = P\big(- \bigm| g^{-1}(s)\big) \quad \textit{for each state } s \in S$$

(Here $g^{-1}(s)$ is the preimage of s under g: that is, $g^{-1}(s) = \{s' \in S \mid g(s') = s\}$.)

Proof. According to a deterministic guess model, *planning* to follow the rule f is tantamount to actually *following* the rule $f \circ g$. That is, if you are in s, then you are sure to follow the recommendation that f gives for the state you guess, $g(s)$. In particular, the definitions tell us:

$$EA(M_f) = EA(P, f \circ g)$$

Let E be the partition induced by g: that is, for each state s,

$$E(s) = g^{-1}(g(s)) = \{s' \in S \mid g(s') = g(s)\}$$

For any rule f, the rule $f \circ g$ is E-determined. Thus Theorem 1 tells us that for any rules f and f', if

$$f \circ g = \mathrm{cond}_{P,E}$$
$$f' \circ g \neq \mathrm{cond}_{P,E}$$

then

$$EA(M_f) = EA(P, f \circ g) = EA(P, \mathrm{cond}_{P,E}) > EA(P, f' \circ g) = EA(M_{f'})$$

Thus a rule f is optimal iff for each state s,

$$f(g(s)) = \mathrm{cond}_{P,E}(s) = P\left(- \mid g^{-1}(g(s))\right) \qquad \square$$

C Stochastic guess models

In Sections II and III we considered situations where the agent's credences were still determined by a guess, but this guess was not *determined* by the actual state of the world (the position of the clock, the person entering the cafe) but was probabilistically related to the state.

Definition 2. A *stochastic guess model* is a planning model $M_{(-)}$ for which there exists a joint probability distribution $P \in \mathbb{P}(S \times S)$, the *guess distribution*, such that, for each rule $f : S \to \mathbb{P}S$, state $s \in S$, and opinion $Q \in \mathbb{P}S$,

$$M_f(s, Q) = P\{(s, t) \mid f(t) = Q\} = \sum_{t : f(t) = Q} P(s, t)$$

Intuitively, the guess distribution $P(s, t)$ represents the probability of being in a state s while guessing that you are in state t. In that case, if you planned to follow the rule f, the opinion you end up holding is $f(t)$. So the probability of having opinion Q in state s is the same as the probability of being in s and guessing some state t such that $f(t) = Q$.

For simplicity we will again restrict attention to models satisfying a regularity condition. Call a distribution $P \in \mathbb{P}(S \times S)$ *state-regular* iff for each state, the probability that you are in it is positive; that is, $P(\{s\} \times S) > 0$ for each state s.

Every deterministic guess model is also a stochastic guess model. If a planning model $M_{(-)}$ arises as a deterministic guess model from a prior P and guess function g, then it is also arises as a stochastic guess model from the guess distribution that assigns probability $P(s)$ to each pair $(s, g(s))$, and probability zero to all pairs not of this form.

Theorem 3. *For a stochastic guess model with (state-regular) guess distribution $P \in \mathbb{P}(S \times S)$, the uniquely optimal doxastic rule is the function $f : S \to \mathbb{P}S$ such that, for any state $t \in S$,*

$$f(t) = P\big((-, t) \,\big|\, E(t)\big)$$

where $E(t) = \big\{(s, t) \,\big|\, s \in S\big\}$.

In other words, the optimal rule recommends for the state t the probabilities that result from, first, *conditionalizing* the guess distribution $P \in \mathbb{P}(S \times S)$ on the proposition "my guess is t", and, second, *marginalizing* that result to ignore the guess component.

The proof is another application of Theorem 1. This time there are some extra complications, because we are moving between two different state spaces: the original state space S, and the enriched "guess space" $S \times S$. Since the accuracy scoring rule A we use to assess doxastic rules is only defined for the original state space, to apply Greaves and Wallace's Theorem we first need to construct an "enriched" scoring rule for the "guess space".

We can move back and forth between the basic state space S and the enriched guess space $S \times S$. For any opinion $Q \in \mathbb{P}S$ and any state $t \in S$, we can "lift" Q to an opinion defined on *pairs* of states:

$$Q^t(s, t') := \begin{cases} Q(s) & \text{if } t = t' \\ 0 & \text{otherwise} \end{cases}$$

Basically, this copies the distribution Q onto $E(t)$, which is a copy of S within $S \times S$. We can also use a state $t \in S$ to "lower" an opinion $Q \in \mathbb{P}(S \times S)$ to the opinion $Q(-, t) \in \mathbb{P}S$. Lifting an opinion and then lowering it again takes you back to the same opinion you started with: for any opinion $Q \in \mathbb{P}S$ and state $t \in S$,

$$Q^t(-, t) = Q$$

We can also lift a *rule* $f : S \to \mathbb{P}S$ to a rule defined for pairs of states $f^* : S \times S \to \mathbb{P}(S \times S)$, namely:

$$f^*(s, t) := f(t)^t$$

Given a pair of an actual state s and a guess t, the enriched rule is only sensitive to the guess, ignoring the actual state, and it only assigns non-zero probabilities to the pairs where the second coordinate matches the guess.

Theorem 3 can be restated more simply in these terms: what we want to show is that, for a guess distribution P, a rule f is optimal iff

$$f^* = \mathrm{cond}_{P,F}$$

where F is the partition of $S \times S$ consisting of each of the sets $E(t) = \{(s, t) \mid s \in S\}$.

Lemma 1. *For any scoring rule $A : S \times \mathbb{P}S \to \mathbb{R}$, there is a "lifted" scoring rule*

$$A^* : (S \times S) \times \mathbb{P}(S \times S) \to \mathbb{R}$$

such that for any stochastic guess model $M_{(-)}$ with guess distribution P, and any rule $f : S \to \mathbb{P}S$,

$$EA(M_f) = EA^*(P, f^*)$$

Furthermore, if A is strictly proper, then while A^* need not be strictly proper in general, each lifted *opinion Q^t is self-recommending with respect to A^*, for any $Q \in \mathbb{P}S$ and $t \in S$.*

Proof. For $(s, t) \in S \times S$ and $Q \in \mathbb{P}(S \times S)$, let

$$A^*\big((s, t), Q\big) := A\big(s, Q(-, t)\big)$$

For any plan $f : S \to \mathbb{P}S$, we have by the definitions

$$\begin{aligned} EA^*(P, f^*) &= \sum_{s,t} P(s, t)\, A^*\big((s, t), f(t)^t\big) \\ &= \sum_{s,t} P(s, t)\, A\big(s, f(t)^t(-, t)\big) \\ &= \sum_{s,t} P(s, t)\, A(s, f(t)) \qquad = EA(M_f) \end{aligned}$$

For the second part, let $R \in \mathbb{P}(S \times S)$ and suppose $Q^t \neq R$. Then, by the definitions (and in particular the fact that $Q^t(s, t') = 0$ for $t \neq t'$),

$$\begin{aligned} EA^*(Q^t, R) &= \sum_{s,t'} Q^t(s, t')\, A^*\big((s, t'), R\big) \\ &= \sum_{s} Q(s)\, A^*\big((s, t), R\big) \\ &= \sum_{s} Q(s)\, A\big(s, R(-, t)\big) \quad = EA(Q, R(-, t)) \end{aligned}$$

Thus, since Q is self-recommending with respect to A,

$$\begin{aligned} EA^*(Q^t, R) &= EA(Q, R(-, t)) \\ &< EA(Q, Q) = EA(Q, Q^t(-, t)) = EA^*(Q^t, Q^t) \end{aligned}$$

So Q^t is self-recommending with respect to A^*. □

Proof of Theorem 3. It suffices to show that a rule $f : S \to \mathbb{P}S$ is optimal iff

$$f^* = \mathrm{cond}_{P,F}$$

where P is the guess distribution, and F is the partition whose cells consist of sets of pairs (s, t) with the same guess coordinate t.

Let A^* be the lifted epistemic utility function as in Lemma 1. Then since $f^*(s, t) = f(t)^t$, the lemma tells us that each opinion in the range of $f^* = \mathrm{cond}_{P,F}$ is self-recommending with respect to A^*. We can also check that, for any rule $g : S \to \mathbb{P}S$ distinct from f, we also have $f^* \neq g^*$. It also follows from the definitions that g^* is F-determined. Thus Theorem 1 tells us that

$$EA(M_f) = EA^*(P, f^*) = EA^*(P, \mathrm{cond}_{P,F}) > EA^*(P, g^*) = EA(M_g) \qquad \square$$

We can also put the conclusion of Theorem 3 another way. First, we state a standard characterization of conditionalization.

Lemma 2. *For probability distributions $P, Q \in \mathbb{P}X$ and any set $E \subseteq X$ such that $P(E) > 0$, we have $Q = P(- \mid E)$ iff:*

(a) $Q(s) = 0$ *for* $s \in X \setminus E$, *and*

(b) $Q(s) : Q(s') = P(s) : P(s')$ *for* $s, s' \in E$.

(The ratio notation $x : y = x' : y'$ is a convenient alternative for $xy' = x'y$.)

Corollary 1. *A rule f is optimal for a stochastic guess model given by a guess distribution P iff for all states $s, s', t \in S$:*

$$f(t)(s) : f(t)(s') = P(s,t) : P(s',t)$$

Proof. Recall that Theorem 3 says that f is optimal iff for $t \in S$,

$$f(t)(s) = P\big((s,t) \,\big|\, E(t)\big) \quad \text{where } E(t) = \big\{(s,t) \,\big|\, s \in S\big\}$$

By Lemma 2 this holds iff

(a) $f(t)(s) = 0$ for $(s,t) \notin E(t)$

(b) $f(t)(s) : f(t)(s') = P(s,t) : P(s',t)$ for (s,t) and (s',t) in $E(t)$.

Since (s,t) is always in $E(t)$, (a) is vacuous and (b) simplifies to the statement above. □

D Conditions for Conditionalization

Our final result states the conditions under which the optimal rule for a stochastic guess model is to conditionalize on a given partition.

Theorem 4. *Consider a stochastic guess model with (state-regular) guess distribution $P \in \mathbb{P}(S \times S)$. Let $P_0 \in \mathbb{P}S$ be the marginalized distribution on states,*

$$P_0(s) = P\big\{(s,t) \,\big|\, t \in S\big\} = \sum_t P(s,t)$$

Let $c(s,t)$ be the conditional probability of guessing state t while in state s:

$$c(s,t) := \frac{P(s,t)}{P_0(s)}$$

Let E be a partition of S. Then the conditionalization rule $\mathrm{cond}_{P_0,E}$ is optimal iff both of the following two conditions hold:

(a) $c(s,t) = 0$ whenever s and t are not in the same cell of E; and

(b) $c(s,t) = c(s',t)$ whenever s, s', t are all in the same cell of E.

Proof. By Corollary 1, the conditionalization rule $\mathrm{cond}_{P_0,E}$ is optimal iff for all states $s, s', t \in S$,

$$\mathrm{cond}_{P_0,E}(t)(s) : \mathrm{cond}_{P_0,E}(t)(s') = P(s,t) : P(s',t) \qquad (*)$$

Let $s \sim_E t$ mean that states s and t are in the same cell of E. There are two cases to consider.

(a) We do not have s, s', and t all in the same E-cell; without loss of generality, say $s \nsim_E t$. Then Lemma 2 (a) tells us

$$\text{cond}_{P_0,E}(t)(s) = 0$$

So (*) holds for all cases of this form iff

$$P(s,t) = 0 \quad \text{for all } s \nsim_E t$$

This is equivalent to condition (a) of the theorem statement.

(b) If $s \sim_E s' \sim_E t$, then by Lemma 2 (b),

$$\text{cond}_{P_0,E}(t)(s) : \text{cond}_{P_0,E}(t)(s') = P_0(s) : P_0(s')$$

So (*) holds for all cases of this form iff

$$P_0(s) : P_0(s') = P(s,t) : P(s',t) \quad \text{for all } s \sim_E s' \sim_E t$$

This is equivalent to condition (b) of the theorem statement.

In short, (*) holds in all cases iff both conditions of the theorem hold. □

References

Briggs, R. A., and Richard Pettigrew. 2018. "An Accuracy-Dominance Argument for Conditionalization." *Noûs* 54 (1): 162–181. doi:10.1111/nous.12258.

Carnap, Rudolf, and Richard C. Jeffrey. 1971. *Studies in Inductive Logic and Probability.* Berkeley, CA, USA: University of California Press.

Christensen, David. 2010. "Rational Reflection." *Philosophical Perspectives* 24 (1): 121–140. doi:10.1111/j.1520-8583.2010.00187.x.

Dorst, Kevin, and Matthew Mandelkern. Forthcoming. "Good Guesses." *Philosophy and Phenomenological Research.* doi:10.1111/phpr.12831.

Easwaran, Kenny. 2013. "Expected Accuracy Supports Conditionalization—and Conglomerability and Reflection." *Philosophy of Science* 80 (1): 119–142. doi:10.1086/668879.

Edwards, Ward. 1982. "Conservatism in Human Information Processing." In *Judgment under Uncertainty: Heuristics and Biases,* edited by Daniel Kahneman, Paul Slovic, and Amos Tversky, 359–369. Cambridge University Press. doi:10.1017/CBO9780511809477.026.

Elga, Adam. 2013. "The Puzzle of the Unmarked Clock and the New Rational Reflection Principle." *Philosophical Studies* 164 (1): 127–139. doi:10.1007/s11098-013-0091-0.

Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein. 1977. "Knowing with certainty: The appropriateness of extreme confidence." *Journal of Experimental Psychology: Human perception and performance* 3 (4): 552.

Gallow, J. Dmitri. 2021. "Updating for Externalists." *Noûs* 55 (3): 487–516. doi:10.1111/nous.12307.

Greaves, Hilary. 2013. "Epistemic Decision Theory." *Mind* 122 (488): 915–952. doi:10.1093/mind/fzt090.

Greaves, Hilary, and David Wallace. 2006. "Justifying Conditionalization: Conditionalization Maximizes Expected Epistemic Utility." *Mind* 115 (459): 607–632. doi:10.1093/mind/fzl607.

Hawthorne, John, and Amia Srinivasan. 2013. "Disagreement Without Transparency: Some Bleak Thoughts." In *The Epistemology of Disagreement: New Essays,* edited by David Phiroze Christensen and Jennifer Lackey, 9–30. Oxford University Press.

Holguín, Ben. Forthcoming. "Thinking, Guessing, and Believing." *Philosophers' Imprint:* 1–34.

Horowitz, Sophie. 2019. "Accuracy and Educated Guesses." *Oxford Studies in Epistemology* 6. doi:10.1093/oso/9780198833314.003.0004.

James, William. 1896. "The Will to Believe." *The New World* 5:327–347.

Jeffrey, Richard C. 1983. "Bayesianism With A Human Face." In *Testing Scientific Theories,* edited by John Earman, 133–156. University of Minnesota Press.

Joyce, James M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65 (4): 575–603. doi:10.1086/392661.

Lasonen-Aarnio, Maria. 2015. "New Rational Reflection and Internalism About Rationality." *Oxford Studies in Epistemology* 5. doi:10.1093/acprof:oso/9780198722762.003.0005.

———. n.d. "Perspectives and Good Dispositions." *Philosophy and Phenomenological Research.* Forthcoming.

Leitgeb, Hannes, and Richard Pettigrew. 2010a. "An Objective Justification of Bayesianism I: Measuring Inaccuracy." *Philosophy of Science* 77 (2): 201–235. doi:10.1086/651317.

———. 2010b. "An Objective Justification of Bayesianism II: The Consequences of Minimizing Inaccuracy." *Philosophy of Science* 77 (2): 236–272. doi:10.1086/651318.

Levinstein, Benjamin Anders. 2017. "A Pragmatists Guide to Epistemic Utility." *Philosophy of Science* 84 (4): 613–638. doi:10.1086/693444.

Moss, Sarah. 2011. "Scoring Rules and Epistemic Compromise." *Mind* 120 (480): 1053–1069. doi:10.1093/mind/fzs007.

Pettigrew, Richard. 2016. *Accuracy and the Laws of Credence.* Oxford University Press UK.

———. 2020. "What is Conditionalization, and Why Should We Do It?" *Philosophical Studies* 177 (11): 3427–3463. doi:10.1007/s11098-019-01377-y.

Pollock, John L. 1986. *Contemporary Theories of Knowledge.* Hutchinson.

Quine, W. V. 1986. "Reply to White." In *The Philosophy of W. V. Quine,* edited by Lewis E. Hahn and Paul A. Schilpp. La Salle, Illinois: Open Court.

Salow, Bernhard. 2018. "The Externalists Guide to Fishing for Compliments." *Mind* 127 (507): 691–728. doi:10.1093/mind/fzw029.

Schafer, Karl. 2014. "Doxastic Planning and Epistemic Internalism." *Synthese* 191 (12): 2571–2591. doi:10.1007/s11229-014-0412-7.

Schoenfield, Miriam. 2017. "Conditionalization Does Not Maximize Expected Accuracy." *Mind* 126 (504): 1155–1187. doi:10.1093/mind/fzw027.

———. 2018. "An Accuracy Based Approach to Higher Order Evidence." *Philosophy and Phenomenological Research* 96 (3): 690–715. doi:10.1111/phpr.12329.

Schwarz, Wolfgang. 2018. "Imaginary Foundations." *Ergo: An Open Access Journal of Philosophy* 5. doi:http://dx.doi.org/10.3998/ergo.12405314.0005.029.

Srinivasan, Amia. 2015. "Are We Luminous?" *Philosophy and Phenomenological Research* 90 (2): 294–319. doi:10.1111/phpr.12067.

van Fraassen, Bas C. 1989. *Laws and Symmetry*. Oxford University Press.

Vul, Edward, and Harold Pashler. 2008. "Measuring the Crowd Within: Probabilistic Representations Within Individuals." *Psychological Science* 19 (7): 645–647. doi:10.1111/j.1467-9280.2008.02136.x.

Williamson, Timothy. 1996. "Cognitive Homelessness." *Journal of Philosophy* 93 (11): 554–573. doi:jphil199693117.

Williamson, Timothy. 2000. *Knowledge and its Limits*. Oxford University Press.

———. 2008. "Why Epistemology Cannot Be Operationalized." In *Epistemology: New Essays*, edited by Quentin Smith. Oxford University Press.

———. 2011. "Improbable Knowing." In *Evidentialism and its Discontents*, edited by T. Dougherty. Oxford University Press.

———. 2014. "Very Improbable Knowing." *Erkenntnis* 79 (5): 971–999. doi:10.1007/s10670-013-9590-9.