FINE-GRAINED EVIDENCE

Stephen Yablo*

January 4, 2026

Abstract: Bayesian conditionalization is rigid: learning E fixes p(E) at 1 while preserving probabilities conditional on E. Non-rigid update is preferable when, in the course of learning *that* E is true, we change our views about *how*—by way of which truthmakers ϵ. A Jeffrey-style generalization of Bayes—active conditioning—is developed which gives learning events a handle on p($\epsilon|E$) and p(E) both. E brings a truthmaker-incorporating "probasition" to the table, rather than simply an intension. Confirmation relations go hyperintensional as a result. E_is true in the same worlds may not license the same updates, if their truth flows from different sources.

1 INTRODUCTION

You remember Bayes. He proposed a rule for adjusting our credences on learning that E. H had beforehand a probability conditional on E: $p_{old}(H|E)$ (= $p_{old}(H \wedge E)/p_{old}(E)$). Bayes advises us to be as confident now of H as we were earlier of H conditional on E:

(B-CON) $p_{new}(H) = p_{old}(H|E) = p_{old}(H \wedge E)\ /\ p_{old}(E)$. (1)

You remember Brentano. Mental states and activities are, he claims, *directed* at something. They have an *intentional object* or *subject matter*. Counterexamples have been alleged: undirected anxiety, nameless dread. But if we stick to propositional attitudes, the idea seems right:

(B-ABT) To believe/want/doubt/hope.....that E is (in part) to consider E's subject matter $\mathbf{e}$. (2)

These ideas are connected insofar as credence is a propositional attitude. If one has to consider $\mathbf{e}$ to believe E, then likewise presumably to believe it to such and such a degree.

Bayes' rule is strangely uncurious about $\mathbf{e}$. It inputs a set $\mathbf{E}$ of worlds, period, the ws where E is true. *Something* must be going on in those ws with regard to $\mathbf{e}$ to earn them a place in the set. And having just found E to hold in our world specifically, we may have views about what it is — about how our world qualifies.[1] B-CON ignores all this. Subject matter is treated as an irrelevant accretion, with no more bearing on H than the font in which E is rendered.[2]

*Lots of people to thank: David Builes, Lucas Champollion, Cian Dorr, Kevin Dorst, Kit Fine, Branden Fitelson, Bas van Fraassen, Caspar Hare, Sally Haslanger, Stephan Krämer, Ofra Magidor, John MacFarlane, Sven Neth, Sonia Pavel, Alexander Pruss, Michael Rescorla, Jack Spencer, Robert Stalnaker, Roger White, Seth Yalcin, Martín Abreu Zavaleta. Thanks. I am grateful to the reviewers for some extremely valuable comments (looking at you, Reviewer Five). Versions of this material were presented at Ohio State University, Berkeley, Oxford, the 2024 Eastern Division APA Meetings, and truthmaker conferences in Geneva and Munich.

[1] If we *knew* how it qualified — by being F — then we'd be updating on it rather than E. This does not prevent us from having views about how it *probably* qualifies.

[2] $\mathbf{e}$ in this paper is the set of E's truthmakers ϵ. Truthmakers are sui generis fact-like entities. E-learning is registering that an ϵ obtains (perhaps by perceiving one). Atomic sentences get their truthmakers (falsemakers) from the model. The rest are assigned recursively: $\tau \Vdash \neg Q$ iff $\tau \dashv\mid Q$; $\tau \Vdash$

2 RELEVANCE

I called B-CON "strangely" uncurious; but curiosity would be strange, if co-intensionals E_1, E_2 always licensed the same updates, their aboutness differences notwithstanding. This is one of the things to be explored. Here though is how aboutness properties *could* be relevant..

A subject matter $\mathbf{m}$ is independent of H if p($H|M$)=p(H) for every way M that matters can stand $\mathbf{m}$-wise. E is unlikely to support H if $\mathbf{e}$ is H-independent.[3] But $\mathbf{e}_1$ and $\mathbf{e}_2$ may differ in this respect; $\mathbf{e}_1$ is independent of H, $\mathbf{e}_2$ is not. Should $\mathbf{e}_2$-ways be sometimes positively relevant, and never negatively, that could give E_2 an evidential advantage over E_1, so that it alone supports H.[4]

Objection: Won't the E_is' topical differences prevent them from holding in the same worlds? Reply: A sentence's subject matter is to do with the aspects of w it takes note of and makes demands on. Intensions are the worlds that meet the demands. Demands visited on different aspects of w can still be satisfied by the same ws. *No Fs are G* is about which of the Fs are G; *No Gs are F* is about which of the Gs are F. A world answers *None* to the first question iff to the second. Still, more Fs are encountered in checking whether *No Fs are G* than *No Gs are F*. One can imagine them bearing differently on *Some Fs are G′* as a result.

That was a "what if?", not an argument. And there's an argument on the other side. p(•$|E_1$) = p(•$|E_2$), since probability is intensional. What could it mean for E_1 to be better news for H than E_2, if p($H|E_1$) = p($H|E_2$)? So we have our work cut out for us. There will have to be more to evidential support than probabilification, if subject matter is to be evidentially relevant.[5]

3 UNDERPINNINGS

A sentence's subject matter is tied up with how it is liable to be true. Worlds are alike where $\mathbf{s}$ is concerned if S is true in the same way(s) in them. $\mathbf{s}$ is identified formally with the set of those ways: the set of its truthmakers σ.[6]

Subject matters cannot be assessed for probability, but truthmakers can. p(ϵ^2) is the probability that the cloth is blue, p($\epsilon^2\epsilon^5$) the probability that it is blue and teal. This comes in handy, as p(E) is a function of the probabilities of E's truthmakers ϵ, and conjunctions thereof. Truthmakers-alone-and-conjoined will be *t-makers*, written $\dot{\epsilon}$. Here is p(E) as a function of the probabilities of its t-makers:[7]

$$\text{p}(E) = \Sigma_i\ \text{p}(\epsilon^i) - \Sigma_{i\neq j}\ \text{p}(\epsilon^i\epsilon^j) + \Sigma_{i\neq j\neq k}\ \text{p}(\epsilon^i\epsilon^j\epsilon^k) - \Sigma_{i\neq j\neq k\neq l}\ \text{p}(\epsilon^i\epsilon^j\epsilon^k\epsilon^l) + \ldots \tag{3}$$

The simplest case is $E = A\vee B$, where A is made true precisely by α, and B precisely by β:[8]

$$\text{p}(A\vee B) = \text{p}(\alpha) + \text{p}(\beta) - \text{p}(\alpha\beta) \tag{4}$$

$R\vee S$ iff $\tau \Vdash R$ or $\tau \Vdash S$; $\tau \Vdash R\wedge S$ iff $\tau = \rho\wedge\sigma$ for some ρ, $\sigma \Vdash R$, S respectively. Truthmakers have intensions; $|\tau|$ is the set of worlds where τ obtains. Truth-functional combinations of τs have intensions too: $|\bar{\rho}|$ is $|\rho|$'s complement, $|\rho\wedge\sigma| = |\rho|\cap|\sigma|$, $|\rho\vee\sigma| = |\rho|\cup|\sigma|$. Unusually for truthmakers, ours are going to be coarse-grained: $|\tau| = |\tau'| \Rightarrow \tau=\tau'$. They individuate like sets of worlds, but aren't identified with them. τs are part of the causal order so presumably fact-like.

[3] There are issues here about conglomerability, the principle that p($H|E$) lies in a certain range if p($H|E^i$) lies in that range for each cell E_i of a partition <E_i> of E ([31]).

[4] Is evidential support a relation between sentences, or propositions? Both. E_2 differs evidentially from E_1 because $\boldsymbol{E}_2$, the subject-matter-enhanced proposition it expresses, differs evidentially from $\boldsymbol{E}_1$. $\boldsymbol{E}$ can be construed as the ordered pair ($\mathbf{E}$, $\mathbf{e}$) of E's intension and its subject matter. This is inelegant since $\mathbf{E}$ is recoverable as $\cup_\epsilon\ |\epsilon|$. But it simplifies the discussion.

[5] Shout-out to other work barking up similar trees: [17], [18], [22]: ch 3, [25], [3], [2]. (And [33].)

[6] [32], [34]

[7] E stands in probabilistic relations with disjunctions of ϵs, too. Trivially p(E) = p($\epsilon^1\vee\ldots\vee\epsilon^m$) as both sides measure the same set of worlds. $|E|$ = the E-worlds = $|\epsilon^1\vee\ldots\vee\epsilon^m| = \cup_i\ |\epsilon^i|$, since E is true somehow or other. Less trivially, p($A\vee B\vee C$) = p($\alpha\vee\beta$) + p($\beta\vee\gamma$) - p(β) - p($\alpha\gamma$) + p($\alpha\beta\gamma$). $\dot{\epsilon}$ ranges over E's t-makers — its truthmakers closed under $\wedge$. Section **10** brings in E's *dt*-makers $\mathring{\dot{\epsilon}}$ = its t-makers closed under $\vee$.

[8] $\alpha\beta$, their conjunction, is sometimes allowed as a third truthmaker. Here we give it the lesser status of t-maker. Truthmakers are *exact* in Fine's sense ([11]); they're wholly positively relevant. $\alpha\gamma \nVdash A\vee B$ since γ is irrelevant. Still less does $\alpha\bar{\beta} \Vdash A\vee B$; $\bar{\beta}$ plays for the other team.

Where does p(E) "come from" when E is $A \vee B$? (4) gets us thinking of α, β, and $\alpha\beta$ as helper-fish bumping E around in the probabilistic waters. All one sees from above when E is learned is that p(E) has broken the surface. It couldn't have done that without the cooperation of the $\dot{\epsilon}$s, but Bayesian updaters deal in certainties. They don't know, anyway are not to consider, what the helper-fish might have been doing. Things get messy when one does consider it, as there are any number of possibilities.

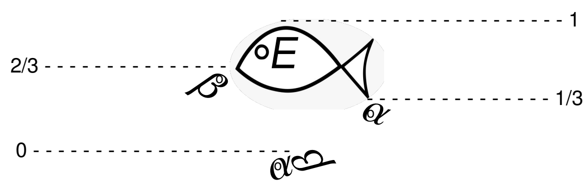

Figure 1: E breaking surface

<p(α), p(β), p($\alpha\beta$)> might for instance be <$\frac{1}{3}$, $\frac{2}{3}$, 0 >. It could also be <$\frac{1}{4}$, $\frac{3}{4}$, 0 >, <$\frac{3}{5}$, $\frac{3}{5}$, $\frac{1}{5}$>, <$\frac{3}{5}$, $\frac{4}{5}$, $\frac{2}{5}$> and so on indefinitely. There are continuum many t-maker-probability routes to p($A \vee B$)=1.

This might remind you of Frege on the underdetermination of sense by reference. There is no backward path, he says, from **t**, a term t's referent, to t, its sense. **t** does *constrain* the sense; t has got to be a mode of presentation of that very item. But there are lots of senses that do this, and ...t...'s cognitive content is sensitive to the differences. The corresponding constraint for us is that Σ p(ϵ^i) - Σ p($\epsilon^i\epsilon^j$) + Σ p($\epsilon^i\epsilon^j\epsilon^k$) - ... has got be 1. This massively underdetermines the p($\dot{\epsilon}$)s, needless to say. There is no backward (downward) path from p(E)=1 to E's probabilistic underpinnings

Underpinnings may be ignored if E's evidential value is the same regardless. But, one might as well ignore what a thing is made of (gold, wood, mud) when assigning it a monetary value. E's status as evidence — the fact that p(E)=1 — is constituted by the $\dot{\epsilon}$s being thus and so probable, as its truth in w is made up of the t-makers that obtain there. $\dot{\epsilon}$s carrying more of the burden deserve more of a say, surely. If β favors H, then α (being only half as likely) would have to strongly disfavor it, for H not to be supported on balance by $A \vee B$.

4 RIGIDITY

T-maker neglect seems at first like a fumble or oversight, the kind of thing B-CON-ers could easily set right. But it's written into the rule's DNA. Bayesian update "just is" the update rule that preserves probabilities conditional on E: $p_n(X|E)=p_o(X|E)$. It's in the jargon the one and only *rigid* update rule. E's subcases $\acute{E}$ — propositions holding in a proper subset of the E-worlds — appreciate by rigidity at a fixed E-determined rate,

$$\begin{aligned} p_n(\acute{E}) &= p_o(\acute{E}|E) && \textsc{b-con} \\ &= p_o(E \wedge \acute{E})/p_o(E) && \text{defn of conditional probability} \\ &= p_o(\acute{E}) \times \underline{1/p_o(E)} && p_o(E \wedge \acute{E}) = p_o(\acute{E}) \text{ since } \acute{E} \text{ implies } E \end{aligned} \tag{5}$$

Now, its t-makers $\dot{\epsilon}$ are certainly subcases of E. So they too appreciate at the underlined rate. Our "mistake" it seems was to think of $\dot{\epsilon}$s as having their probabilities set independently, then fixing E's probability via (3). Really they trail along after E. This is why Bayesians feel they can disregard underpinnings. There was only ever one option: $p_n(\dot{\epsilon})=p_o(\dot{\epsilon})/p_o(E)$. Underpinnings are in effect *stipulated*. E and its t-makers fly in a fixed probabilistic formation that E-learning cannot disturb. Only the scale changes, and that uniformly throughout.

But, but ... what do you mean, underpinnings are stipulated? Who gave the classical updater that power? $p_n(\dot{\epsilon})$ is not a theoretical plaything to be adjusted as convenient. It's a psychological reality that depends on the learning scenario. It had better depend on something, or "proper degree of belief on learning *E*" loses all meaning.[9] An example to bring this out.

> COINS Three coins are tossed: Al's silver dollar, Betty's loonie (Canadian dollar), and Cleo's one euro coin. We've bet on all three landing heads (H=ABC). The news E arrives that $A \vee B$. E's t-makers are α = dollar-heads, β = loonie-heads, and $\alpha\beta$ = dollar-and-loonie-heads.[10] p(E) was originally 3/4 (the prior probability of two tails is 1/4) but now becomes 1.

α, β, and $\alpha\beta$ can't stay where they were, or p(α)+p(β)-p($\alpha\beta$) is stuck at 3/4. p(E) will have to be supplied with new underpinnings. Any number can be imagined, one per triple of reals r, s, $t \in [0,1]$ such that $r+s-t = 1$. Here are four to get us started:

	$A \vee B$	α	β	$\alpha\beta$	ABC
prior probabilities	3/4	1/2	1/2	1/4	1/8
posterior probabilities	1	↓	↓	↓	
UP_0	1	2/3	2/3	1/3	1/6
UP_1	1	3/4	3/4	1/2	1/4
UP_2	1	7/8	7/8	3/4	3/8
UP_3	1	9/10	9/10	4/5	2/5

Table 1: Underpinnings for $p_o(A \vee B)=3/4 \rightsquigarrow p_n(A \vee B)=1$

[[Insert Table 1 here.]]

Where is the Bayesian on this table? $p_n(\alpha)$ according to B-CON is $p_o(\alpha)$ divided by $p_o(A \vee B)$. That last is 3/4 (the chance of two tails was 1/4), so we're looking at 1/2÷3/4 = 2/3. $p_n(\beta)$ is the same. $p_n(\alpha\beta) = p_o(\alpha\beta)/p_0(A \vee B) = 1/4 \div 3/4 = 1/3$. 2/3, 2/3, 1/3 is UP_0. So the Bayesian (classical) underpinnings are UP_0. $p_n(ABC)$ is given as 1/6 on the UP_0 row since that is half of $p_n(AB)$ =1/3 (Cleo's toss is independent of the others). B-CON agrees; $p_o(ABC|A \vee B)= p_o(ABC)/p_o(A \vee B) = \frac{1}{8} \div \frac{3}{4} = \frac{1}{6}$.

Further down on the table we see other assignments to α, β, and $\alpha\beta$ whereby $p_n(A \vee B)$=1. $p_n(ABC)$ remains at 1/6 for the Bayesian because $p_o(ABC)$ and $p_o(A \vee B)$ are still 1/8 and 3/4. But this is no longer the right result. $p_n(ABC)$ ought to be 1/4 — $p_n^{\checkmark}(ABC)$=1/4 — if $p_n(AB)$=1/2 (UP_1), 3/8 if $p_n(AB)$=3/4 (UP_2), and 2/5 if $p_n(AB)$=4/5 (UP_3). UP_0-UP_3 work equally well as far as the math goes. There is so far nothing to choose between them. And yet the Bayesian somehow knows.

5 BACKSTORIES

E's new probability is supported (underpinned) by its t-makers' probabilities $p_n(\dot{\epsilon})$. But, where do t-makers get their probabilities? $p_n(\dot{\epsilon})$ has figured so far as an independent variable bearing on $p_n(H)$. ($p_n(H) = p_n(\alpha\beta)/2$ when H is

[9] A word about that "proper." Update is something we do, like deductive inference. It can be done well or badly. Update rules help us organize our procedures with a view to doing it well, much as inference rules (modus ponens, etc.) do on the deductive side. B-CON is a tool for self-understanding and self-governance. It's a good tool if its recommendations make sense on reflection. Of course there's a question of what its recommendations are. That a good result *can* be reached by B-CON (by reverse-engineering an E and priors $p_o(\bullet)$ such that $p_n(H) = p_o(H|E)$) is only so interesting from our perspective. A just-so story can almost always be concocted by packing clues into E which $p_o(\bullet)$ somehow knows how to decipher. That may be useful for some purposes — self-governance is not one of them. An *improper* $p_n(H)$ is a credence it would be procedurally unreasonable to adopt. B-CON's *recommendations* are the $p_n(H)$s it suggests we adopt (not the retrospectively likeable ones accommodated after the fact). A recommended update is not automatically proper.

[10] $\alpha\bar{\beta}$ isn't on the list for two reasons. Formally it's because $\alpha\bar{\beta}$ is not an $A \vee B$-truthmaker (those are α and β) nor a conjunction of them. (See note 2.) Philosophically it's because t-makers like truthmakers are wholly *positively* relevant, and $\bar{\beta}$ plays for the other team.

ABC.) But, $\dot{\epsilon}$ may be itself among the hypotheses H whose probability we are trying to figure out. E's underpinnings in that second guise are *dependent* variables. They depend on what I'll be calling a *backstory* BK. Backstories specify how the information reaches the agent, that we come away expecting E to hold this way rather than that..[11]

Learning takes many forms; we'll need to impose some structure. E is delivered by an event $\mathcal{E}$. Learning events $\mathcal{E}$s are sorted according to which t-makers are given how much of a boost. Here is the Al and Betty story again, with an observation thrown in as delivery vehicle.

> COINS, ELABORATED Coins are tossed side by side on a stage. Is that Al's on the left, in which case Betty's is on the right? Or maybe it's the other way round; the two are difficult to tell apart. *Something* is seen that makes clear that $A \vee B$, leaving us undecided however on A and on B.

Three scenarios will be considered. The first and Bayes-friendliest has the coins disappearing behind a curtain before either lands. A sign emerges (by prior arrangement) just if either has come up heads: *Neither? No!*, it says. Call this UNSEEN since neither coin is seen landing.. Righty is still behind the curtain in the second scenario, but Lefty is seen to land heads (ONESEEN). Was that Al's US dollar we were looking at, or Betty's loonie? It definitely wasn't both. The last scenario has two sightings, first of Lefty_1 and then Lefty_2. They may or may not be the same coin. "Both" are in the position they assumed on landing: heads up. Was it the silver dollar we saw (twice), or the loonie (twice), or both (one after the other)? This is TWOSEEN.

5.1 *UNSEEN*, THE BAYESIAN PICK

UNSEEN, or BK_0, is the null case. One learns *that* $A \vee B$, period. No contact is made with E's truthmakers, nor is any news received about them, beyond that at least one obtains. Bare learning, this might be called, or pure learning-that. It should be the rule, not the exception, if Bayesian update is to be generally appropriate.

The curtains draw shut before either coin has a chance to land. A sign appears announcing that it's not the case that neither landed heads.

Our source, the sign-holder, may have seen the coins land. But they could equally have heard about it from a third individual, who heard about it from a fourth,... Or perhaps no coins were observed. A buzzer sounds when the number of heads is 0, and they noticed its absence.

Figure 3: Not the case that neither coin landed heads

[11] "Shouldn't the backstory be counted into our evidence? In which case we're not updating just on E." I agree that we're not updating just on E. But there's a question of what should be added. Option 1: A description Z of the learning circumstances such that $p_n(\dot{\epsilon}) = p_o(\dot{\epsilon} \mid E \wedge Z)$. Option 2: An assessment of which $\dot{\epsilon}$s appear to the learner to be in view. Orthodoxy likes the first option. There's nothing definitively *wrong* with it, but we'll be pursuing the second. Potshots will be taken now and then at the first. (E.g., in note 9: that B-CON can *accommodate* an update isn't saying much. Think of all the inferences that can be set out as modus ponenses. Note 14: Z may register subpersonally. But it is too much to be negotiated at a personal level.) If you're not convinced, fine. No harm in laying out an alternative for your consideration.

The sign leaves us with three possibilities: $\alpha\beta$, $\alpha\overline{\beta}$, and $\overline{\alpha}\beta$. They are incompatible and all of the same probability x. $3x = 1$ since one of the three definitely obtains, hence x=1/3. How probable are $A\vee B$'s t-makers? $\alpha\beta$ was one of the original three, so p($\alpha\beta$) is 1/3. p(α) = p($\alpha\beta$) + p($\alpha\overline{\beta}$) = 1/3+1/3 = 2/3; similarly for p(β). $\text{p}_n(\alpha) = \text{p}_o(\alpha)/\text{p}_o(A\vee B)$ = 1/2 ÷ 3/4 = 2/3. $\text{p}_n(\beta) = \text{p}_o(\beta)/\text{p}_o(A\vee B)$ = 1/2÷3/4 = 2/3. $\text{p}_n(\alpha\beta) = \text{p}_o(\alpha\beta)/\text{p}_o(A\vee B)$ = 1/4 ÷3/4 = 1/3. All as B-CON predicted. BK_0 is how the Bayesian tells the story.

5.2 *ONESEEN*, AND A FAMOUS FALLACY

UNSEEN puts the probability of two heads at 1/3. A different answer is sometimes suggested (cf., Martin Gardner's "two children" puzzle ("Mathematical Games," [13])).

1. $A\vee B$ ... p($A\vee B$)=1.
2. p(*A coin landed heads*)=1.
3. p($A\wedge B$) = p(*The other coin did as well*).
4. p(*The other coin landed heads*) = $\frac{1}{2}$.
5. So p($A\wedge B$) = $\frac{1}{2}$.

Gardner initially thought the reasoning here was fallacious. It seems to break down at line 3. *What* other coin? The case is underdescribed, he later decided. p($A\wedge B$) might or might not be 1/2, depending on where $A\vee B$ comes from. Imagine we learned it as follows (BK_1):

The left-hand curtain stays open. Lefty (either Al's coin or Betty's) is seen to land heads. The other curtain closes on Righty while it is still in the air.

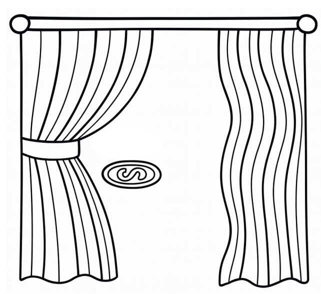

Figure 3: Lefty lands heads; Righty is hidden.

Al's coin lands heads, if it does, either seen or unseen. p(α) then is p($\mathcal{S}[\alpha]$) + p($\alpha\wedge\neg\mathcal{S}[\alpha]$). p($\mathcal{S}[\alpha]$) = p(*Lefty is Al's coin*) = 1/2, and p($\alpha\wedge\neg\mathcal{S}[\alpha]$) is p($\alpha|\neg\mathcal{S}[\alpha]$)×p($\neg\mathcal{S}[\alpha]$) = 1/2 × 1/2 = 1/4. So, p(A) = p(α) = 1/2 + 1/4 = 3/4. Similarly p(B) = 3/4. $A\wedge B$'s probability is A's plus B's - p($A\vee B$) = 3/4 + 3/4 - 1 = 1/2. Which gives us UP_1: $<r, s, t> = <3/4, 3/4, 1/2>$.

Back now to the Gardner argument. Line 4 could be true or false, depending on how "the other coin" is picked out. If it's Righty, then yes, p(*The other coin landed heads*)=1/2. Suppose however that *Headsy* is introduced as a name for (i) the unique heads if there is one, (ii) one of the two chosen randomly if both land heads; and that *Othersy* $=_{df}$ Al's coin if Headsy is Betty's, and Betty's coin if Headsy was Al's. Othersy definitely landed tails on the not implausible hypothesis that Headsy was a forced choice (clause (i) applied). It landed heads iff Headsy was freely chosen (clause (ii)). p((ii)) = 1/3, so that is the probability too of Othersy landing heads.[12]

[12] "Lefty and Othersy might be the same coin!" That is a different problem, akin to "How can Julius's probability of inventing the zip be higher than that guy's, when Julius is that guy?" It's clearer what is going on than what to do about it. Julius so-called earned that title by a process that was certain to deliver a zip-inventor. Othersy earned its title by a process that was likely to deliver a tails. That guy's status as that guy is independent of his inventor-properties; Righty's status as Righty is independent of how it landed. (See [8], [25]:147ff.)

5.3 *TWOSEEN*, A UNIVERSAL UNDERPINNINGS MACHINE

The probability of both landing heads ≤ 1/2 if only one coin is seen. UP$_2$ and UP$_3$ put it over 1/2, so we'll need a second sighting. The left curtain draws shut after our initial observation, reopening a moment later on "another" heads. Lefty$_1$ could be Al's coin or Betty's, and likewise Lefty$_2$. And Lefty$_1$ might or might not be Lefty$_2$. The probability α was seen (early or late) is judged, based on the Lefty$_i$s' appearances and the time between views, to be a; the probability β was seen is judged to be b.

Two sightings are made of heads-up coins. Was that (i) α twice, (ii) β twice, or (iii) one of each, in either order? p(α-*seen*) = p($i \vee iii$). p(β-*seen*) = p($ii \vee iii$).

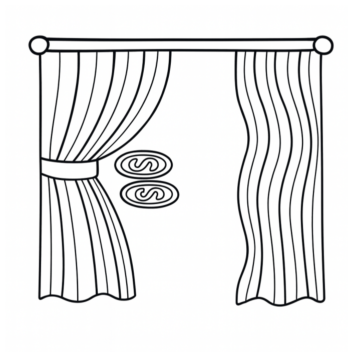

Figure 4: Lefty$_1$ is seen and then Lefty$_2$; "both" are heads-up.

What is the probability that Al's coin landed heads? It's the sum of two other probabilities: that of Al's coin being seen in a heads up position, and that of its being heads up unseen. Abbreviating α-*seen* to $\mathcal{S}[\alpha]$, the first is p($\mathcal{S}[\alpha]$) = a. The second is p($\bar{\mathcal{S}}[\alpha]$) times p($\alpha|\bar{\mathcal{S}}[\alpha]$) = (1-$a$)/2. All in all

$$p(\alpha) = a + (1-a)/2 = (a+1)/2. \tag{6}$$

β similarly could have obtained seen, with probability b, or unseen, with probability (1-b)/2. So

$$p(\beta) = b + (1-b)/2 = (b+1)/2. \tag{7}$$

$\alpha\beta$'s probability (since p($A \vee B$) = 1 = p(α)+p(β)-p($\alpha\beta$) is the sum of α's with β's, minus 1.

$$p(\alpha\beta) = (a+1)/2 + (b+1)/2 - 1 = (a+b)/2. \tag{8}$$

These equations yield UP$_2$ — <7/8, 7/8, 3/4> — if a and b are imagined to be 3/4, UP$_3$ if a=b=4/5.[13] All UP$_i$s can be obtained this way on which $\alpha\beta$ is at least half-probable For <p(α), p(β), p($\alpha\beta$)> = <r, s, t (= (r+s) - 1)>, let

- a, the probability α is seen, be 2r-1, and
- b, the probability β is seen, be 2s-1.

a = 2r-1 ⇒ r = (a+1)/2, which by (6) is p(α). b = 2s-1 ⇒ s = (b+1)/2, which by (7) is p(β). r+s - 1 = (a+b)/2, which by (8) is p($\alpha\beta$). For <.9, .6, .5> and <.6, .9, .5>, set one of a, b to .8 and the other to .95; for <.8, .8, .6>, let both be .9. UP$_i$s where p($\alpha\beta$)<1/2 are obtainable by supposing that tails was seen twice rather than heads.

[13] UP$_2$: p(α) = p(β) = (3/4 + 1)/2 = 7/8. p($\alpha\beta$) = 7/8 + 7/8 - 1 = 3/4. UP$_3$: p(α)=p(β) = (4/5 + 1)/2 = 9/10. p($\alpha\beta$) = 9/10 + 9/10 - 1 = 4/5.

6 "BATTING 1.000"

The last few sections have been variations on a theme. There was first a backstory BK assigning probabilities to $\copyright$ meaning $\dot{\epsilon}$. That backstory then generated underpinnings UP, probability assignments to the $\dot{\epsilon}$s whose combined effect is to raise p(E) to 1. BK_0 generated UP_0, BK_1 generated UP_1, and so on — as laid out in Table 2:

[Insert Table 2 here.]

	$\copyright[\alpha]$	$\copyright[\beta]$	$\copyright[\alpha\beta]$		$\alpha\vee\beta$	α	β	$\alpha\beta$	
prior probabilities	N/A	N/A	N/A		3/4	1/2	1/2	1/4	
posterior probabilities	↓	↓	↓		↓	↓	↓	↓	
BK_0	0	0	0	⇒	1	2/3	2/3	1/3	UP_0
BK_1	1/2	1/2	0	⇒	1	3/4	3/4	1/2	UP_1
BK_2	3/4	3/4	1/2	⇒	1	7/8	7/8	3/4	UP_2
BK_3	4/5	4/5	3/5	⇒	1	9/10	9/10	4/5	UP_3

Table 2: Underpinnings from backstories

The probabilities Bayes would have us adopt on learning $A\vee B$ are p(α)=p(β)=2/3, p($\alpha\beta$)=1/3. This is good advice should $A\vee B$ be learned as in UNSEEN. But not if we observe an unidentified coin landing heads (ONESEEN), or two heads-up coins that might or might not be the same (TWOSEEN). B-CON to go by the chart is batting .250; its advice is wrong in three of four cases. Even that is too generous. One can make the average as low as one likes by adding, say, $\text{UP}_4 = <\frac{1}{2}, \frac{1}{2}, 0>$, $\text{UP}_5 = <\frac{3}{5}, \frac{3}{5}, \frac{1}{5}>$, $\text{UP}_6 = <\frac{3}{5}, \frac{2}{3}, \frac{4}{15}>$, $\text{UP}_7 = <\frac{4}{5}, \frac{4}{5}, \frac{3}{5}>$, $\text{UP}_8 = <\frac{5}{6}, \frac{5}{6}, \frac{2}{3}>$ and so on.

So, then: classical updaters get it wrong ≥ 99.999...% of the time.[14] They are wrong because (i) "it" is a moving target ($p_n(ABC)$ can be many things), and (ii) the factors moving it — the probabilities of E's t-makers — are not on B-CON's radar. There may be a standard reply to this, I don't know. Here is the one that occurs to me.

> B-CON comes like any tool with a manual. Update only on Es known for certain, we read in Chapter 1. That's not enough, according to Chapter 2: "the mere fact that X knows E does not entitle him to believe H to the degree r" (r being p($H|E$); ([7]:211)). For X may know as well an E' such that $E\wedge E'$ probabilifies H less/more than E alone. This is the "total evidence requirement":[15]
>
> (TE) Don't update on E, if it omits a relevant E': $p_o(H|E)\neq p_o(H|E\wedge E')$ (9)

[14] Aka 100% of the time. Bayesians will say, "One's true evidence is $(A\vee B)\wedge Z_i$ (i =0-3), Z_i laying out that such and such explanations were received when we entered the theater from a thus and so trustworthy-seeming source etc etc etc. $p_o(AB|(A\vee B)\wedge Z_i)$ is the number you're looking for, and it varies with i as desired." Response: This is the kind of just-so story alluded to earlier (note 9). It leaves a lot to the imagination, and the rest we are expected to take on faith. AB is 1/3 (1/2, 3/4, 4/5) probable, conditional on $(A\vee B)\wedge Z_0$ (Z_1, Z_2, Z_3)? This is more of a wish list than a supportable hypothesis. Z_is of the required type would be bafflingly complex; yet one is supposed to see in advance their probable implications for H. (*Both heads* is 3/4 probable if: a philosopher with a firm handshake says that BLAH ... and ... and ... and ... and....heads-up coins are seen five seconds apart.) Compare Jeffrey's doubts ([16], 165) about the availability of an E capturing "the precise quality of one's visual experience," reflection on which reveals it to confer .7 probability on *It's green*. A rational through-line *can* be drawn, we shall see, from $A\vee B$ plus (not Z_i but) a probability distribution over the t-makers plausibly on display when $A\vee B$ is learned.

[15] Note 14 asks in effect whether TE blocks the counterexamples already. "X has to be *aware* of the changing backstories for them to license different responses. She has to learn the conditions E' under which she is learning E. She should be updating on $E\wedge E'$." This has its points as a defensive, "catch me if you can"-type maneuver. I want to put it aside, and move on to (IE), both for reasons already mentioned (note 14) and because TE is nearly unappeasable. So easy are incrementally relevant E's to come up with that even perfectly worthy Es are disqualified. Step 1: Choose at random an also-learned I that's omissible on account of being H-irrelevant. Step 2: Form I's disjunction with H ($\neg H$). Step 3: Count E out on the ground that both are (on modest assumptions) H-relevant; $(p(H|E\wedge(I\vee\neg H)) < p(H|E) < p(H|E\wedge(I\vee H))$. How are $I\vee H$ and $I\vee\neg H$ to be blocked? B-CON is for inactive (inert) evidence in a sense we're getting to. $I\vee H$ obtained by inference from I is active. Inertness comes first and runs deeper; TE cannot even be stated correctly without inertness controls on E'.

Here is what Chapter 2 was trying to get at, explains Chapter 3. Don't update on E, unless $p(E)\nearrow 1$ is the *sole legitimate input to your system of credences*. All other credal adjustments have got to trace back to that initial E-impulse. ONE way for them not to trace back is that more than E has been learned. ANOTHER is that E alone is learned, but probability is reapportioned over its subcases $\acute{E}$ in a way E doesn't prepare us for.[16]

Call $\acute{E}$ *inert* if $p(\acute{E})$ remains the same fraction of $p(E)$ as before, when E is learned the contextually indicated way (via $\mathcal{E}$); otherwise $\acute{E}$ is *active*. E is itself inert (otherwise active) if all its subcases are. When $\mathcal{E}$ conveys precisely *that E* holds, offering no clues about how, E is inert and we're good to go. When clues are offered, E is unfit for duty due to its active subcases:

$$\text{(IE) Don't update on } E\text{, if it has active subcases } \acute{E} : p_n^{\checkmark}(\acute{E}) \neq p_o(\acute{E})/p_o(E) \tag{10}$$

($p_n^{\checkmark}(\acute{E})$ is your proper, $\mathcal{E}$-licensed, credence in $\acute{E}$.) $\acute{E}$ in practice will always be a t-maker $\dot{\epsilon}$, e.g., α, β, and $\alpha\beta$ when E is $A\vee B$. ($p_n^{\checkmark}(\alpha)>p_n^{\checkmark}(\beta)$ if the coin looks to be likelier Al's than Betty's.) The extra credal changes occurring when $\mathcal{E}$ teaches us more than E are caught by TE. IE takes on the extra changes occurring when $\mathcal{E}$ drops hints about how E probably holds.

Suppose B-CON is used only on *circumspect* $\mathcal{E}$s, those delivering just the fact that E. What becomes of the counterexamples in Table 2 (BK_1-BK_3)? α, β, and $\alpha\beta$ gain more probability in those scenarios than $A\vee B$ can account for. $p(\alpha\beta)$, originally at 1/4, "should" go to $1/4 \times 1/p_o(A\vee B) = 1/4 \times 4/3 = 1/3$. It's pushed higher (to 1/2, in ONESEEN) by the hint that drops when a *particular* coin is seen coming up heads. The counterexamples are blocked if hints are disqualifying. $A\vee B$ *is* admissible when we're told in so many words that *Not: neither landed heads*. It is inadmissible in the other three cases — the ones that B-CON gets wrong.

Can counterexamples really be written off in this way, by marking them "inadmissible"? It feels a little high-handed. That wasn't the intention, of course. IE was supposed to be just another box to tick before setting $p_n(H)$ to $p_o(H|E)$. $p_o(H|E)$ is 2/3? Go ahead, provided that $p_o(H|E\wedge E')$ is 2/3 as well. ϵ^3 doubles in probability? Fine, provided that $p_o(E) = 1/2$, and every other t-maker does too. Terms and conditions apply. You fed it an active E? That's on you; you were warned. On inert Es, B-CON is batting 1.000.

7 INERTNESS

The inertness gambit was charged with high-handedness. It has now entered its plea: not guilty by association with total evidence TE. I am not convinced. TE for all its difficulties (note 15) is a workable requirement. IE is not for reasons we now explain.

Testing: E is tested for inertness by checking $p_n^{\checkmark}(\dot{\epsilon})$ — $\dot{\epsilon}$'s deserved probability — against $p_o(\dot{\epsilon})/p_o(E)$. $p_n^{\checkmark}(\dot{\epsilon})$ is obtained how? Not with B-CON; it can't be trusted until E has been certified inert. No such circularity arises for TE. A totality check involves only $p_o(H|E\wedge E')$ and $p_o(H|E)$. $p_n^{\checkmark}(H)$ doesn't come into it.

Triviality: Inert Es *just are* Es such that $p_n^{\checkmark}(\dot{\epsilon}) = p_o(\dot{\epsilon}|E)$, as B-CON says. "B-CON restricted to inert evidence gives the right advice" amounts to: B-CON gives the right advice except when it doesn't. Any rule, even outright fallacies, does that much. Affirming the consequent ($A\rightarrow C, C \therefore A$) never fails if avoided when A is untrue. "I don't have a truth oracle; how do I know when to avoid it?" You don't have a correct-credences oracle either, or an inertness oracle. Same question.

Disregard: From the cardinalities it would seem that E is usually active. Is most of our evidence to be disregarded? You might say that TE has a similar issue, in that most Es are relevantly incomplete. But,

[16] Recall that B-CON is *rigid* (section 4): probabilities conditional on E are preserved. From $p_n(\acute{E}|E) = p_o(\acute{E}|E)$ it follows that all $\acute{E}$s appreciate at the same rate, viz. $1/p_o(E)$.

incomplete Es are not simply discarded. One repairs the omissions, and then conditions on $E \wedge E'$. Active Es are so far unsalvageable.

Further on the same point, learning is preceded, typically, by inquiry. Inquiry into whether E calls on us to engage with the world's **e**-ish aspects. This would seem to involve (since $\mathbf{e} = \{\epsilon \mid \epsilon \Vdash E\}$) reviewing its truthmakers with a view to which of them might obtain. None of them definitely *does*, or E would not be the strongest thing learned. But the exercise is bound to be educational about their probabilities. It's a problem then if E becomes inadmissible as p(ϵ) falls under the spell of outside influences. ("It was supposed to trail along after p(E)!" Odd that the act of learning E would make it not update material.)

Active Es *should* be discarded, if they're intrinsically misleading. But "active" just means that insight has been gained into how E probably holds. That is a good thing, one would think, not a stain on E's reputation. The problem with active Es is extrinsic rather. B-CON doesn't know what to do with them.

Kind of a jam B-CON has gotten us into. Has the time come to look for another rule? As it is, we are like the patient in a Henny Youngman joke: "Doctor, it hurts when I go like this." "So, don't go like that!" Not the response we were hoping for, especially if it's when we breathe that it hurts. One wants a diagnosis, and to be fixed so that the hurting stops. "Reverend Bayes, your rule lands me sometimes with the wrong credences." "So, use it the other times!" Again not so helpful. Especially if to find those other times, one needs a rule that works all the time.

8 ACTIVE LEARNING (1): UNCERTAINTY

The rule we are after would take active Es in stride. Rather than stipulating the probabilities of E's t-makers, it would take them on board and consider their implications. Three obstacles stand in the way. The first has nothing to do with t-makers. Bayes' rule inputs (i) conditional probabilities p($H|E$), and (ii) monadic/unconditional, certainties (Es known for sure). It outputs (iii) monadic probabilities p(H). What it does *not* take as input are monadic probabilities. And that's what t-maker probabilities p($\dot{\epsilon}$) would be. How does one feed the probability of anything, t-maker or not, into an update rule?

Uncertain inputs need attention anyway, for a reason noted by Ramsey. If "I think I perceive or remember something but am not sure, this would seem to give me some ground for believing it" ([29]:86).[17] I should believe to degree .65, say, that I fell off a horse on my fifth birthday (F), since I sort of remember doing so. I'd be sure of F, if it was definitely falling off (φ) I remembered. But it might have been *worrying* I'd fall off when the horse gave a giant kick (κ). $\mathcal{R}[\varphi]$ and $\mathcal{R}[\kappa]$ are all I have to go on. That I'm certain of neither (p($\mathcal{R}[\varphi]$), p($\mathcal{R}[\kappa]$)<1 means .65 is *not* "the degree of belief in [F] which it would be rational for me to have [given] the probability relation between [F] and the things I know for certain" (86).

How does a hazy memory make for tentative acceptance? Imagine it gradually growing stronger until it becomes crystal clear that $\mathcal{R}[\varphi]$. B-CON initially takes no notice; p(F) remains at .1 as p($\mathcal{R}[\varphi]$) rises from .55 through .77 to .99. Not until $\mathcal{R}[\varphi]$ takes that final .01-sized step does it jump from .1 to 1 = $p_o(F|\mathcal{R}[\varphi])$.

That cannot be right. The support F receives from $\mathcal{R}[\varphi]$ at the end did not come out of nowhere. It was gradually accumulating as p($\mathcal{R}[\varphi]$) grew. Shouldn't p(F) have been growing too? It would have been, had $p_n(F)$ been calculated like this — by taking a weighted average of $p_o(F|\mathcal{R}[\varphi])$ and $p_o(F|\mathcal{R}[\kappa])$:

$$p_n(F) = p_n(\mathcal{R}[\varphi]) \times p_o(F|\mathcal{R}[\varphi]) + p_n(\mathcal{R}[\kappa]) \times p_o(F|\mathcal{R}[\kappa]). \tag{11}$$

$p_o(F|\mathcal{R}[\varphi])=1$, and we can take $p_o(F|\mathcal{R}[\kappa])$ to be .2 (perhaps I was right to be worried). $p_n(F)$ is $.55 \times p_o(F|\mathcal{R}[\varphi]) + .45 \times p_o(F|\mathcal{R}[\kappa])) \approx .65$, then $.77 \times p_o(F|\mathcal{R}[\varphi]) + .23 \times p_o(F|\mathcal{R}[\kappa])) \approx .82$, then $.99 \times p_o(F|\mathcal{R}[\varphi]) + .01 \times p_o(F|\mathcal{R}[\kappa]) \approx .99$, and finally $p_o(F|\mathcal{R}[\varphi])=1$.

[17] Ramsey might not approve of the use we'll be making of his example, a reviewer points out.

Now, $\mathcal{R}[\varphi]$ is throughout the main thing favoring F. But it's just over half likely at first, and so hard to think of as "evidence." The evidence properly so-called is a distribution $\vec{\boldsymbol{E}}$ over $\mathcal{R}[\varphi]$ and $\mathcal{R}[\kappa]$: $\mathrm{p}(\mathcal{R}[\varphi])$=.55 and $\mathrm{p}(\mathcal{R}[\kappa])$=.45. ($E$ comes to express different $\vec{\boldsymbol{E}}$s — different *probasitions*, we'll be saying — as $\mathrm{p}(\mathcal{R}[\varphi])$ evolves.) Here there are two alternatives, $\mathrm{R}[\varphi]$ and $\mathrm{R}[\kappa]$, but there could be several, as in Jeffrey's familiar example:

> *Observation by candlelight* The agent inspects a piece of cloth by candlelight, and gets the impression that it is green, although he concedes that it might be blue or even (but very improbably) violet. If G, B, and V are the propositions that the cloth is green, blue, and violet, respectively, then the outcome of the observation might be that, whereas originally his degrees of belief in G, B, and V were .30, .30, and .40, his degrees of belief in those same propositions after the observation are .70, .25, and .05. ([16], 165).

The alternatives for Jeffrey are external-world hypotheses G, B, and V, not, as you might expect from the previous case, hypotheses $\mathcal{S}[\gamma]$, $\mathcal{S}[\beta]$, and $\mathcal{S}[\nu]$ about what is visually presented. This will be an important choice point later on (section **9**); set it aside for now.

Bayes' rule could perhaps accommodate these cases, given an E which captured "the precise quality of [one's] visual experience in looking at the cloth," or the force and vivacity at a given time of one's φ-recollection. E would be learned with probability 1; Jeffrey's .7, .25, and .5 would be derived as $\mathrm{p}_o(G|E)$, $\mathrm{p}_o(B|E)$, and $\mathrm{p}_o(V|E)$. No such E is available, in Jeffrey's view:

> the best we can do is to describe, not the quality of the visual experience itself, but rather its effects on the observer, by saying, "After the observation, the agent's degrees of belief in G, B, and V were .70, .25, and .05" ([16], 165-6)

There ought to be a kind of conditionalization that treats the distribution as evidence in its own right. $\mathrm{p}_n(H)$ would be the average of $\mathrm{p}_o(H|G)$, $\mathrm{p}_o(H|B)$, and $\mathrm{p}_o(H|V)$, weighted by the likelihood of G, B, V — so, $.70\times\mathrm{p}(H|G) + .25\times\mathrm{p}(H|B) + .05\times\mathrm{p}(H|V)$.

Before we get to the rule, there's a question of motivation. Jeffrey update is called for, people say, when and because nothing is learned with certainty. But something is, or easily could be, learned with certainty in his own examples. The only reason $G\vee B\vee V$ wasn't, in *Observation by candlelight*, is that Jeffrey has set things up so that it was already known; $\mathrm{p}_o(G) + \mathrm{p}_o(B) + \mathrm{p}_o(V) = 1$. Let $\mathrm{p}_o(V)$ be .3 instead of, as he has it, .4, and $G\vee B\vee V$ is .9-likely until the candlelight observation $\mathcal{S}$ is made. Someone "transcribing a lecture in a noisy auditorium... might think he had heard the word 'red,' but still think it possible that the word was actually 'led"' ([16], 166). Was it that Chamberlain got LED all over on his visits with Hitler, or RED all over? Hard to tell. Certainly though Chamberlain got led *or* red all over in Munich. $L\vee R$ is in all likelihood new knowledge. $\mathcal{H}$ takes $\mathrm{p}(L\vee R)$ from who knows what to 1.

The problem with updating Bayes's way has been misidentified if E was learned with certainty. It's that other things were "learned" with non-certainty, for instance, that the tablecloth was green rather than violet. $B\vee G\vee V$ (E) would have been a fine thing to (classically) update on, had additional information not come in that tilted the odds in G's favor.

A disjunction X of incompatible X^is is learned; insight is gained too into the probability of X obtaining X^j-ly rather than X^k-ly. H's new probability is going to be a weighted average of its old ones conditional on the X^is; $\mathrm{p}_o(H|X^j)$ counts for more than $\mathrm{p}_o(H|X^k)$ if $\mathrm{p}_n(X^j)>\mathrm{p}_n(X^k)$. $\mathrm{p}_o(H|X^i)$ becomes in effect a random variable whose value we are trying to estimate:

$$\text{J-CON} \qquad \mathrm{p}_n(H) = \mathbb{E}[\mathrm{p}_o(H|X^i)] = \Sigma_i\, \mathrm{p}_n(X^i) \times \mathrm{p}_o(H|X^i) \tag{12}$$

This can be made to look more like B-CON by bringing in probasitions $\vec{\boldsymbol{X}}$: assignments to the X^is of probabilities summing to 1. Probasitional learning is learning X by, or while, distributing probability $\vec{\boldsymbol{X}}$-ly over its subcases X^i. If you learn simply that X, by all means update classically: $\mathrm{p}_n(H) = \mathrm{p}_o(H|X)$. J-CON kicks in when your evidence is $\vec{\boldsymbol{X}}$:

J-CON (with probasitions) $\quad \mathrm{p}_n(H) = \mathrm{p}_o(H|\vec{X}) = \mathbb{E}[\mathrm{p}_o(H|X^i)] = \Sigma_i\ \mathrm{p}_n(X^i) \times \mathrm{p}_o(H|X^i)$ (13)

X is *learned* only if $\mathrm{p}_o(X) < 1$. But it's the same idea regardless. A kind of information — graded, or probasitional — has come in that B-CON can't handle. J-CON lets us use it rather than lose it.

9 ACTIVE LEARNING (2): AVAILABILITY

A rule A-CON for active learning would allow E's t-makers to appreciate at their own rates.[18] J-CON with its openness to probabilistic inputs gets us part of the way there. But not all. The $\dot{\epsilon}$s for one thing don't tend to form a partition. Partitionality is addressed in the next section. Here we ask about the cells/subcases taken individually. What do $|X_i|$'s members have in common? Or, as we might also put it, what are $\boldsymbol{X}$ (and $\vec{X}$) about? This is the choice point mentioned earlier.

Start as usual with E — our new information, delivered by $\mathcal{E}$. $\boldsymbol{X}$ could be about (i) the same worldly matters as E, or (ii) what $\mathcal{E}$ is *showing* us of those worldly matters. Jeffrey takes the first option. He wants in the candlelight case to update on a distribution over worldly hypotheses B, G, V (blue, green, violet). Ramsey we've represented as going the other way, dividing his credences between hypotheses $\mathcal{R}[\varphi]$ and $\mathcal{R}[\kappa]$ about what the world is memorially showing of itself via $\mathcal{R}$. (This corresponds in the Jeffrey case to assigning probabilities to hypotheses $\mathcal{S}[\beta]$, $\mathcal{S}[\gamma]$, $\mathcal{S}[\nu]$ (seeing blue, green, violet) about the colors being seen.) Ramsey's approach (conditioning on X^is that speak to what is remembered, or seen) has advantages, it turns out. The work we were trying to do with $\mathrm{p}(\dot{\epsilon})$, the probability of $\dot{\epsilon}$ obtaining, is better done by $\mathrm{p}(\mathcal{O}[\dot{\epsilon}])$, the probability of being presented with $\dot{\epsilon}$ by the event bringing us E.

Why would Ramsey-type distributions be better? $\mathrm{p}(X^i)$ has got to be available pre-update; subcase probabilities are what J-CON expects to be given. But, we can give it only what observation has given us. $\dot{\epsilon}$'s probability will have to be made available by $\mathcal{O}$, then, if it is to play the subcase role. If t-maker probabilities had to be *calculated* or *figured out*, then J-CON, which starts where the calculation leaves off, would be foisting on us work that our update rule ought to be doing. Numbers should not have to be handed over beforehand that it takes update-type work to find.

So, then: Jeffrey-cells can be $\dot{\epsilon}$-based only if $\mathrm{p}(\dot{\epsilon})$ is made available by observation.[19] I don't doubt that probabilities can be grokked; they just come to us when we observe that E. But what are the grokked probabilities *of*? Hypotheses about the external world, e.g., *The tablecloth is green*? One *sees* the tablecloth on this view as .7-probably green. Or hypotheses about observation, to the effect that $\mathcal{O}$ puts this or that external-world fact before us. I *take* it to be .7 probable that $\mathcal{S}$ presents a green tablecloth. The difference may seem slight, and perhaps is in the tablecloth case. But the "taking" model is more powerful and of wider application. Learning events $\mathcal{O}$ are never clearly *not* involved in the alternatives whose probabilities we grok, and there is sometimes no way around them.

Take for example our determination in ONESEEN that $\mathrm{p}(\alpha)$=3/4. Al-heads did not *look* 3/4 probable; we were just as likely to have been looking at Betty-heads (β). A fact that *was* definitely seen is Lefty-heads (λ). But it was seen as half-likely to be Al-heads, not 3/4. 3/4 was reached by calculation. To $\mathrm{p}(\alpha\text{-seen})$=1/2 we added the probability (1/4) of an *unseen* coin landing heads which turned out to be Al's. $\mathrm{p}(\alpha)$ was figured out on the basis of $\mathcal{S}[...]$-probabilities made available by observation:

$$\begin{array}{lr} 1.\ \mathrm{p}(\mathcal{S}[\lambda]) = 1 & \text{obvious} \\ 2.\ \mathrm{p}(\mathcal{S}[\alpha]) = 1/2 & \mathrm{p}(\lambda \equiv \alpha) = 1/2 \\ 3.\ \mathrm{p}(\alpha) = \underbrace{\mathrm{p}(\alpha|\mathcal{S}[\alpha]) \times \mathrm{p}(\mathcal{S}[\alpha])}_{1\times 1/2} + \underbrace{\mathrm{p}(\alpha|\bar{\mathcal{S}}[\beta]) \times \mathrm{p}(\bar{\mathcal{S}}[\beta])}_{1/2\times 1/2} = 3/4 & \alpha \text{ obtains seen or unseen} \end{array} \tag{14}$$

[18] Subject to the condition that $\Sigma_i\ \mathrm{p}(\epsilon^i)$ - $\Sigma_{i\neq j}\ \mathrm{p}(\epsilon^i\epsilon^j) + \Sigma_{i\neq j\neq k}\ \mathrm{p}(\epsilon^i\epsilon^j\epsilon^k)$ - $\Sigma_{i\neq j\neq k\neq l}\ \mathrm{p}(\epsilon^i\epsilon^j\epsilon^k\epsilon^l) + ... = 1$.

[19] See [24], [26], [4] on probabilities figuring in an observation's representational content vs the observation serving as a basis for probabilistic surmises. Morrison raises the interesting possibility of an experience representing that one should where *it's* concerned be 70% confident that the cloth is violet.

These kinds of transitions are not supposed to be a prequel to update; if we're at the point of wringing real-world probabilities out of observations, the update process is already under way. (14) moreover is how α's probability has got to be reached. There just is no direct, non-calculational, route to p(α) = 3/4 from *Here is something with half a chance of being α.*

None of this can be blamed on J-CON; it's just that we have to be careful about X's subject matter. One thinks first of **the truth in E**, which groups worlds by the $\dot{\epsilon}$s obtaining in them. But p($\dot{\epsilon}$) remains oftentimes to be determined. X should be about the prior something that does the determining. That we've said is the event $\mathcal{E}$ bringing us E; let's try then **the truth in $\mathcal{E}$**. The two run in many ways parallel:[20]

	groups worlds by	w goes with v iff
the truth in E	how E is true	the same ϵs hold
the truth in $\mathcal{E}$	how $\mathcal{E}[E]$ is true (how E's truth is presented)	the same $\mathcal{E}[\epsilon]$s hold (the same ϵs are presented)

Table 3: Two partitions

The truth in $\mathcal{E}$ has the advantage of being relatively accessible. I'm a better judge of what I am currently apprehending than of which facts are out there to be apprehended. $\alpha\beta$'s probability is anyone's guess in ONESEEN, but p($\mathcal{S}[\alpha\beta]$) is clearly 0; I know myself to be looking at a single coin. p($\mathcal{S}[\alpha]\vee\mathcal{S}[\beta]$) is just as clearly 1, and now we are off and running. $\mathcal{S}$ can't in the relevant sense mean ϵ unless ϵ obtains (as Grice said about smoke and fire), therefore p(α)≥p($\mathcal{O}[\alpha]$). p(α) adds to p($\mathcal{S}[\alpha]$) the probability of Al's coin landing heads unobserved. That's 1/4, so p(α)=1/2+1/4=3/4. The more accessible subject matter gives us a line of sight on the less accessible one.

10 ACTIVE LEARNING (3): MEANINGS

A familiar complaint about Jeffrey update is that it makes observation into a black box from which probability distributions drop by magic.[21] There's no telling, or even asking, whether a distribution $\vec{X}$ is licensed by $\mathcal{O}$, for J-CON isn't told about $\mathcal{O}$; learning events are the I-know-not-what from which $\vec{X}$ somehow results. This is troubling if, as we've been saying, t-maker probabilities have to be figured out. Figuring can be done well or badly; one may well find it suspicious that the results have to be taken on faith.[22] A decent update rule ought to show its work.

There is something right about this — but something is being missed. True, $\mathcal{O}$ is not presented to J-CON as $\vec{X}$'s source. That doesn't mean it is not presented at all. $\vec{X}$ might be *about* $\mathcal{O}$, $\mathcal{O}$ figuring in the propositions that probability is distributed over. This was the Ramseyian proposal at the end of section 9. $\mathcal{O}$ is a window, as we're thinking of it, on the world of truthmakers. One sees (senses, detects) *something* through it that gives one knowledge that E, but what? No window is perfect, and some ϵs might be off in the distance, partly occluded, or hard to tell apart. Some amount of guesswork is called for, but of a familiar kind. Asking which ϵ one is (perhaps confusedly) glimpsing is like wondering who that is in a photograph, where a remark might be coming from, or what those spots might mean.

What sorts of things *can* be presented? Individual truthmakers, certainly. $\mathcal{S}$ in BK_1 is an event either of α-seeing or β-seeing. Conjunctions of truthmakers too, as when α and β are seen one after the other in BK_2. The list can't end there, however, for there are other ways of learning E than being presented with an ϵ. The *Neither? No!* sign in BK_0

[20] $\mathcal{E}[E]$ is short for $\mathcal{E}[\epsilon^1\vee\epsilon^2\vee...\vee\epsilon^m]$. It says that $\vee_i\epsilon^i$ is presented. Dt-makers (disjunctions of t-makers) are discussed in the next section.

[21] "Jeffrey assumes that somehow perceptual experiences, together with the agent's previous epistemic state, cause new credences for certain statements. But, "patterns of irritation of our sensory surfaces aren't reasons or evidence for any of our beliefs, any more than irritation of the mucous membrane of the nose is a reason for sneezing" (Jeffrey 1968: 176). ... [T]he input parameter describes what he takes to be the output of a perceptual-psychological process: the post-experience credence for these statements. [These] are compulsory in the sense that "I cannot decide to have a different degree of belief in the proposition, any more than I can decide to walk on air" (Jeffrey 1968: 176). "For this reason, Jeffrey conditionalization (exclusively) prescribes how to adjust one's other credences to the compulsory post-experience credences ..." ([5], 2802-3)

[22] Crime labs are not supposed to throw away the samples on which rely in court.

testifies to a disjunction of ϵs. Contact is made with $\alpha\vee\beta$ but not α or β. Or, $\mathcal{O}$ might pick up on Al and Betty throwing the *same*, period: $\alpha\beta\vee\bar{\alpha}\bar{\beta}$. Nothing more definite was observed, even through a glass darkly.

E's t-makers $\dot{\epsilon}$ are its truthmakers ϵ closed under conjunction. Its *dt*-makers $\mathring{\epsilon}$ take this one step further; they're its t-makers closed under disjunction.[23] They share with t-makers and truthmakers the properties of (i) guaranteeing E's truth, and (ii) containing nothing irrelevant to E's truth. They differ in not always being *categorical* — in presenting us sometimes with options. $\mathcal{O}$'s possible meanings are drawn from the class of dt-makers $\mathring{\epsilon}$ of the E that is learned.

The weakest $\mathring{\epsilon}$ (the minimal wholly relevant guarantor) is $\vee_i\epsilon^i$. $\mathcal{O}$ has got to mean IT, we assume — that E is true somehow or other — to be a vehicle for E-learning. The rest represent between them all possible ways of elaborating, or expanding on, that minimal $\mathring{\epsilon}$. Elaboration involves either lopping off disjuncts, or strengthening them with additional ϵs, or both at once.[24] The $\mathring{\epsilon}$s form between them a bounded lattice with $\vee_i\epsilon^i$ (= $\mathring{\epsilon}_\vee$) at the bottom and $\wedge_i\epsilon^i$ ($\mathring{\epsilon}_\wedge$ at the top.[25] So, in the three-truthmaker case:

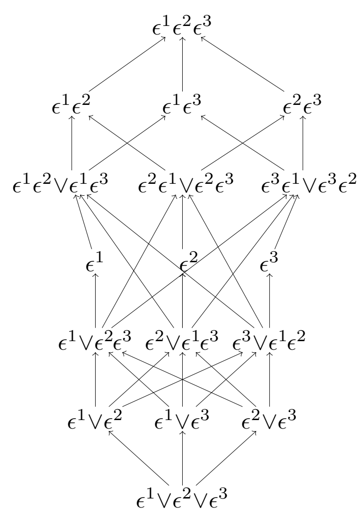

Figure 5: E's dt-makers $\mathring{\epsilon}$ = $\mathcal{O}$'s possible meanings

These are the things an event $\mathcal{O}$ of E-learning can mean. Meaning is downward-closed: if $\mathcal{O}[\mathring{\epsilon}]$, then it means as well all E's weaker dt-makers $\mathring{\chi}$. $\mathcal{O}[\mathring{\epsilon}]$-based cells are as a result not partitional; they overlap on worlds where $\mathcal{O}$ means more than one thing.[26] But we can count worlds equivalent in which $\mathcal{O}$ means the *same* things. This if meanings are downward-closed is the same as grouping by strongest, or total, meaning.

$$\begin{array}{lll} \mathcal{O}[\mathring{\epsilon}] & =_{Df} & \mathcal{O} \text{ means } \mathring{\epsilon} \text{ possibly inter alia.} \\ \mathcal{O}[\![\mathring{\epsilon}]\!] & =_{Df} & \mathcal{O} \text{ means } \mathring{\epsilon} \text{ in toto. that is, } \forall \text{ dt-makers } \mathring{\chi}\text{: } \mathcal{O}[\mathring{\chi}] \text{ iff } \mathring{\epsilon} \vDash \mathring{\chi}. \end{array} \tag{15}$$

$<\mathcal{O}[\mathring{\epsilon}]>$ cannot be relied on to partition the $\mathcal{O}$-worlds, but $<\mathcal{O}[\![\mathring{\epsilon}]\!]>$ comes through every time. Probability is going to be distributed over hypotheses of the form, $\mathcal{O}$ has thus and such as its total meaning

[23] The $\mathring{\epsilon}$s are all and only disjunctions (possibly 1-ary) of conjunctions (possibly 1-ary) of ϵs.

[24] There is nothing in $\vee_i\,\epsilon^i$ about truthmakers *not* holding, a fortiori nothing capable of being further elaborated. $\epsilon^2\bar{\epsilon}^3$ is not a dt-maker for E since it "expands" on something $\vee_i\epsilon^i$ never said.

[25] E's dt-makers correspond in Fine's system to $\vee_i\epsilon^i$'s "exact entailers" closed under conjunction ([12]). A disjunction is exactly entailed by its sub-disjunctions.

[26] $\mathcal{O}[\mathring{\epsilon}_\vee]$'s cell includes all the others; $\mathcal{O}$ means $\mathring{\epsilon}_\vee$ everywhere.

11 ACTIVE CONDITIONING

Bayes' rule has no idea what may have occurred "under the surface" to lift p(E) to 1. It has no other option but capitulate to this ignorance by leaving p($\dot{\epsilon}|E$) at its former value. J-CON takes an important first step towards correcting this by allowing the probabilities of the $\dot{\epsilon}$s to be supplied directly as inputs (section **8**). But Jeffrey expects (i) a partition, which the $\dot{\epsilon}$s normally aren't (section **9**). and (ii) probabilities to be already available that remain in fact to be determined (section **10**).

Learning events to the rescue. We update not on the E that was learned, but on the event $\mathcal{O}$ of learning it — more precisely, on its total meaning as best we can judge. If we had it in hand, $p_n(H)$ would be $p(H|\mathcal{O}[\![\text{IT}]\!])$. As it is, we treat $p(H|\mathcal{O}[\![\mathring{\epsilon}]\!])$ as a random variable ($\mathring{\epsilon}$ itself as a function from worlds to dt-makers) and guesstimate its value in the usual way:[27]

$$\text{A-CON} \qquad p_n^{\mathcal{O}}(H) = \Sigma_{\mathring{\epsilon}}\ p_n(\mathcal{O}[\![\mathring{\epsilon}]\!]) \times p_o(H|\mathcal{O}[\![\mathring{\epsilon}]\!]) \tag{16}$$

This makes precise the dependence posited earlier of proper credence in H on $\mathcal{O}$'s probability of meaning this or that. $p_n^{\mathcal{O}}(H)$ is (our best guess as to) H's probability conditional on the true $\mathcal{O}[\![\mathring{\epsilon}]\!]$.

The switch to total meanings gets us a partition, but at what costl? Underpinnings p($\mathring{\epsilon}$) were obtained (in sections **5**ff.) from backstories. Backstories gave the probabilities of *inter-alia*-meaning assignments $\mathcal{O}[\mathring{\chi}]$, not the total-meaning assignments $\mathcal{O}[\![\mathring{\epsilon}]\!]$ we are using now. The old distribution is recoverable from the new, however: $p_n(\mathcal{O}[\mathring{\chi}]) = \Sigma_{\mathring{\epsilon} \vDash \mathring{\chi}}\ p_n(\mathcal{O}[\![\mathring{\epsilon}]\!])$. $\mathcal{O}$'s probability of meaning $\mathring{\chi}$ inter alia is the summed probabilities of $\mathring{\epsilon}$s that imply $\mathring{\chi}$ being meant in toto. The exact same worlds are involved in both cases: $|\mathcal{O}[\mathring{\chi}]| = \cup_{\mathring{\epsilon} \vDash \mathring{\chi}} |\mathcal{O}[\![\mathring{\epsilon}]\!]|$.[28]

12 TOPIC-SENSITIVITY

It was suggested in section **2** that E_is true in the same worlds could come apart evidentially. A nice idea, but we have yet to see any actual examples. A single E licensing this update or that, depending on how it is learned: *that* we've seen. A-CON derives the update differences from differences in the likely meanings of the learning events $\mathcal{O}_i$:

$$\begin{array}{l} E \text{ learned through } \mathcal{O}_1 \text{ boosts } p(H) \text{ more than } E \text{ learned through } \mathcal{O}_2 \\ \qquad \text{because } \mathcal{O}_1 \text{ is likelier to mean } \mathring{\epsilon} \text{ (which supports } H\text{): } p(\mathcal{O}_1[\mathring{\epsilon}]) > p(\mathcal{O}_2[\mathring{\epsilon}]). \end{array} \tag{17}$$

From (17), however, it's a short step to E_1 coming apart evidentially from E_2. For $\mathcal{O}_1$ could be an event of E_1-learning — an $\mathcal{E}_1$ — and $\mathcal{O}_2$ an event of E_2-learning —an $\mathcal{E}_2$. And it could be the E_is' semantic differences (the fact that one expresses $\boldsymbol{E}_1$ and the other $\boldsymbol{E}_2$) that explain why events of E_1-learning are likelier to mean $\mathring{\epsilon}$:

$$\begin{array}{l} E_1 \text{ as a rule boosts } p(H) \text{ more than } E_2 \\ \qquad \text{because } \mathcal{E}_1\text{-type events are likelier to mean } \mathring{\epsilon} \text{ (which supports } H\text{) : } p(\mathcal{E}_1[\mathring{\epsilon}]) > p(\mathcal{E}_2[\mathring{\epsilon}]). \end{array} \tag{18}$$

Why would sentences get into the business of signaling how they were learned? This is a large question on which a lot has been written.[29] The key for us is that observations are not for private consumption; they are made to be shared. Audiences can't be left to guess at the probability of particular $\mathring{\epsilon}$s having been spotted, or they won't know the observation's upshot for H. Were UN-, ONE-, and TWOSEEN reported the same way, any number of credences could be appropriate in *Both threw heads*. Better surely if there were different styles of report, such that the $\mathring{\epsilon}$s we were hearing *about* could be guessed from the one we heard. English to some extent obliges:

[27] A-CON bears comparison to Konek's K-CON, in [19], to which I'm indebted.

[28] Proof. [$\supseteq$] $\mathcal{O}[\![\mathring{\epsilon}]\!] \Rightarrow \mathcal{O}[\mathring{\epsilon}] \Rightarrow \mathcal{O}[\mathring{\chi}]$ (by downward-closure). [$\subseteq$] To be the strongest thing meant in w, $\mathring{\epsilon}$ must imply also-meant $\mathring{\chi}$. □

[29] See the *Wikipedia* page on "evidentials," [1], and [27].

R_0. *Not: neither coin landed heads* conveys that p($\mathcal{H}[\alpha]$)=p($\mathcal{H}[\beta]$)=0 (use this to report on UNSEEN)
R_1. *A certain coin landed leads* conveys that p($\mathcal{H}[\alpha]$), p($\mathcal{H}[\beta]$)>0[30] (... ONESEEN)
R_2. *A certain number of coins landed heads* conveys that p($\mathcal{H}[\alpha]$), p($\mathcal{H}[\beta]$), p($\mathcal{H}[\alpha\beta]$)>0.[31] (... TWOSEEN)

These are intensionally equivalent, but not evidentially.[32] *Both landed heads* is better supported by R_2 than by R_1, and R_1 supports it better than R_0. R_2's advantage over the others appears to be rooted in truthmaker differences. Two landing heads ($\alpha\beta$) is *allowed* by *A certain coin landed heads*, even welcomed, but overkill from a truthmaker perspective; it's not among the possibilities specifically envisaged. Two landing heads does make *A certain number landed heads* true.

More generally let S and T be co-intensional sentences about **s** and **t**, learned via $\mathcal{S}$ and $\mathcal{T}$ respectively. Here are the lines of influence by which topical differences make (or can make) for evidential differences:

s ≠ **t**
⇒ the σ ≠ the τs
⇒ the $\mathring{\sigma}$s ≠ the $\mathring{\tau}$s
⇒ $\Sigma_{\mathring{\sigma}}$ (p$_n$($\mathcal{S}[\![\mathring{\sigma}]\!]$)×p$_o$($H|\mathcal{S}[\![\mathring{\sigma}]\!]$)) floats free of $\Sigma_{\mathring{\tau}}$ (p$_n$($\mathcal{T}[\![\mathring{\tau}]\!]$×p$_o$($H|\mathcal{T}[\![\mathring{\tau}]\!]$))
⇒ conditioning (the A-CON way) on $\vec{\boldsymbol{S}}$ leads potentially to a different credence in H than conditioning on $\vec{\boldsymbol{T}}$
⇒ S and T differ as evidence.

(19)

Suppose for instance that **s** ⊊ **t**, so that S's dt-makers form a proper subset of T's. Should $\mathring{\sigma}$ and $\mathcal{S}[\![\mathring{\sigma}]\!]$ be always neutral on H, S will be H-neutral too (p($H|\vec{\boldsymbol{S}}$)= p(H)). This leaves the new $\mathring{\tau}$s (those not in **s**) and $\mathcal{T}[\![\mathring{\tau}]\!]$s free to probabilify H. And now T according to A-CON supports H (p($H|\vec{\boldsymbol{T}}$)> p(H)) — *unless* each new $\mathring{\tau}$ had zero chance of being presented.[33]

13 HYPERINTENSIONALITY

A certain number of coins landed heads (R_2) supports *Both did* more than *A certain coin landed heads* (R_1) does. This is enabled in the first instance by $\vec{\boldsymbol{R}}_1$'s distinctness from $\vec{\boldsymbol{R}}_2$. But it pays to dig deeper. They're distinct because $\boldsymbol{R}_1 \neq \boldsymbol{R}_2$, because $\mathbf{r}_1 \neq \mathbf{r}_2$, because $\alpha\beta$ (both coins landing heads) is a truthmaker only for R_2. One can't account in this way for R_1's advantage over R_0. They appear to agree in their truthmakers, differ only in which $\mathring{\rho}$s are dangled before us as maybe encountered. Here is a cleaner case, from Krämer ([20]).

Your ankle hurts and you have it examined. *It is sprained, or sprained and broken*, the doctor says ($A \vee AB$). *Sprained* had been half likely (p$_o$(A)=.5), and you're now certain of it since A is implied by $A \vee AB$. *Broken* too had been half likely, and now becomes likelier. But p(B) rises less—only, let's say, to .75. Had you been told instead, *Your ankle is sprained, or sprained and contused*, it would have been *Contused* (C) that went from .5 to .75. Both would have stayed at .5, had the doctor said simply *Sprained*. (Assume independence; p($B|A$)=p(B) and p($C|A$) = p(C).)

This is unexpected from a classical perspective. B's new probability of .75 is not obtained by Bayes-updating on $A \vee AB$, for p($B|A \vee AB$)=p($B|A$)=p(B)=.5 again. ($|A \vee AB|$ = $|A|$ = $|A \vee AC)|$.)

[30] Also that p($\exists x\ \mathcal{H}$[heads(x)])=1 ([14]), [15] [21], [10]
[31] And that p($\exists x\ \mathcal{H}$[heads(x)])=1.
[32] Our full evidence, if one of these reports were actually made, would include the specific words uttered, the manner of delivery, and more. This gives the Bayesian an opening. Fine-grained hearing events $\underline{\mathcal{H}}$ could be posited, with everything relevant to the probability that $\underline{\mathcal{H}}[\alpha]$ written into their essence. p$_o$($\alpha|\underline{\mathcal{H}}$) could take the place of p$_n$($\mathcal{H}[\alpha]$), if one could tell with certainty which $\underline{\mathcal{H}}$ had occurred and knew beforehand α's probability conditional on each. Some may find this attractive. It has to me a retrograde, positivist, feel, harking back to the days of inputting appearances rather than how things appear to be. (This ends the series of potshots announced in note 11.)
[33] Why speak in that case of learning that T, as opposed to S?

Griceans will say that context pragmatically strengthens E to an E^+ that includes E's implicatures.[34] $p_n(B)$ isn't $p_o(B|A{\vee}AB)$ because $A{\vee}AB$ is not all we've learned. But, $A{\vee}AB$'s only clear implicature is that the ankle might be sprained *and* broken ($\Diamond AB$).[35] $p_o(AB)$ was already at 1/4 ($p_o(A)$=$p_o(B)$=1/2 and the two are independent). Shouldn't that be enough for $p_o(\Diamond AB)=1$?[36] If so then $(A{\vee}AB)^+$ doesn't probabilify B any more than $A{\vee}AB$ does:

$$\begin{aligned}
&p_o(B|\,(A\vee AB)^+) \\
&= p_o(B|\,(A\vee AB)\wedge\Diamond AB) \\
&= p_o(B|A\vee AB) && \text{.... } p_o(\Diamond AB)=1 \\
&= p_o(B|A) && \text{.... } |A|=|A\vee AB| \\
&= p_o(B) && \text{....} B \text{ is independent of } A \\
&= 1/2.
\end{aligned} \tag{20}$$

What *would* raise the probability of a broken ankle, if not $\Diamond AB$? That's what we're coming to. But $\Diamond\circledcirc[\alpha\beta]$ seems worth a look. That one might in learning $A{\vee}AB$ be *picking up* on $\alpha\beta$ says more for B than $\alpha\beta$'s persistence as a possibility.

$A{\vee}AB$ never supports B, on the pragmatic approach, unless context steps in to strengthen it. Our roughly opposite view is that it always supports B, unless context steps in to weaken it.[37] $p_n(H)$ for A-CON is H's expected probability conditional on "the true $\circledcirc[\![\mathring{\epsilon}]\!]$." The true $\circledcirc[\![\mathring{\epsilon}]\!]$ when $E_1 = A$ is $\circledcirc[\![\alpha]\!]$ (α is the only thing $\circledcirc$ *can* mean). So, assuming that B is independent also of $\circledcirc[\![\alpha]\!]$,

$$p_n(B) = \underbrace{p_n(\circledcirc[\![\alpha]\!])}_{1} \times \underbrace{p_o(B|\circledcirc[\![\alpha]\!])}_{1/2} = 1/2 \tag{21}$$

Compare now $E_2 = A{\vee}AB$. Its dt-makers (up to equivalence) are α and $\alpha\beta$. Those are the only total meanings available, so $\vec{\boldsymbol{E}}_2$ is a distribution over $\circledcirc[\![\alpha]\!]$ and $\circledcirc[\![\alpha\beta]\!]$; they're assigned respectively k and 1-k.

$$p_n(B) = \underbrace{p_n(\circledcirc[\![\alpha]\!]) \times p_o(B|\circledcirc[\![\alpha]\!])}_{k\times 1/2} + \underbrace{p_n(\circledcirc[\![\alpha\beta]\!]) \times p_o(B|\circledcirc[\![\alpha\beta]\!])}_{(1-k)\times 1} = 1 - k/2 \tag{22}$$

1 - k/2 exceeds 1/2 unless k=1, that is, $p_n(\circledcirc[\![\alpha]\!]) = 1$ and $p_n(\circledcirc[\![\alpha\beta]\!]) = 0$. The only way then for $A{\vee}AB$ NOT to support B is for its delivery vehicle $\circledcirc$ to have NO chance of testifying to $\alpha\beta$.[38]

That can happen. But it takes work to arrange, e.g., one observes that A, and then picks AB arbitrarily out of a long list of possible second disjuncts. $A{\vee}AB$ just by its content — just by virtue of expressing $\boldsymbol{A{\vee}AB} = (\mathbf{A}, \{\alpha, \alpha\beta\})$ — raises for A-CON the question of what $p_n(\circledcirc[\alpha\beta])$ should be. B-favoring answers outnumber B-neutral ones by $2^{\aleph_0}$ to 1. $p_n(Broken)$ is 3/4, as suggested above, if $\circledcirc$ is undecided between α and $\alpha\beta$; its total meaning has half a chance of either. For 1-$p_n(\circledcirc[\![\alpha]\!])$/2 is in that case $1 - \frac{1}{2}/2 = 1 - \frac{1}{4} = \frac{3}{4}$. But B is supported to *some* extent as long as $\circledcirc$ bringing news with $\alpha\beta$ in its subject matter is not an absolute fluke —the alternative to absolute-fluke-hood being, a sign that $\alpha\beta$ is among the things $\circledcirc$ could mean.

To be built out of different atoms ({A, B} vs {A, C} is one way for PC-equivalents to come apart evidentially. A-CON predicts this and provides a mechanism. What about PC-equivalents with the *same* atoms? Let E for instance be $A{\equiv}B{\vee}B{\equiv}C$. E can be rendered up to logical equivalence as

[34] Manner implicatures, presumably, since $A{\vee}AB$ and A "say the same."
[35] Mandelkern [23] puts $\Diamond AB$ in the literal content.
[36] The example is easily tweaked if not.
[37] By neutralizing the second disjunct.
[38] "(22) assumes that $p_o(B|\circledcirc[\![\alpha]\!]) = 1/2$. But it could be on account of β's absence that only α was observed. So 1/2 is too big." This is fair enough. But $p_o(B|\circledcirc[\![\alpha]\!])$ only has to exceed 0 for $A{\vee}AB$ to support B more than A does. ("Adverse inferences" come up again in the next section.)

E_1: $AB \vee \bar{A}\bar{B} \vee BC \vee \bar{B}\bar{C}$
E_2: $AB \vee \bar{B}\bar{C} \vee \bar{A}C$
E_3: $\bar{A}\bar{B} \vee BC \vee A\bar{C}$

They are not *exactly* equivalent since each E_i has its own stock of truthmakers (given by their disjuncts). E_1 is the most faithful in this respect to E. Both are made true by $\alpha\beta$, $\bar{\alpha}\bar{\beta}$, $\beta\gamma$, and $\bar{\beta}\bar{\gamma}$. E_3 however has $\alpha\bar{\gamma}$ as a truthmaker and E_2 has $\bar{\alpha}\gamma$.

Though true for their own reasons, the E_s involve the same atoms. They make a good test case, then, for the idea that PC-equivalents can differ evidentially due to how their (shared) atoms are put together. E_1's truthmakers are neutral on $A{\equiv}C$ (H) if A-C are *Al threw heads* etc; $\mathrm{p}_o(A{\equiv}C|\alpha\beta) = 1/2$, the same as $\mathrm{p}_o(A{\equiv}C)$. $\alpha\bar{\gamma}$ and $\bar{\alpha}\gamma$ are inconsistent with $A{\equiv}C$, giving E_2 and E_3 the appearance of favoring $A{\not\equiv}C$.

What does A-CON say? Al, Betty, and Cleo have tossed their coins. The two we can see (it's not clear whose they are) both show heads. $\mathrm{p}(A{\equiv}C|\vec{\boldsymbol{E}}_1)$ according to A-CON is

$$\underbrace{\mathrm{p}_n(\circledcirc[\![\alpha\beta]\!])}_{1/4}\times\underbrace{\mathrm{p}_o(A \equiv C \mid \circledcirc[\![\alpha\beta]\!])}_{1/2} + \underbrace{\mathrm{p}_n(\circledcirc[\![\bar{\alpha}\bar{\beta}]\!])}_{1/4}\times\underbrace{\mathrm{p}_o(A \equiv C \mid \circledcirc[\![\bar{\alpha}\bar{\beta}]\!])}_{1/2}$$
$$+ \underbrace{\mathrm{p}_n(\circledcirc[\![\beta\gamma]\!])}_{1/4}\times\underbrace{\mathrm{p}(A{\equiv}C|\circledcirc[\![\beta\gamma]\!])}_{1/2} + \underbrace{\mathrm{p}_o(\circledcirc[\![\bar{\beta}\bar{\gamma}]\!])}_{1/4}\times\underbrace{\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\bar{\beta}\bar{\gamma}]\!])}_{1/2} = 1/2 \tag{23}$$

($\mathrm{p}_n(\circledcirc[\![\alpha\beta]\!])$=1/4 by indifference; $\circledcirc$ might be picking up on any of four scenarios of which $\alpha\beta$ is one. $\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\alpha\beta]\!])$ = 1/2 because that's $A{\equiv}C$'s unconditional probability (it holds in two of four cases), and $\circledcirc[\![\alpha\beta]\!]$ is irrelevant to whether $A{\equiv}C$.) $\mathrm{p}(A{\equiv}C|\vec{\boldsymbol{E}}_2)$ meanwhile is

$$\underbrace{\mathrm{p}_n(\circledcirc[\![\alpha\beta]\!])}_{1/3}\times\underbrace{\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\alpha\beta]\!])}_{1/2} + \underbrace{\mathrm{p}_n(\circledcirc[\![\bar{\beta}\bar{\gamma}]\!])}_{1/3} \times \underbrace{\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\bar{\beta}\bar{\gamma}]\!]}_{1/2} + \underbrace{\mathrm{p}_n(\circledcirc[\![\bar{\alpha}\gamma]\!])}_{1/3}\times\underbrace{\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\bar{\alpha}\gamma]\!])}_{0} = 1/3 \tag{24}$$

($\mathrm{p}_n(\circledcirc[\![\alpha\beta]\!])$ etc are 1/3 this; $\circledcirc$ might be picking up on any of three facts, corresponding to E_2's three disjuncts. $\mathrm{p}(A{\equiv}C|\circledcirc[\![\alpha\beta]\!])$=1/2 by virtue of $\circledcirc[\![\alpha\beta]\!]$'s irrelevance.) What pulls the sum down from 1/2 to 1/3 is that $\mathrm{p}_o(A{\equiv}C|\circledcirc[\![\bar{\alpha}\gamma]\!])$=0, $\bar{\alpha}\gamma$ entailing $A{\not\equiv}C$. $\mathrm{p}(A{\equiv}C|\vec{\boldsymbol{E}}_3)$ is 1/3 too; the tie-breaker this time is $\alpha\bar{\gamma}$.

Now, *all* these statements, E and E_1 included, probabilify $A{\not\equiv}C$.[39] $\mathrm{p}_o(A{\not\equiv}C|E) = 2/3 > 1/2 = \mathrm{p}_o(A{\not\equiv}C)$. A- and B-CON disagree, then, on the proper response when word arrives that $A{\equiv}B{\vee}B{\equiv}C$. Whose advice are we to follow? That depends, as usual, on the delivery vehicle $\circledcirc$: on how E is learned. If passively — $\mathrm{p}_n(\circledcirc[\mathring{\epsilon}]) > 0$ only when $|\mathring{\epsilon}| = |E|$ — then the rules agree: $\mathrm{p}_n(A{\not\equiv}C)$=2/3. If $\mathrm{p}_n(\circledcirc[\![\alpha\beta]\!])$= $\mathrm{p}_n(\circledcirc[\![\beta\gamma]\!])$=1/2 (the learning is active) then $\mathrm{p}_n(A{\not\equiv}C)$ should be 1/2 (by (23)).

"Active learners" in the everyday sense have a *style* of learning that they bring to every pedagogical occasion. That is not how we're using the term in this paper. A given E can be learned either way (actively or inertly); one employs whichever rule is appropriate. The abstract, decontextualized, question, "Does $A{\equiv}B{\vee}B{\equiv}C$ support $A{\not\equiv}C$?" has no clear answer. If you feel yourself pulled both ways on it, this could be why. One needs to know how the information came in.

14 A PRUSSIAN PUZZLE

> Infinitely many perfectly rational people each roll an independent and fair die. Things are arranged so they can't see how the rolls come out, and yet each is asked to guess if her roll came out six. If a person guesses correctly, she gets a treat; otherwise, she gets an electric shock ([28]:98).

[39] $A{\not\equiv}C$ implies them.

They all of course guess against having rolled six. To have rolled one, two, three, four, or five is five times as likely. An angel now appears who has looked at most (if not all) of the dice. Her report:

K Almost everyone rolled six — all but finitely many of you. (25)

Should they guess again? I am part of this group, as it happens; consider the matter from my perspective. *Almost all of us* is a lot! Never mind for now whether it's enough to lift p(*I rolled six*) to 1/2 (so that I withdraw the guess). It definitely shouldn't stay at 1/6. That would be to reckon *Almost all of us are* ... irrelevant to *I am* .... The angel might as well have said *A dog is barking*. Wait, it was almost *none* of us that rolled six? One report is as irrelevant as the other.

Strange as it seems, a case can be made that they *are* irrelevant. A single die is not enough to make the difference between infinitely many and finitely many. Almost all of *US* rolled six in the exact same worlds as

L Almost all of *THEM* rolled six, (26)

THEM being the rollers other than me. *L* is irrelevant to *I rolled six* (*M*), for it ignores me and the dice are independent and fair.[40] *K* is co-intensional for similar reasons with

N Almost all of us rolled *STIX*, (27)

where to roll stix is to roll six if not me, otherwise three. *N* would seem if anything to support *I didn't roll six*, since rolling stix amounts in my case to rolling three.

Pruss's puzzle may not look a whole lot like the sprained ankle case, but bear with me. $A \vee AB$ can be paired up with *K* by putting *L* and *M* in for *A* and *B*. The sentence $L \vee LM$ that results is exactly equivalent to *Almost everyone rolled six*; both are made true by all and only attributions of six-rolling to cofinitely many of us. If $A \vee AB$ can support *B* despite its co-intensionality with *A*, then $L \vee LM$ should be able to support *M* despite its co-intensionality with *L*. *Almost everyone rolled six*, with *LM* as a (virtual) disjunct, should be able to support *M* as well.

Let's make sure we have got this right, that *K* ($\cong L \vee LM$) really does stand to *L* as $A \vee AB$ to *A*. $A \vee AB$ has two sorts of truthmaker, corresponding to its two disjuncts. There are

- those it gets from *A* .... there's just the one, α; and
- those it gets from *AB* .... again just the one, $\alpha\beta$.

$L \vee LM$ has two sorts as well. There are

- those it gets from *L* .... one (λ) per cofinite bunch of rollers ≠ me, to the effect that *they* rolled six
- those it gets from *LM*... one ($\lambda\mu$) for each λ, to the effect that *they* rolled six and me too.

$A \vee AB$ makes itself *B*-relevant by having truthmakers (like $\alpha\beta$) that entail *B*. (Its other truthmaker α is *B*-independent.) $L \vee LM$ makes itself *M*-relevant by having truthmakers $\lambda\mu$ that entail *M*, such as ω = everyone rolling six. ($L \vee LM$'s other truthmakers are *M*-independent.) *K* does it by having the exact same truthmakers as $L \vee LM$.

The infinitary case is in one respect more compelling. $A \vee AB$ could be heard as a needlessly complicated way of saying that *A*. *Almost everyone rolled six* is not in anyone's book a gratuitous twist on *Almost everyone distinct from*

[40] Builes ([6]) has a closely related puzzle, discovered independently. He notes a possible connection with subject matter: "I believe the paradox shows that our intuitive conceptions of inadmissible evidence and independent evidence are sensitive to facts about evidential aboutness in an interesting way" (114). Dorr et al respond in [9]. A uniform distribution isn't possible on the assumption of countably many rollers. If this matters (I'm not sure it does), we can let the group be uncountable.

Steve rolled six. The difference at a semantic level is that $A \vee AB$ has a minimal truthmaker with the same intension: $|\alpha| = |A \vee AB|$. its non-minimal truthmakers ($\alpha\beta$) would to that extent not be missed were they dropped. Try that with $L \vee LM$'s truthmakers, or K's and there would be nothing left. They're *all* non-minimal.[41] You would have to hack away with malice aforethought at $\boldsymbol{K}$ ((using an axe "with my name on it") to stop it supporting M.

Is it a problem that A-CON uses dt-makers $\mathring{\epsilon}$ rather than truthmakers ϵ? These come to the same when $E = K =$ *Almost everyone threw six* (cofinite subset-hood is preserved under finite union and intersection). Estimating $p(M|\mathbb{O}[\![\kappa]\!])$ *is* estimating $p(M|\mathbb{O}[\![\mathring{\kappa}]\!])$. K's truthmakers κ divide we have seen into the $\lambda\mu$s, which have *me* rolling six, and the λs, which ignore me. $p(M|\mathbb{O}[\![\lambda\mu]\!]) = p(M|\mathbb{O}[\mu])$ and $p(M|\mathbb{O}[\![\lambda]\!]) = p(M|\bar{\mathbb{O}}[\mu])$, so we can partition simply on $\mathbb{O}[\mu]$, $\bar{\mathbb{O}}[\mu]$. Our credence in M on learning K is

$$\begin{aligned} &\mathbb{E}[p(\mathbb{O}[\mu])] \times \underbrace{p(M|\mathbb{O}[\mu])}_{1} + \mathbb{E}[p(\bar{\mathbb{O}}[\mu])] \times \underbrace{p(M|\bar{\mathbb{O}}[\mu])}_{m} \\ &\quad = \mathbb{E}[p(\mathbb{O}[\mu])] + m \times (1\text{-}\mathbb{E}[p(\mathbb{O}[\mu])]) = (1\text{-}m) \times \mathbb{E}[p(\mathbb{O}[\mu])] + m \end{aligned} \tag{28}$$

$\mathbb{E}[p(\mathbb{O}[\mu])]$, the expected probability of my being among the cofinite multitudes (seen to have rolled six), is at least 1/2. (I am not *likelier* among the finite few.) Assuming the worst to make life difficult, let $\mathbb{E}[p(\mathbb{O}[\mu])] = 1/2$, (28) then gives $(1\text{-}m)/2 + m = (m+1)/2$ as my updated credence in having thrown six. That's at least 1/2, up from $p_o(M) = 1/6$. So p(*I threw six*) becomes in the *worst case scenario* three times larger when the news comes in we almost all rolled six.

Suppose the news had been different; I am told instead that *Almost all of THEM rolled six* (L). An angel looking for L-truthmakers λ is not going to bother about μ (which is independent of λ). I assume therefore that my die wasn't checked:[42] $\mathbb{E}[p(\mathbb{O}[\mu])] = 0$. No adverse inferences can be drawn ($\mathbb{O}[\mu]$ wasn't in the cards anyway), so $p(M|\bar{\mathbb{O}}[\mu]) = p(M) = 1/6$. $p_n(M)$ when I'm told that *Almost all the others rolled six* is

$$\mathbb{E}\underbrace{[p(\bar{\mathbb{O}}[\mu])]}_{1} \times \underbrace{p(M|\bar{\mathbb{O}}[\mu])]}_{1/6} = 1/6 \tag{29}$$

Our conclusion so far: *Almost all of THEM rolled six* leaves p(*I rolled six*) at 1/6. *Almost all of US rolled six* by contrast at least triples M's probability.

When does my probability of rolling six *more* than triple? That depends on the value of $m = p(M|\bar{\mathbb{O}}[\mu])$. m should not be as high as 1/6, as we *can* now draw an adverse inference; μ wasn't presented, maybe, because it didn't obtain. So $0 \leq m < 1/6$. The rest depends on how far the angel takes her inquiries. $p(M|\bar{\mathbb{O}}[\mu])=0$ if every last die is checked; μ would have been noticed, had it obtained. Some angels are in the habit, though, of knocking off early, content with their existing haul of ϵs. $p(M|\bar{\mathbb{O}}[\mu]) > 0$ if we're dealing with that kind of angel. My die could be one of those unnoticed sixes.

Almost everyone rolled six retains an evidential advantage over *Almost everyone ELSE rolled six* whatever the angel's work habits. $p_n(M)$ is significantly larger ($1/3 + m/2$ larger) when it's K we learn and L. Our initial bet (against rolling six) will have to be withdrawn, since p_n(*I didn't roll six*) $\leq 1/2$. The change of heart is *not* because I am likelier among the cofinite multitudes than the finite few. That is a tempting assumption, but it wasn't needed to make the case that $p(M|\vec{\boldsymbol{K}}) \geq 1/2$. (*Almost everyone rolled stix* likewise raises the probability of having thrown three to $\geq 1/2$.)

[41] [35]

[42] There's a question quite generally of what qualifies E as a statement of what was learned. Possibly this, that E's truthmakers ϵ are the facts or potential facts between which the learning event is undecided. $\mathbb{O}$ is clear that some ϵ or other holds ($p(\mathbb{O}[\vee_i \epsilon^i]=1)$, but not about any sub-disjunction ($p(\mathbb{O}[\vee_{i \neq j} \epsilon^i]) < 1$) (otherwise the report would have been stronger). L is chosen over K when μ-involving κs were not potentially glimpsed.

15 CONCLUSION

Learning, on a recently popular Bayesian metaphor, is having information "flung" at you by the world ([30], 1155). I like the metaphor and would merely note that "fling" takes plural objects.[43] The world presents us with E, I'm suggesting, by flinging multiple possibilities $\mathring{\epsilon}$ at us, all of them bearing E out in some way.

Multi-messages don't wear their meanings on their sleeves, which makes the update process less deterministic. No algorithm is going to decide for us which $\mathring{\epsilon}$s to take how seriously, and hence which probasition has been learned. But the process is not unconstrained, either. One knows a lot about the likely sources of various sorts of message. And where a learning event $\oslash$ offers no clues — $\mathrm{p}(\oslash[\![\vee_i \epsilon^i]\!]) = \mathrm{p}(\oslash[\![E]\!]) = 1$ — A-CON has us classically updating, not quite on E, but on having learned that E via $\oslash$: $\mathrm{p}_n(H) = \mathbb{E}[\mathrm{p}_o(H|\oslash[\![\mathring{\epsilon}]\!])]$

Still, why upset the applecart in this way? A number of reasons were given. B-CON may turn us down; active learners need not apply. A-CON lets us update on E however it's learned. (No need to avert our eyes from the facts that determine its truth-value.) B-CON lets the intensionality of probability percolate through to evidential relations. A-CON respects the intensionality of probability. but does not allow it to force our hand; co-intensional E_is can differ evidentially by putting different $\mathring{\epsilon}$s forward as the ones potentially in view. B-CON misses intuitive evidential differences, e.g., between *¬No Fs are G*, *A certain F is G*, and *Some number of Fs are G*. A-CON is alive to these differences. It is sensitive even to fluctuations in the evidential value of a fixed E, as the learning event highlights different truthmakers. And then, there's the puzzle of the Prussian dice.

References

[1] A. Y. Aikhenvald. *Evidentiality*. OUP Oxford, 2004.

[2] B.-M. Benjamin. Identity and aboutness. *Journal of Philosophical Logic*, 50(6):1471–1503, 2021.

[3] F. Berto and A. Özgün. Indicative conditionals: Probabilities and relevance. *Philosophical Studies*, 178(11):3697–3730, 2021.

[4] N. Block. If perception is probabilistic, why does it not seem probabilistic? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 373(1755):20170341, 2018.

[5] P. Brössel. Learning from experience and conditionalization. *Philosophical Studies*, 180(9):2797–2823, 2023.

[6] D. Builes. A paradox of evidential equivalence. *Mind*, 129(513):113–127, 2020.

[7] R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, 1962.

[8] C. S. Chihara. The mystery of julius: A paradox in decision theory. *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 80(1):1–16, 1995.

[9] C. Dorr, J. Hawthorne, and Y. Isaacs. Solving a paradox of evidential equivalence. *Mind*, 130(520):1159–1182, 2021.

[10] C. Ebert and S. Hinterwimmer. *Different kinds of specificity across languages*, volume 92. Springer Science & Business Media, 2012.

[11] K. Fine. Truthmaker semantics. In B. Hale, A. Miller, and C. Wright, editors, *A Companion to the Philosophy of Language*, chapter 22, pages 556–577. Wiley, 2017.

[12] K. Fine and M. Jago. Logic for exact entailment. *The review of symbolic logic*, 12(3):536–556, 2019.

[13] M. Gardner. The two children problem, "Mathematical Games". *Scientific American*, 1959.

[14] J. Hintikka. The semantics of "a certain". *Linguistic Inquiry*, 17(2):331–336, 1986.

[15] N. Hornstein. "A certain" as a wide-scope quantifier: a reply to Hintikka. *Linguistic inquiry*, 19(1):101–109, 1988.

[16] R. C. Jeffrey. *The logic of decision*. University of Chicago Press, 1990.

[17] S. Kaufmann. Conditioning against the grain. *Journal of Philosophical Logic*, 33(6):583–606, 2004.

[18] J. Khoo. Probabilities of conditionals in context. *Linguistics and Philosophy*, 39(1):1–43, 2016.

[19] J. Konek. The art of learning. In J. Gendler, Tamar;; Hawthorne, editor, *Oxford Studies in Epistemology*, volume 7. Oxford University Press, 2019.

[20] S. Krämer. Mighty belief revision. *Journal of Philosophical Logic*, 51(5):1175–1213, 2022.

[21] A. Kratzer. Scope or pseudoscope? are there wide-scope indefinites? In S. Rothstein, editor, *Events and Grammar*, pages 163–196. Kluwer, Dordrecht, 1998.

[22] H. Leitgeb. *The stability of belief: How rational belief coheres with probability*. Oxford University Press, 2017.

[43] Schoenfield remarks on this too. She has a different application in mind (a plurality of things learned for sure).

[23] M. Mandelkern. Disjunction and possibility. *Philosophical Perspectives*, 38(1):211–221, 2024.

[24] J. Morrison. Perceptual confidence. *Analytic Philosophy*, 57(1):15–48, 2016.

[25] S. Moss. *Probabilistic knowledge*. Oxford University Press, 2018.

[26] J. Munton. Visual confidences and direct perceptual justification. *Philosophical Topics*, 44(2):301–326, 2016.

[27] S. E. Murray. *The semantics of evidentials*, volume 9. Oxford University Press, 2017.

[28] A. R. Pruss. *Infinity, causation, and paradox*. Oxford University Press, 2018.

[29] F. P. Ramsey and D. Mellor (ed.). *Philosophical Papers*. Cambridge University Press, 1990.

[30] M. Schoenfield. Conditionalization does not (in general) maximize expected accuracy. *Mind*, 126(504):1155–1187, 2017.

[31] R. T. Stewart and M. Nielsen. Conglomerability, disintegrability and the comparative principle. *Analysis*, 81(3):479–488, 2021.

[32] S. Yablo. *Aboutness*. Princeton University Press, 2014.

[33] S. Yablo. Ifs, ands, and buts: An incremental truthmaker semantics for indicative conditionals. *Analytic Philosophy*, 57(3):175–213, 2016.

[34] S. Yablo. Precis of Aboutness. *Philosophical Studies*, 174(3):771–777, 2017.

[35] S. Yablo. Relevance without minimality. In D. Kindermann, P. van Elswyk, and A. Egan, editors, *Unstructured Content*. OUP, 2025.