DOI: 10.1111/phpr.12947

Philosophy and
Phenomenological Research

ORIGINAL ARTICLE

Epistemic probabilities are degrees of support, not degrees of (rational) belief

Nevin Climenhaga

Dianoia Institute of Philosophy,
Australian Catholic University

Correspondence
Nevin Climenhaga, Dianoia Institute of Philosophy, Australian Catholic University.
Email: nevin.climenhaga@acu.edu.au

Abstract

I argue that when we use 'probability' language in epistemic contexts—e.g., when we ask how probable some hypothesis is, given the evidence available to us—we are talking about *degrees of support*, rather than *degrees of belief*. The epistemic probability of A given B is the mind-independent degree to which B supports A, not the degree to which someone with B as their evidence believes A, or the degree to which someone would or should believe A if they had B as their evidence. My central argument is that the degree-of-support interpretation lets us better model good reasoning in certain cases involving old evidence. Degree-of-belief interpretations make the wrong predictions not only about whether old evidence confirms new hypotheses, but about the values of the probabilities that enter into Bayes' Theorem when we calculate the probability of hypotheses conditional on old evidence and new background information.

1 | INTRODUCTION

Orthodox Bayesianism is a theory about *degrees of belief*. Orthodox Bayesians say that degrees of belief should be probabilistically coherent, and perhaps obey other norms as well, such as diachronic conditionalization, the Principal Principle, and the Principle of Indifference. On standard presentations, the more of these norms you accept, the more *objective* a Bayesian you are, whereas the fewer you accept, the more *subjective* a Bayesian you are. But Bayesians of all these sorts share an assumption that degrees of belief, or 'credences', are the primary objects of interest to formal epistemology. Standard presentations call (coherent) credences "subjective probabilities", suggesting that in epistemic contexts 'probability'-language refers to these psychological entities.[1]

I seek to upend this orthodoxy. I will argue that the probabilities we reason about in epistemic contexts are *degrees of support*: mind-independent relations between propositions that determine what degrees of belief are rational, but are not themselves degrees of belief. By 'epistemic contexts', I mean contexts in which we reason about things like how probable a scientific theory is, or how strongly a theory predicts some evidence, or to what degree some evidence confirms a theory. It is standard to distinguish between the "epistemic probabilities" reasoned about in these kinds of contexts and the "physical probabilities" theorized about in, e.g., quantum mechanics (see, e.g., Romeijn 2022: sec. 2). My thesis here is only about epistemic probabilities. The relation between epistemic and physical probabilities is an important question, but one beyond the scope of this essay.

To be clear, I do not deny that 'probability' can be used in a technical sense to refer to degree of (rational) belief. Some philosophers use the term in this way in particular formal contexts, just as some mathematicians use 'probability' in a technical sense to refer to any quantity that satisfies Kolmogorov's axioms. But just as the latter technical usage does not show normalized areas (which satisfy Kolmogorov's axioms) to be a referent of 'probability' in ordinary language, the former technical usage does not show degrees of (rational) belief to be a referent of 'probability' in ordinary language. My focus here is on the entities picked out by pretheoretic uses of 'probability'-language (in epistemic contexts).

The main question that this paper addresses is a descriptive one: what are epistemic probabilities? However, I take the answer to this question to also have normative implications. If epistemic probabilities were degrees of support, but degrees of support were not that interesting, we might prefer to start using 'probability'-language to pick out a different quantity (perhaps adopting the technical usage of personalist philosophers). But as I characterize degrees of support, they are very important: they tell us what degrees of belief we ought to have. Consequently, we could not do epistemology without theorizing about them. As such, I take the argument of this paper to suggest, not only that we pretheoretically use 'probability'-language to refer to degrees of support, but that we ought to continue to do so.

This paper proceeds as follows. In section 2, I explain what degrees of support are, and argue that debates among orthodox Bayesians about the norms governing rational degrees of belief can be recast as debates about the nature of degrees of support and the way in which they constrain rational degrees of belief. As such, commitment to a particular position in these debates is not a reason to favor a degree-of-belief interpretation of epistemic probability over a degree-of-support interpretation. In section 3, I argue that natural probabilistic reasoning in certain cases involving old evidence is inconsistent with a degree-of-actual-belief interpretation of the probabilities reasoned about in those cases, but consistent with a degree-of-support interpretation of those probabilities. In section 4, I argue that natural probabilistic reasoning in two further old evidence

[1] The language of 'subjective probability' goes back to early Bayesians such as de Finetti (1931). The general presentation described above can be found in many contemporary philosophical presentations of Bayesianism. For example, Meacham (2014: 1185) follows the presentation above almost exactly, except that he uses the terms "permissive" and "impermissive" rather than "subjective" and "objective"; and Jon Williamson (2010: 15–16) delineates three different versions of Bayesianism (strictly subjective Bayesianism, empirically based subjective Bayesianism, and objective Bayesianism) based on how many constraints they put on rational credences.

cases is inconsistent with an interpretation of the relevant probabilities as rational initial degrees of belief, but consistent with a degree-of-support interpretation of those probabilities. In section 5, I argue that not only is the degree-of-support interpretation consistent with how we reason in these cases, it can also *explain* the propriety of that reasoning. Finally, in section 6, I give some reasons to extrapolate from the preliminary conclusion that the probabilities in these old evidence cases are degrees of support to the more general conclusion that all epistemic probabilities are degrees of support.

2 | DEGREES OF SUPPORT

The degree-of-support interpretation of probability[2] interprets probabilities as relations between propositions. It understands the probability of A given B as the degree to which B supports A, and the unconditional probability of A as the degree to which A is supported by a priori truths or tautologies. Entailment is a limiting case of this support relationship: if B entails A, then B supports A to a maximal degree, and $P(A|B) = 1$.[3] These support relations are not defined in terms of or reducible to degrees of belief. Instead, degrees of support rationally constrain degrees of belief. Below I discuss various possible bridge principles from the former to the latter; as a first pass, the idea is that, if $P(A|B) = r$, then someone with B as her evidence ought to be confident in A to degree r.

The precise nature of probabilistic support relations is left open by this minimal characterization. For the purposes of this paper, the distinctive claims of the degree-of-support interpretation are that probabilities are mind-independent relations between propositions and that probabilities constrain rational degrees of belief.[4] The key difference between the degree-of-support interpretation and degree-of-belief interpretations is that the latter define probabilities in terms of (rational, counterfactual, or actual) degrees of belief, while the former takes probabilities to be fixed independently from degrees of belief and to determine what degrees of belief are rational.[5]

I adopt this intentionally thin characterization of the degree-of-support interpretation here, not because I have no further opinions about the nature of probabilistic support relations, but because I want to be as ecumenical as possible in defending the claim that probabilities are degrees of support. For example, the claim that the conditional probability of A given B is the degree to which B supports A is sometimes combined with a rejection of Kolmogorov's ratio analysis of conditional probabilities, which defines $P(A|B)$ as the ratio $P(A\&B) / P(B)$, and the

[2] For ease of exposition, I mostly drop the qualifier 'epistemic' in the remainder of this paper.

[3] Degree-of-support theorists who endorse the ratio analysis of conditional probabilities (see below) may wish to make an exception for contradictions, saying that, e.g., $C\&{\sim}C$ does not support everything maximally. This is because $P(A|C\&{\sim}C) = P(A\&C\&{\sim}C) / P(C\&{\sim}C) = 0/0$, and so the ratio analysis leaves $P(A|C\&{\sim}C)$ undefined. There are also hard questions about whether understanding entailment as maximal support makes degrees of support unable to guide the reasoning of non-logically omniscient agents—see, e.g., Swinburne 2001: ch. 3—but these issues are beyond the scope of this paper.

[4] So characterized, a number of philosophers have endorsed a degree-of-support interpretation of probability, or something very similar. These include Keynes (1921), Jeffreys (1939), Carnap (1950), Timothy Williamson (2000), Swinburne (2001), Franklin (2001), Jaynes (2003), Hawthorne (2005), and Maher (2006).

[5] Rowbottom (2008: 342) offers a similar characterization of the difference between Keynes's (1921) "logical" interpretation and Jon Williamson's (2005) "objective Bayesian" interpretation: "the former defines rational degrees of belief in terms of probabilities … whereas the latter interprets probabilities as rational degrees of belief."

adoption of a non-standard axiomatization of probability that instead takes conditional probabilities as primitive (e.g., Hawthorne 2005: 288n9; Jaynes 2003: ch. 1–2).[6] However, one could endorse both the degree-of-support interpretation and the ratio analysis, as Timothy Williamson (2000: ch. 10) does.[7] Williamson holds that the probability of a hypothesis on one's evidence is the mind-independent degree to which "the evidence tells for or against the hypothesis" (209), rather than one's actual or hypothetical degrees of belief, but also holds that this conditional probability is defined as a ratio of unconditional probabilities. These unconditional probabilities come from a mathematical probability distribution that "measures something like the intrinsic plausibility of hypotheses prior to investigation" (211).[8]

In my view, both proponents and critics of the claim that probabilities are degrees of support have been too quick to combine this claim with logically independent assumptions about the nature of degrees of support. In particular, arguments for and against the degree-of-support interpretation have often assumed that degrees of support must be metaphysically necessary, knowable a priori, unique, and point-valued. That is, like the laws of logic and mathematics, degrees of support do not depend on contingent facts about the world, and are knowable apart from empirical investigation; and they can be precisely quantified, so that for any propositions A and B, B will support A to some precise degree r.

Consider necessity and apriority. Carnap (1950) infamously held that the degree to which B supports A is purely a matter of syntax: the form of the propositions A and B is enough to determine the degree to which B supports A. Contemporary defenders of the degree-of-support interpretation, by contrast, tend to hold that support relations are semantic: the degree to which B supports A is determined by the content of A and B, and not merely their form. Support relations are then analogous to natural language entailments, which are a function of the content of the propositions involved, and not merely their logical form.[9]

[6] It is somewhat unclear whether Jaynes understands probabilities to be degrees of support or rational degrees of belief. His main project is deriving objective probabilistic rules that an ideal reasoner should follow—suggesting a rational-degree-of-belief interpretation. And he cautions against projecting epistemology onto ontology (Jaynes 2003: 22). But by 'ontology' here Jaynes may only mean physical reality, not abstract reality (cf. Rowbottom 2008: 343). And Jaynes draws analogies with deductive logic that suggest he thinks his rules for ideal uncertain reasoning follow from logical or quasi-logical relations.

[7] There are also some philosophers who endorse the claim that conditional probabilities are primitive but do not endorse the degree-of-support interpretation. Hájek (2003) defends the former claim for probability understood as subjective degree of belief. And de Finetti claimed that all probabilities were conditional, but held that they were conditional partly on a person's state of mind: "every prevision and, in particular, every evaluation of probability, is conditional; not only on the mentality or psychology of the individual involved, at the time in question, but also, and especially, on the state of information in which he finds himself at that moment" (de Finetti 1970: 113). The relata of support relations, by contrast, are mind-independent entities. The degree-of-support interpretation of probability holds that conditional probabilities are relations between the propositions on the left-hand and right-hand side of the conditionalization bar, |, rather than an agent's (actual, hypothetical, counterfactual, or ideally rational) degrees of belief when (actually, hypothetically, counterfactually, or ideally rationally) in a state of mind.

[8] My thanks to an anonymous editor for pushing me to clarify the relation between the degree-of-support interpretation and the claim that conditional probabilities are primitive. I revisit the question of whether conditional probabilities are primitive in note 32.

[9] For example, according to Swinburne (2001: 64), the probability of q given r "has a value determined by the content of q and r, which measures the total force of r with respect to q; to which we try to conform our judgments of inductive probability on evidence but about the value of which we may make mistakes." Hawthorne (2005: 285) writes that Keynes and Carnap tried to "logicize" probability "through syntactic versions of the principle of indifference. But logical form alone cannot determine reasonable values for prior probabilities, as examples employing Goodmanian grue-predicates

Both the syntactic and semantic conceptions of support relations take them to be necessary and a priori. But there are alternative conceptions available. For example, one could adopt a frequentist analysis of degrees of support, on which the degree to which B supports A is something like the frequency with which A is true when B is true, or propositions like A are true when propositions like B are true. This could be the frequency in the actual world, or in some subset of possible worlds. This frequency would be a contingent, empirical fact. It would be knowable either a posteriori or not at all. The possibility of such a conception shows that we need not assume that support relations are necessary or knowable a priori.

There are philosophical objections that could be raised against this empirical conception of degrees of support. For example, one might worry that for one to be rationally required to conform one's degrees of belief to degrees of support, one must be able to tell how much one's current evidence supports any given proposition.[10] But this objection relies on a philosophically controversial premise, one that some epistemic externalists might deny. My goal here is not to suggest that they would be correct in doing so. It is simply to point out that this dispute is downstream from the one about the proper interpretation of probability. If I am right, there are good reasons to take probabilities to be degrees of support whether we go on to characterize degrees of support as a priori/internalist or empirical/externalist. We should all understand probabilities to be relations between propositions; we can go on to debate the precise nature of those relations later.

Similar remarks go for the assumption that degrees of support must be unique and point-valued. Philosophers skeptical that all probabilities are point-valued could hold that some degrees of support are imprecise or spread out, as Keynes (1921) did. Philosophers inclined towards "permissivism" about rational degree of belief could hold that there is not one unique degree to which B supports A, but rather many degrees, and that agents are free to choose which degree of support to adopt as a guide to their degree of belief.[11]

Indeed, the degree-of-support interpretation is compatible with a variety of different norms for degrees of belief.[12] Let's assume that degrees of support are point-valued, but may be non-unique. Here are some theories that imply progressively more demanding norms for rational degrees of belief. Note that in these claims 'probability function' means one of the degree-of-support functions (from an ordered pair of propositions to a number representing the degree to which the latter supports the former), and not just a mathematical probability function.

illustrate." Williamson (2000: 211) similarly asserts that, unlike Carnap, his probability measure "is *not* assumed to be syntactically definable. … The difference between green and grue is not a formal one." For his part, Franklin (2001) characterizes support relations as "logical", and uses the metaphor of "partial entailment", but (contra Carnap) holds that logical relations in general are not purely syntactical, noting that this view "has enough difficulties even in deductive logic" (287)—for example, in showing the incompatibility of "this is red" and "this is blue" (290).

[10] Titelbaum (2022: sec. 5.1.2) suggests that if there is a unique "hypothetical prior" function to which agents are rationally required to conform, we must be able to determine what it is a priori.

[11] The arguments given for the degree-of-support interpretation in sections 3–5 below appeal to the familiar kind of urn-sampling examples in which the probabilities have precise and unique values. But these arguments do not presuppose that all other probabilities are precise or unique.

[12] Titelbaum (2022: sec. 5.1.2) observes that the distinction between "objective" vs. "subjective" Bayesians crosscuts two different disputes: a normative disagreement over whether there is one uniquely rational credence distribution given a body of evidence, and a semantic disagreement over whether "'probability' talk expresses or reports the degrees of confidence of the individuals doing the talking" or has "truth-conditions independent of the attitudes of particular agents". Put in these terms, the degree-of-support interpretation implies semantic objectivism, but is neutral on normative objectivism.

(i) There are multiple probability functions $P_1(.)$, $P_2(.)$, etc. An agent's credence function at a time, Cr(.), is rational just in case there exists a probability function $P_i(.)$ such that Cr(.) = $P_i(.|X)$, where X is the agent's evidence at that time.
(ii) There are multiple probability functions $P_1(.)$, $P_2(.)$, etc. An agent's credence function at a time, Cr(.), is rational just in case there exists a probability function $P_i(.)$ such that (a) Cr(.) = $P_i(.|X)$, where X is the agent's evidence at that time and (b) at all earlier times, $Cr_{old}(.)$ = $P_i(.|Y)$, where Y was the agent's evidence at that earlier time.[13]
(iii) There are multiple probability functions $P_1(.)$, $P_2(.)$, etc., but they are indexed to agents, such that for any agent, there is some unique probability function $P_i(.)$ such that that agent's credence function at any time, Cr(.), is rational just in case Cr(.) = $P_i(.|X)$, where X is the agent's evidence at that time.
(iv) There is one unique probability function P(.). An agent's credence function at a time, Cr(.), is rational just in case Cr(.) = P(.|X), where X is the agent's evidence at that time.

Claim (i) implies a strong permissivism, on which your credences must conform to one of the probability functions, but which probability function you conform to is entirely up to you. Claim (ii) imposes a diachronic consistency requirement, on which the probability function you conform to now must be one you have conformed to in the past. This implies that a rational agent who does not lose evidence will satisfy the traditional Bayesian requirement of diachronic conditionalization. Claim (iii) implies intrapersonal uniqueness, on which, although different agents can rationally conform their credences to different probability functions, there is only one probability function that any particular agent can rationally conform her credences to.[14] Finally, claim (iv) implies interpersonal uniqueness, on which all agents are rationally required to conform their credences to the one unique probability function.[15]

These theories do not exhaust logical space. One could complicate matters further by varying how many degree-of-support functions there are,[16] allowing agents to have imprecise credences (see, e.g., Moss 2015a), or by allowing agents to have gaps in their credences (see, e.g., Eder forthcoming: sec. 4). But the above claims illustrate how one could endorse common Bayesian commitments about rational degree of belief while interpreting probabilities as degrees of support, and interpreting reasoning about probabilities as reasoning about degrees of support. Adopting a degree-of-support interpretation of probability does not commit us to any particular position in

[13] If one's credences in the past were not conformed to any of the probability functions, then (ii) implies that one cannot now be rational. This is a general problem for diachronic consistency principles: see Meacham 2016 for discussion.

[14] Kelly (2013), Meacham (2014), and Jackson (2021) distinguish between interpersonal and intrapersonal uniqueness. By relativizing to time, my formulations further distinguish intrapersonal uniqueness from diachronic consistency. Diachronic consistency requires you to conform to the same probability function at different times, but (unlike intrapersonal uniqueness) it leaves up to you which probability function that is.

[15] Claim (iv) is similar to Timothy Williamson's (2000: 220) ECOND and Hedden's (2015: 470) Synchronic Conditionalization.

[16] This is speaking a bit loosely. Plausibly, if there is more than one degree-of-support function, there are infinitely many, because for any distinct degree-of-support functions $P_1(.)$ and $P_2(.)$ and numbers m and n such that $P_1(A|B) = m$ and $P_2(A|B) = n$, there must be another degree-of-support function and number r such that $P_3(A|B) = r$ and r is a number between m and n. But the set of degree-of-support functions could still be more or less expansive in that for any conditional probability P(A|B) on which the different support functions disagree, the conditional probabilities assigned by the different support functions could range over a more or less expansive subset of [0,1]—e.g., [0.45, 0.5] vs. [0.35, 0.6].

recent debates about uniqueness,[17] or about the existence of diachronic norms governing rational credences.[18] Because we can adopt any of the normative positions in these debates within a framework that takes degrees of support to constrain rational degrees of belief, supporters of these positions have no reason to reject a degree-of-support interpretation of probability on normative grounds.

To reiterate: I am not claiming that conceptions of degrees of support as frequentist, subjectivist, imprecise, or derivative (from unconditional probabilities) are ultimately tenable. But we do not need to build in claims about primitiveness/derivativeness, necessity/contingency, aposteriority/apriority, precision/imprecision, or uniqueness/non-uniqueness into our initial concept of degrees of support. We can explore whether probabilities are degrees of support or degrees of belief without settling these further questions about the properties of degrees of support, leaving them as matters for future philosophical investigation and argumentation.

The upshot of this is that skepticism about the existence of a support relation with particular properties (such as apriority or precision), or about a particular bridge principle from degree of support to degree of belief, is not good grounds for skepticism about the existence of any probabilistic support relation at all. For example, if one agrees with Ramsey's (1926: 162–163) famous criticism of Keynes "that if we take the two propositions 'a is red', 'b is red', we cannot really discern more than four simple logical relations between them", and that none of these are degree of support, we should not immediately conclude that degree-of-support relations do not exist, for they may exist and simply not be *logical* in the sense Ramsey had in mind. Instead, if we find that positing degrees of support is necessary to explain the soundness of ordinary probabilistic reasoning, we should conclude that they do exist, but that, if Ramsey is right, they do not have all of the properties that (according to Ramsey, at least) Keynes ascribed to them.

3 | A NEW OLD EVIDENCE PROBLEM

This section proceeds as follows. First, I present a case—OLDEVIDENCE1—in which we naturally use Bayes' Theorem to calculate the posterior probability of a hypothesis. I then argue that the probabilities that enter into this application of Bayes' Theorem cannot be our actual degrees of belief, but can be degrees of support.

Here is the case:

> OLDEVIDENCE1
>
> Clark is taken to a room with an urn. Upon entering the room at t_1, Clark is told that the urn contains three balls, each of which is either black or white. Clark has not yet considered any hypotheses about the composition of the urn. At t_2, he decides to draw a ball from the urn. He draws a white ball and then returns it to the urn.
>
> At t_3, Clark is told more about the contents of the urn: this urn was selected by coin flip from two urns. The first urn contained 2 black balls and 1 white ball, while the second urn contained 1 black ball and 2 white balls.

[17] For defenders of uniqueness theses, see White (2005, 2013), Christensen (2007), Moss (2015a), and Hedden (2015). For opponents, see Kelly (2013), Meacham (2014), Schoenfield (2014), and Jackson (2021).

[18] Hedden (2015) and Moss (2015b) defend "time-slice rationality", the thesis that there are only synchronic norms. Recent defenders of diachronic norms governing credences include Briggs (2009), Podgorski (2016), and Carr (2016).

So, at t_1, Clark learns *K_1: This urn contains three balls, each of which is either black or white.*

At t_2, Clark learns *W: The ball drawn out of the urn is white.*

At t_3, Clark learns *K_2: The urn was selected by coin flip from U_1 (2 black, 1 white) and U_2 (1 black, 2 white).*

After t_3, having now learned that U_1 or U_2 are the only two possibilities for the urn's contents, Clark wonders: how probable are U_1 and U_2?

There is now a large literature on the so-called "old evidence problem" for Bayesianism. This literature goes back to Glymour (1980), who argued that on orthodox Bayesianism, facts learned some time ago are unable to confirm hypotheses formulated later, because after a fact is learned it has probability 1 conditional on anything. According to the standard Bayesian analysis of confirmation, E confirms H relative to K iff P(H|E&K) > P(H|K). In OLDEVIDENCE1, it seems that W confirms U_2, relative to K_2 (which, since it entails K_1, is logically equivalent to the conjunction K_1&K_2). The standard Bayesian analysis of confirmation will deliver this result just in case $P(U_2|W\&K_2) > P(U_2|K_2)$, which is true iff $P(W|U_2\&K_2) > P(W|U_1\&K_2)$. However, if these last two probabilities are interpreted as Clark's degrees of belief upon first entertaining U_1 and U_2 at t_3, then they should both be equal to 1, since at that time Clark had already learned W, and so his credence in it will be 1 conditional on anything.

This creates a *prima facie* problem for the combination of the standard Bayesian analysis of confirmation with a degree-of-belief interpretation of probability, for together these appear to render the incorrect verdict that W does not confirm U_2 relative to K_2, when it clearly does. There are a number of proposed solutions to this and related problems for Bayesian confirmation theory in the literature, which I will not canvas here (though some will come up in the discussion below). For now, the important thing to note is that this classic problem is only a problem for orthodox Bayesians who interpret probabilities as degrees of belief and is not a problem for heterodox Bayesians who interpret probabilities as degrees of support (see Rosenkrantz 1983: 85 and Hawthorne 2005). Degrees of support are atemporal: they are not affected by the order in which propositions conditioned on are learned. The value of P(A|B) stays constant regardless of when, or whether, A and B are learned. So there is no problem in assigning $P(W|U_2\&K_2)$ a higher value than $P(W|U_1\&K_2)$, on the degree-of-support interpretation.

The problem I wish to focus on is different. I am concerned, not with whether W confirms U_2, but with the posterior probability of U_1 and U_2, relative to W and Clark's other background knowledge—and, crucially, on how Clark can *figure out* those probabilities.

Let us continue the story above as follows:

OLDEVIDENCE1 (continued)

Applying Bayes' Theorem, Clark reasons as follows:

$$P(U_1|W\&K_1\&K_2) = P(U_1|W\&K_2) = \frac{P(U_1|K_2)\,P(W|U_1\&K_2)}{P(U_1|K_2)\,P(W|U_1\&K_2) + P(U_2|K_2)\,P(W|U_2\&K_2)}$$

$$= \frac{(1/2)(1/3)}{(1/2)(1/3) + (1/2)(2/3)} = \frac{1/6}{1/2} = \frac{1}{3}$$

> On the basis of this calculation, Clark concludes that the probability of U_1 on his total evidence is 1/3, and the probability of U_2 on his total evidence is 2/3.

Clark, it seems, has reasoned impeccably. I shall now argue that the propriety of Clark's reasoning is easier to reconcile with the degree-of-support interpretation of probability than a degree-of-belief interpretation. My argument proceeds as follows:

1. The probabilities mentioned in Clark's reasoning in OLDEVIDENCE1 have the values assigned to them in Clark's application of Bayes' Theorem.
2. If the probabilities in OLDEVIDENCE1 have these values, then these probabilities are degrees of support.
3. The probabilities in OLDEVIDENCE1 are degrees of support. [from (1)–(2)]

I take premise (1) to be pretheoretically obvious—not pretheoretic in the sense that it does not rely on any knowledge of the mathematics of probability, but in the sense that it does not rely on any philosophical assumptions about the nature of probability. Clark's application of Bayes' Theorem is of the kind that might appear in an introductory textbook on probability, in which the given values of the probabilities would be assumed without argument. (1) is an intuitive datum that can be used to judge the adequacy of different interpretations of probability, but does not antecedently assume any interpretation.

In the rest of this section I will present a preliminary argument for premise (2). Consider Clark's application of Bayes' Theorem above. If the values assigned to the probabilities on the right-hand side of this equation are correct (that is, if (1) is true), then these probabilities cannot be interpreted as Clark's degrees of belief upon considering U_1. This is because W is old evidence: Clark did not just learn it, but instead knew it before learning K_2, which prompted him to consider U_1. As such, if we interpret the probabilities above as Clark's conditional degrees of belief at t_2 or t_3, then they do not have the values given. Instead, the likelihoods all have probability 1, because Clark already knows W to be true. This further means that the priors cannot have probability 1/2, because if they did, this would wrongly imply that the posterior probability of U_1 is 1/2, which it is not—it is 1/3. So, if the probabilities on the right-hand side of Bayes' Theorem are Clark's conditional degrees of belief at t_2 or t_3, they do not have the values given above. Hence, if they do have the values given above, they are not Clark's conditional degrees of belief at t_2 or t_3.

Could these probabilities be Clark's conditional degrees of belief at t_1, or at some other earlier time?[19] No. For we stipulated that Clark has not considered U_1 or U_2 before learning K_2. As such, he has no degree of belief in them, either unconditional or conditional, nor does he have degrees of belief in other propositions conditional on them. But if these conditional degrees of belief are

[19] "Backtracking" to earlier credences is one common solution to the classic old evidence problem for Bayesian confirmation theory. I consider another common solution, appealing to counterfactual credences, below. It is worth pausing here to note the inapplicability of two other proposed solutions to the classic old evidence problem to my old evidence problem. One of these is to model confirmation in cases of old evidence as based on learning logical or explanatory facts (Garber 1983, Jeffrey 1983a; see Sprenger 2015 for criticisms). Even if, in OLDEVIDENCE1, Clark learns some such fact (e.g., that U_1 would, if true, explain W), adding this fact to Clark's evidence at t_3 would not change the conditional degrees of belief he has at that time, and so would not let us assign the intuitively correct values to the probabilities in Clark's application of Bayes' Theorem if we interpret those probabilities as Clark's conditional degrees of belief at that time. Another solution, championed by Christensen (1999), involves employing a non-standard Bayesian measure of confirmation. (For criticisms of Christensen, see Eells & Fitelson 2000 and Climenhaga 2013.) This solution will not help here because my problem is not dependent on how we measure confirmation.

undefined, then they do not have the values given above. Hence, if the probabilities on the right-hand side of Bayes' Theorem do have the values given above, they are not Clark's conditional degrees of belief at *any* time.

If, on the other hand, we interpret these probabilities as degrees of support, there is no problem in assigning them their natural values. Degrees of support are atemporal and not constrained by what a person knows at a given time. So a degree-of-support interpretation is consistent with Clark's reasoning.

If the probabilities on the right-hand side of Bayes' Theorem are degrees of support rather than degrees of belief, then so are the probabilities on the left-hand side. Bayes' Theorem is only a *theorem* if the kinds of probabilities in the equation are held constant, so that they are relative to a single probability distribution. (If we interpreted the probabilities on the right-hand side of an instance of Bayes' Theorem as your degrees of belief, and the probabilities on the left-hand side as my degrees of belief, then there is no guarantee that the equality will or should hold.) So in order for Clark's application of Bayes' Theorem to make sense, the probabilities on the left- and right-hand side must be given the same interpretation.

But before reaching the heterodox conclusion that the probabilities in OLDEVIDENCE1 are degrees of support, we should consider alternative orthodox interpretations of these probabilities. One relatively minor way to amend degree-of-belief Bayesianism, in keeping with some solutions to the classic old evidence problem,[20] is to interpret the probabilities in OLDEVIDENCE1, not as Clark's *actual* degrees of belief, but as the conditional degrees of belief he *would* have had at some earlier time had he considered U_1 or U_2, or the probabilities that he *should* have had at that time.[21]

Supposing that the probabilities in OLDEVIDENCE1 are the conditional degrees of belief that Clark would/should have had at some earlier time, what is that earlier time? t_1? Some earlier time before Clark entered the room? The probabilities Clark is reasoning with do not come with "temporal tags" that tell him what times they are relative to. Moreover, are *all* the probabilities we reason with the conditional degrees of belief we would/should have had at some earlier time, or only some? If not, how do we tell when we are reasoning about current degrees of belief, and when we are reasoning about historical degrees of belief?

The least arbitrary answer to these questions is that, in general, one's epistemic probabilities are the credences one should have had or would have had had one considered all the possibilities at the beginning of one's epistemic life—that is, at the moment at which one first had credences. Bayesians who have recognized the psychological implausibility of an agent who starts out with well-defined credences over all imaginable propositions have moved to talking of "hypothetical priors" here, which for our purposes we can understand as the credence distribution that one should have had or would have had upon considering all possibilities at the beginning of one's epistemic life.[22]

[20] See, e.g., Howson 1991 and Howson & Urbach 2006: 297–301. For criticisms, see Chihara 1987, Maher 1996, and Sprenger 2015.

[21] One might alternatively say that they are the degrees of belief that Clark would or should have *now* were his background evidence different (as opposed to the degrees of belief he would or should have had at a time when his background evidence *was* different). I discuss this view in note 23.

[22] Meacham (2016) helpfully distinguishes several different ways of understanding and conditionalizing on "ur-priors". In addition to interpreting them as initial credences (or the initial credences an agent ought to have had), he considers interpreting them purely functionally, as any probability function (in the mathematical sense) to which one's credences conform over time, and interpreting them as an agent's "evidential standards". The purely functional interpretation is not sufficient to avoid the problems developed below, as there will be mathematical probability functions to which a coherent

In this section I have argued for the following three claims. First, the probabilities on the right-hand side of the above application of Bayes' Theorem can be interpreted as degrees of support, and if they are, then we should interpret the probabilities on the left-hand side as degrees of support as well. Second, these probabilities cannot be interpreted as Clark's actual degrees of belief at any time. Third, if they are interpreted as counterfactual or rationally required degrees of belief at some earlier time, they need to be the degrees of belief that Clark would or should have had at the beginning of his epistemic life.

In what follows I will focus on the latter possibility (that probabilities are rational initial degrees of belief), but all of my criticisms apply to the former possibility (that probabilities are counterfactual initial degrees of belief) as well. In section 4, I will present two new cases, OLDEVIDENCE2 and OLDEVIDENCE3, involving *permanently* old evidence and *necessarily* old evidence, where we end up assigning the wrong values to the relevant probabilities if we interpret them as rational initial credences.

4 | PERMANENTLY OLD EVIDENCE AND NECESSARILY OLD EVIDENCE

I argued in section 3 that the probabilities in OLDEVIDENCE1 cannot be interpreted as Clark's actual degrees of belief at any time, whereas they can be interpreted as degrees of support. In this section, I will present two cases in which the probabilities also cannot be interpreted as rational initial degrees of belief. Consider first:

> OLDEVIDENCE2
>
> As soon as he was born, Ernest was placed in a large urn. A short while later, at t_1, Ernest was pulled out of the urn holding a white ball. The very first thing he learned, and the first thing he remembers, is that he was holding a white ball when drawn out of the urn. Call this proposition W′. Right after learning W′, at t_2, Ernest got his first credences.
>
> Later in his life, at t_3, Ernest is told a bit about his origins. He learns that before he was born, a coin was flipped. If the coin landed heads, he would be placed in an urn with 2 black balls and 1 white ball (U_1). If it landed tails, he would be placed in an urn with 1 black ball and 2 white balls (U_2). Call this proposition K′.
>
> So, at t_1, Ernest learns *W′: The ball drawn out of the urn is white.*

agent's credences conform over time that assign probability 1 to evidence that the agent has possessed her whole life, and so those functions will not allow us to employ that evidence in Bayes' Theorem in the way illustrated in section 4. Whether the evidential standards interpretation of ur-priors can avoid the problems I develop below depends on precisely what evidential standards are. The argument in section 4 will imply that if one's "evidential standards" do not assign probability 1 to evidence one has possessed one's whole life, or to evidence that all agents necessarily possess, then these evidential standards cannot be actual, hypothetical, or counterfactual degrees of belief, and must instead be something like an agent's beliefs about degrees of support. Consequently, if conditionalizing on "evidential standards" avoids the problems I go on to develop, it is only because these standards themselves presuppose the existence of a degree-of-support function.

At t_2, Ernest gets his first credences.

At t_3, Ernest learns *K′: The urn was selected by coin flip from* U_1 *(2 black, 1 white) and* U_2 *(1 black, 2 white).*

After t_3, Ernest wonders: how probable are U_1 and U_2?

Ernest reasons as follows:

$$P\left(U_1|W'\&K'\right) = \frac{P\left(U_1|K'\right)P\left(W'|U_1\&K'\right)}{P\left(U_1|K'\right)P\left(W'|U_1\&K'\right)+P\left(U_2|K'\right)P\left(W'|U_2\&K'\right)}$$
$$= \frac{(1/2)(1/3)}{(1/2)(1/3)+(1/2)(2/3)} = \frac{1/6}{1/2} = \frac{1}{3}$$

On the basis of this calculation, Ernest concludes that the probability of U_1 on his total evidence is 1/3, and the probability of U_2 on his total evidence is 2/3.

Like Clark, Ernest has reasoned impeccably. Unlike in OLDEVIDENCE1, however, we cannot interpret the probabilities on the right-hand side of the application of Bayes' Theorem in OLDEVIDENCE2 as the degrees of belief that Ernest should have had at the beginning of his epistemic life, because in this case the likelihoods all have probability 1, as Ernest already knew W to be true. This further means that the priors cannot have probability 1/2, because if they did, this would wrongly imply that the posterior probability of U_1 is 1/2, which it is not—it is 1/3. So, if the probabilities on the right-hand side of Bayes' Theorem are the conditional degrees of belief that Ernest should have had at the beginning of his life, then they do not have the values given above. Hence, since they do have the values given above, they are not the conditional degrees of belief that Ernest should have had at the beginning of his life.

In response to OLDEVIDENCE2, the orthodox Bayesian might idealize even further, and say that we should interpret the probabilities in our two problems, not as Ernest or Clark's rational initial credences, but as the credences of some ideal agent who can stand in for Ernest and Clark in some appropriate way.[23] But there are further cases in which even the credences of such a hypothetical agent will not give us what we are looking for. For there are some contingent propositions that, necessarily, any rational agent has as evidence. For example, consider the proposition that conscious things exist—call this proposition Conscious. If there is anyone around to have Conscious as evidence, this proposition is true, and at least know*able* by that person, so if that person is ideally rational, she *does* know it. (If you do not like this example, feel free to replace it with another proposition for which this seems more plausible: e.g., that there is something that has evidence,

[23] Eder (forthcoming) considers an argument of Williamson's (2000) against this interpretation. Williamson asks us to consider evidence E that makes probable T&(no one has great credence in T), where T is a complex logical truth. Then the probability of this conjunction on E is high, but an ideal agent would never have high credence in a Moorean-paradoxical proposition such as this. So the ideal agent could not have E as her evidence, and so the probability of this conjunction given E cannot be an ideal agent's credence. One response Eder suggests to this argument is that we instead adopt an interpretation of your epistemic probability of A given B as the credence *you* ought to have in A if you have B as your evidence. The arguments I go on to give against the ideal-agent's-credence interpretation will apply equally well against this interpretation.

or that there is some concrete thing.) But Conscious, like W′, can be relevant to the probability of other propositions. Now consider a third old evidence case:

> OLDEVIDENCE3
>
> Conrad receives the following revelation from God: at the beginning of the universe, God created two urns, one of which he would draw a ball out of. The color of the ball would determine whether or not God would create conscious life. God flipped a coin to choose between the urns. If the coin landed heads, he drew a ball out of the urn with 2 black balls and 1 white ball (U_1). If it landed tails, he drew a ball out of the urn with 1 black ball and 2 white balls (U_2). If God then drew a white ball, he set things up so that conscious things would evolve. If he instead drew a black ball, he set things up so that nothing conscious would evolve.[24]
>
> Call the content of this revelation K′′. In addition to K′′, Conrad already knows that conscious things exists (Conscious). Conrad then wonders: given Conscious&K′′, how probable are U_1 and U_2? He reasons using Bayes' Theorem as follows:

$$P\left(U_1|Conscious\&K''\right) = \frac{P\left(U_1|K''\right)P\left(Conscious|U_1\&K''\right)}{P\left(U_1|K''\right)P\left(Conscious|U_1\&K''\right) + P\left(U_2|K''\right)P\left(Conscious|U_2\&K''\right)}$$

$$= \frac{(1/2)(1/3)}{(1/2)(1/3) + (1/2)(2/3)} = \frac{1/6}{1/2} = \frac{1}{3}$$

> On the basis of this calculation, Conrad concludes that the probability of U_1 on his total evidence is 1/3, and the probability of U_2 on his total evidence is 2/3.

Like Clark and Ernest, Conrad has reasoned impeccably. Unlike in OLDEVIDENCE1 and OLDEVIDENCE2, however, we cannot interpret the probabilities on the right-hand side of the application of Bayes' Theorem in OLDEVIDENCE3 as the hypothetical degrees of belief of some ideally rational agent. For such an agent would always have Conscious as evidence, and so the agent's credence in Conscious conditional on anything would always be 1.[25]

It seems to me that the only response to OLDEVIDENCE3 available to the degree-of-belief theorist is to deny that there is any contingent proposition that our ideally rational agent must have as evidence. This is a strong commitment for the degree-of-belief Bayesian to take on. But perhaps we could motivate this response by arguing that we can "separate out" different aspects of rationality in the following way (cf. Eder forthcoming: sec. 2.3): distinguish ideal *reflection* from ideal *reasoning*, and hold that the former is relevant to what evidence an agent has, while the latter is relevant to what credences that agent has, given that evidence. We could then make our agent an

[24] If you are worried that God would himself be a conscious thing, imagine K′′ describing purely naturalistic laws on which a similar urn drawing process determines whether there is anything conscious.

[25] Some philosophers and physicists have endorsed "anthropic principles" on which propositions like Conscious cannot confirm anything, since they are inevitably old evidence, and so have probability 1 conditional on anything. (See, e.g., Monton 2006 and Pust 2007.) Since the claim that Conscious has probability 1 conditional on anything is inconsistent with the claims that P(Conscious|U_1&K′′) = 1/3 and P(Conscious|U_2&K′′) = 2/3, and these claims are obviously true, we should reject these anthropic principles.

ideal reasoner but not an ideal reflector, so that she can have ideal credences despite having no evidence.

Even if this move is tenable, there is still a serious problem in the neighborhood of OLDEVIDENCE3 facing an ideal-reasoner's-credence-interpretation of probability. There are some propositions, such as ~Conscious, that are *impossible* to have as evidence. But just as Conrad can reason about $P(U_1|\text{Conscious}\&K'')$, he can also reason about $P(U_1|{\sim}\text{Conscious}\&K'')$. This probability is calculable as a function of the same probabilities that are in OLDEVIDENCE3, and is equal to 2/3.[26] But if the probability of X given Y were the credence that an ideal reasoner with Y as her only evidence would have in X, then this probability would be undefined, since no agent could have ~Conscious as evidence.

In reply, the proponent of the ideal-reasoner's-credence interpretation could identify the probability of X given Y with the *conditional credence* in X given Y of an ideal reasoner with no evidence, where that conditional credence is either reducible to unconditional credences in X&Y and Y or is taken as a primitive. This may not deal with all cases, though. For there may be some propositions that not only cannot be possessed as evidence, but cannot be thought at all. For example, perhaps there are propositions about a particular one of Max Black's (1952) two identical spheres in an otherwise-empty universe, but these propositions are not thinkable, because neither of these spheres can be picked out in thought. But it seems that we can still say, for example, that $P(X|X) = 1$ for all propositions X,[27] including any propositions that are not thinkable.[28]

This version of the ideal-reasoner's-credence interpretation also faces a problem the original version did not: there are plausibly some propositions that it is impossible to have any credence in without first having some other proposition as evidence. For example, perhaps an agent cannot have a credence that something is conscious without knowing that she is conscious, because being acquainted with one's own consciousness is a necessary condition for possessing the concept of consciousness. Or perhaps an agent can only have thoughts about a concrete particular after first learning that that concrete particular exists. If there are any cases like this, we will not be able to identify all conditional probabilities with the conditional credences of an ideal reasoner with no evidence.

There thus remain serious problems in the neighborhood of OLDEVIDENCE3 for any interpretation of probability as the credences of an ideal agent. At the least, these interpretations will be saddled with heavy metaphysical commitments about topics besides probability, such as the nature of ideal reasoning and the conditions under which various propositions can be thought.

[26] Proof:

$$\begin{aligned} P(U_1|\sim Conscious\&K'') &= \frac{P(U_1|K'')P(\sim Conscious|U_1\&K'')}{P(U_1|K'')P(\sim Conscious|U_1\&K'') + P(U_2|K'')P(\sim Conscious|U_2\&K'')} \\ &= \frac{P(U_1|K'')[1 - P(Conscious|U_1\&K'')]}{P(U_1|K'')[1 - P(Conscious|U_1\&K'')] + P(U_2|K'')[1 - P(Conscious|U_2\&K'')]} \\ &= \frac{(1/2)(2/3)}{(1/2)(2/3) + (1/2)(1/3)} = \frac{1/3}{1/2} = \frac{2}{3} \end{aligned}$$

[27] Or all non-contradictory propositions: see note 3 above.

[28] In section 6 below, I consider the possibility that in cases calling for Jeffrey conditionalization, we learn "ineffable" propositions that we cannot fully understand. Even if it is possible to have an unthinkable proposition as evidence, though, it is still impossible to have a degree of belief in it, and so we cannot identify its probability conditional on itself with any kind of degree of belief.

In addition to these problems, I have a more general concern with any form of the ideal-agent's-credence interpretation. It appears to me that when it comes to actually *figuring out* what the credences of this ideal agent would be, all our reasoning will be about the relations between the propositions themselves. For example, if we ask what the ideal reasoner's credence in Conscious given U_1&K′′ would be, our reasoning will appeal solely to the fact that this urn has 2 black balls and 1 white ball and that K′′ says that conscious things will exist iff a white ball is drawn. It won't appeal to anything particular about the agent. (If we think to ourselves, "well, an ideally rational agent ought to be somewhat confident that consciousness exists, since she would herself be conscious", or anything of this sort, then we will get the wrong result.) Since the ideal agent and her characteristics are irrelevant to our actual reasoning about these probabilities, identifying these probabilities with this agent's credences rather than relations between the propositions themselves seems unmotivated.

In the next section, I will offer further support for this final worry, by explaining why the degree-of-support interpretation provides a better *explanation* of the reasoning employed in the OldEvidence cases than does any degree-of-belief interpretation.

5 | THE ORDER OF EXPLANATION AND THE ORDER OF LEARNING

In section 3 I gave a case in which a degree-of-actual-belief interpretation implies the intuitively wrong values for the probabilities an agent reasons about, and in section 4 I gave two more cases in which more idealized degree-of-belief interpretations appear to have the same implication. In this section I want to present an independent argument that the probabilities in our three cases are degrees of support, rather than degrees of belief. Focusing on OldEvidence1, I will argue that only the degree-of-support interpretation can *explain* why Clark reasons as he does.

Call the application of a theorem of probability *proper* when the theorem expresses a probability as a function of other probabilities the values of which we can more easily see or determine. We can then ask: when is the application of a theorem of probability proper? For our purposes, we can focus on Bayes' Theorem, which is often written more abstractly as follows:

$$P(H|E) = \frac{P(H)P(E|H)}{P(H)P(E|H) + P(\sim H)P(E|H)}$$

While the question of when Bayes' Theorem is the proper theorem to apply is rarely explicitly raised in discussions of probability, there are two common answers to this question implicitly given when Bayesian epistemologists introduce Bayes' Theorem in their writings.

The first answer is that Bayes' Theorem should be employed when H is a *hypothesis* or *theory* and E *empirical data* that H *predicts* to some degree (see, e.g., Howson & Urbach 2006: 20–22, Joyce 2021: sec. 1, Weisberg 2021: sec. 1.2.2). The basic idea of this advice is that we should apply Bayes' Theorem when H is *explanatorily prior* to E.[29]

[29] See Climenhaga 2020: section 3.4 and forthcoming: section 4. This idea is also present in older Bayesian terminology. As I note in Climenhaga 2020: 3226n20:

> Whereas today philosophers and statisticians follow R.A. Fisher in speaking of posterior probabilities and likelihoods, older writers (e.g., Venn 1866: sec. VI.9) referred to these as "inverse probabilities" and "direct probabilities," respectively. (These terms have occasionally survived, e.g., in Joyce [2021]: sec. 1.) The term

The second answer is that Bayes' Theorem should be employed when E is *new evidence* you are *updating* your credence distribution on. Degree-of-belief Bayesians implicitly suggest this answer when they introduce Bayes' Theorem in the context of the proposed requirement of diachronic conditionalization (sometimes, tellingly, called "Bayes' Rule") that $Cr_{new}(H) = Cr_{old}(H|E)$, where Cr_{new} is the agent's new credence distribution after learning E and Cr_{old} is the agent's old credence distribution before learning E (see, e.g., Hartmann & Sprenger 2010: 612, Strevens 2013: 307, Talbott 2016: sec. 2 and 4.1, Douven 2021: sec. 4). Occasionally this view is made explicit, as when Salmon (1990: 177) says that the "empirical data" E that enters into Bayes' Theorem is "new evidence we have just acquired". Bird and Ladyman (2013: 215) go so far as to write that "The first thing to note about Bayesian conditionalization is that a scientist's new credence in h, $P_{new}(h)$, is *determined by* her old credences $P_{old}(e)$, $P_{old}(h)$ and $P_{old}(e|h)$" (emphasis mine).

Our OLDEVIDENCE cases make clear that these two answers are not equivalent, and that we should prefer the first to the second. That we have just acquired evidence E is neither necessary nor sufficient for it to be proper to apply Bayes' Theorem in a way that "brings out" E on the right-hand side of the equation (as the proposition to the left of the conditionalization bar in the likelihoods).

To see that is it not necessary, consider again Clark's application of Bayes' Theorem:

$$P(U_1|W\&K_1\&K_2) = P(U_1|W\&K_2) = \frac{P(U_1|K_2)\,P(W|U_1\&K_2)}{P(U_1|K_2)\,P(W|U_1\&K_2) + P(U_2|K_2)\,P(W|U_2\&K_2)}$$

$$= \frac{(1/2)(1/3)}{(1/2)(1/3) + (1/2)(2/3)} = \frac{1/2}{1/2} = \frac{1}{3}$$

Clark's application of Bayes' Theorem was proper—it helped him to break down the value of $P(U_1|W\&K_1\&K_2)$ into more tractable probabilities—but the evidence W was old evidence, not evidence he had just learned.

To see that it is not sufficient, imagine that upon learning K_2, Clark had followed the second group of authors' advice to figure out the value of $Cr_{new}(U_1)$, which, according to diachronic conditionalization, should be equal to $Cr_{old}(U_1|K_2)$. He tries to expand this latter probability out through Bayes' Theorem in the way suggested by these authors:

$$P(U_1|K_2) = \frac{P(U_1)\,P(K_2|U_1)}{P(U_1)\,P(K_2|U_1) + P(\sim U_1)\,P(K_2|\sim U_1)}$$

$$= \frac{Cr_2(U_1)\,Cr_2(K_2|U_1)}{Cr_2(U_1)\,Cr_2(K_2|U_1) + Cr_2(\sim U_1)\,Cr_2(K_2|\sim U_1)}$$

If he does this, Clark will be no closer to figuring out what the old probability he should have assigned to U_1 conditional on K_2 is. Even given values for $P(U_1)$ and $P(\sim U_1)$, Clark will have no idea how likely it is that the urn was selected by coin flip from two urns with compositions U_1 and U_2, given U_1 or given $\sim U_1$ (and given that he drew white earlier, which he learned prior to K_2).[30]

"inverse probability" embodied the idea that in employing Bayes' Theorem we are moving "backwards" from effects to causes (Fienberg 2006: 5), and the term "direct probability" connotes a probability the value of which we are able to directly see or determine.

[30] This discussion illustrates another difference between my problem and the classic old evidence problem for Bayesian confirmation theory: the relevant sense in which W is old evidence in OLDEVIDENCE1 is not simply that it was learned

The second answer is thus untenable. Contrariwise, the first answer is supported not only by our own cases but by the kinds of examples other writers standardly use to illustrate Bayes' Theorem. Salmon (1990: 178) illustrates Bayes' Theorem with an example in which H is the hypothesis that this can opener was produced by a machine with a certain propensity for producing defective can openers, and E is the (explanatorily downstream) proposition that this can opener is defective. All four examples (drawing balls from an urn, finding a spider in a batch of bananas, hearing a witness report the color of a taxi, and getting a positive result on a medical test) in the "Bayes' Rule" chapter from Ian Hacking's introductory textbook (2001: ch. 7) likewise conform to this pattern. When we reflect on examples like these, it is clear that, where K is background information in the problem, the order in which we learned E and K does not make a difference to how we should employ Bayes' Theorem. This is just like our OldEvidence cases. All that matters is that E is explanatorily downstream from H and K is not; the order in which we learn E and K is irrelevant.

So far I have argued that the first answer to the question of when it is proper to employ Bayes' Theorem is correct, and the second answer is incorrect. I will now argue that the first answer fits more naturally with the degree-of-support interpretation of probability, while the second fits more naturally with a degree-of-belief interpretation.

The second answer fits more naturally with a degree-of-belief interpretation of probability because it ties Bayes' Theorem to an agent's state of mind prior to the employment of Bayes' Theorem—what credences she has and whether she has just learned evidence E. Unlike the degree-of-belief theorist, the degree-of-support theorist has no antecedent reason to expect that an agent's state of mind would be relevant to the proper application of Bayes' Theorem. (On a degree-of-support interpretation of probability, it is not necessary that one have evidence to reason about probabilities conditional on that evidence, as illustrated by the fact that we can reason about the probabilities in the OldEvidence cases even though we do not have Clark, Ernest, or Conrad's evidence.)

The first answer fits more naturally with the degree-of-support interpretation of probability because it ties Bayes' Theorem to relations between the propositions in the probabilities in the equation. Degree-of-belief interpretations of probability give us no reason to think that these relations should matter to the proper application of Bayes' Theorem. The degree-of-support interpretation, by contrast, does give us such reason. On the degree-of-support interpretation, probability is itself a relation between propositions. This support relation may be partly dependent on other relations, such as explanatory relations. For example, here is a plausible idea open to the degree-of-support theorist:[31] U_1 directly gives a probability to W in virtue of its being the sole proposition influencing its truth. It directly makes W probable to degree 1/3 because of the role it plays in explaining the truth or falsity of W.[32]

---

some time ago, but that it was learned before evidence explanatorily prior to it (that is, K_2). If, for example, instead of learning K_2, Clark had learned that a white ball had been drawn out of the urn a second time, there would be no need to "backtrack".

[31] In Climenhaga 2020, I develop this idea at greater length, and show in a more general way how explanatory priority relations determine when a theorem of probability is the proper one to apply.

[32] If the value of $P(W|U_1)$ is determined by the role U_1 plays in explaining W, this suggests that it is not determined by the ratio $P(W\&U_1) / P(U_1)$, as on the ratio analysis of conditional probability. If this is right, this idea would then support versions of the degree-of-support interpretation that take conditional probabilities to be primitive (e.g., Jaynes 2003, Hawthorne 2005) over those that define them as ratios of unconditional probabilities (e.g., Williamson 2000). By contrast, the arguments in sections 3–4 do not discriminate between these views, and are equally open to both versions of the degree-of-support view.

Philosophers have not clearly seen the objectivist implications of the way we employ Bayes' Theorem partly because we use the term 'prior probability' to refer both to one of the terms in Bayes' Theorem—an *explanatorily* prior probability—and to the probability of a proposition for some agent prior to receiving some evidence—a *temporally* prior probability. The argument of this section suggests that what matters for the proper application of Bayes' Theorem is the *order of explanation*, not the *order of learning*. Conflation of explanatorily prior probabilities with temporally prior probabilities has led to conflation of the order of explanation with the order of learning, and this has made the degree-of-belief understanding of probability appear more credible than it is. When we explicitly distinguish these, we see that in Bayesian reasoning we are thinking about relations between propositions, and not our credences before and after learning some evidence.

There is much more to be said about the proper application of Bayes' Theorem and other theorems of probability, but the foregoing suggests that a tenable solution to these problems will fit more naturally with the degree-of-support interpretation of probability than a (rational initial) degree-of-belief interpretation. We can accurately model reasoning in cases of old evidence using degrees of support, focusing on the explanatory order relating our propositions. We cannot accurately model reasoning in these cases using degrees of belief, focusing on the temporal order of our learning.

6 | EXTRAPOLATING TO OTHER CASES

In section 3, I presented the following argument:

1. The probabilities mentioned in Clark's reasoning in OldEvidence1 have the values assigned to them in Clark's application of Bayes' Theorem.
2. If the probabilities in OldEvidence1 have these values, then these probabilities are degrees of support.
3. The probabilities in OldEvidence1 are degrees of support. [from (1)–(2)]

In section 4 I considered an objection to premise (2), namely that the probabilities in Clark's reasoning could have the values he assigns to them if they are his counterfactual or rational initial credences. I presented two more cases in which the probabilities reasoned about cannot have the values assigned to them if they are credences of any kind—at least, not without substantial additional metaphysical commitments. Insofar as one is persuaded by these cases, they provide inductive support for (2). By considering the three cases in succession, we can see that reasoning with permanently and necessarily old evidence is structurally identical to reasoning with more familiar kinds of old evidence. This makes it plausible that Clark is reasoning about the same kinds of probabilities in OldEvidence1 as Ernest and Conrad are in OldEvidence2 and OldEvidence3—so that (2) is true. In section 5 I then gave an independent argument for (2), namely that the degree-of-support interpretation provides a better explanation of why the probabilities in OldEvidence1 have the values they do than any of the degree-of-belief interpretations.

How far can we extrapolate from this? As stated, (3) only says that the probabilities in this one particular thought experiment are degrees of support. But old evidence of the kind present in OldEvidence1 is quite common. Scientists frequently get evidence relevant to a theory and then later get further evidence explanatorily prior to the earlier evidence—e.g., evidence about the experimental set-up that produced the earlier evidence. Doctors wondering whether a patient has a disease frequently learn things explanatorily prior to evidence they already had that helps them see how that evidence impacts their hypothesis—e.g., they learn about a patient's exercise habits, which interact with the hypothesis of disease to make the observed symptoms probable to some degree. If a scientist or doctor is in such a case and talks about how probable H is on the total available evidence E, this will then need to be the degree to which E supports H, by the above arguments.[33]

Moreover, while OLDEVIDENCE2 and OLDEVIDENCE3 are exotic cases (devised to respond to increasingly idealized degree-of-belief interpretations), OLDEVIDENCE1 is a normal case of probabilistic reasoning. As I said in section 3, it's a case that could appear in an introductory textbook. The order in which Clark learns the evidence was important in arguing for (2), but is not otherwise remarkable, or something that suggests that in asking what the probability of U_1 is, Clark is talking about a different kind of probability from the probabilities we talk about in other epistemic contexts.

In addition, I argued in section 5 that our application of Bayes' Theorem ought to be guided by the explanatory relations among the propositions we are reasoning about, rather than the order in which we learned them, and that this is best explained by probabilities being relations between propositions. Hence, even when our evidence happens to have been learned in an order that makes degree-of-belief interpretations compatible with our applying Bayes' Theorem in the natural way, only the degree-of-support interpretation can explain why we ought to apply Bayes' Theorem in that way.

These considerations give us some reason to infer from the specific claim that the probabilities in the OLDEVIDENCE cases are degrees of support to the claim that the probabilities we reason about in epistemic contexts more generally—whenever we do things like ask how probable a scientific theory is, or how strongly a theory predicts some evidence, or to what degree some evidence confirms a theory—are also degrees of support. This hypothesis is also simpler than the alternative possibility that some epistemic probabilities are degrees of support and others are not, and so arguably preferable on grounds of parsimony.

Still, these pro tanto reasons in favor of a universal degree-of-support theory could be defeated if we could identify other epistemic probabilities that are not plausibly degrees of support. I cannot consider all potential counterexamples to a universal degree-of-support interpretation here, but for illustrative purposes I will consider two particular kinds of probabilities that skeptics might maintain must be degrees of (rational) belief: prior probabilities and probabilities of uncertain evidence.

Hawthorne (2005) argues that interpreting likelihoods as degrees of support solves the old evidence problem for confirmation and explains intersubjective agreement in scientific practice.

[33] Objection: these cases are unlike OLDEVIDENCE1 in that the scientist/doctor has already considered the hypothesis in question before getting the new evidence, opening up the possibility of interpreting the prior probability of the hypothesis in Bayes' Theorem as their earlier degree of belief in that hypothesis. Reply: this is true, but even if (to make the example concrete) the doctor has already considered the hypothesis that her patient is anemic, she will usually not have already considered the probability of his observed fatigue given that he is anemic *and* does not exercise until she learns that he does not exercise. As a result, P(fatigue|anemic&∼exercise) and P(fatigue|∼anemic&∼exercise)—the likelihoods when she uses Bayes' Theorem to calculate P(anemic|fatigue&∼exercise)–will still not be interpretable as her degree of belief that the patient is fatigued at any time, because after observing fatigue her degree of belief in it is 1, and prior to observing it, she had not considered how probable the hypothesis of anemia makes it relative to the proposition that the patient does not exercise. (My thanks to Lydia McGrew for pushing me to clarify this point.)

He notes that one could accept an argument like this and still hold that the prior probabilities that enter into Bayes' Theorem (what I called "explanatorily prior probabilities" in section 5) are degrees of belief, perhaps motivated by the greater extent of disagreement among scientists about prior probabilities. My argument, however, rules out this possibility: in the OLDEVIDENCE cases, both the likelihoods and prior probabilities are degrees of support. Consider, for example, $P(U_1|K'')$ in OLDEVIDENCE3, which is equal to 1/2, and which Conrad uses to calculate that $P(U_1|Conscious\&K'') = 1/3$. $P(U_1|K'')$ cannot be Conrad's actual degree of belief at any time, since Conrad's current degree of belief in U_1 is 1/3, and Conrad had not considered U_1 in the past. And it cannot be the degree of belief Conrad should have had in U_1 at some time, since Conrad has always known that he is conscious, and relative to Conscious&K′′, the degree to which Conrad should believe U_1 is 1/3, not 1/2.

This example does not on its own show that all other prior probabilities are degrees of support as well. But it does show that one cannot defend the existence of degree-of-belief probabilities by distinguishing likelihoods from prior probabilities, and maintaining that while the former are degrees of support, the latter are degrees of belief. If some probabilities are degrees of belief, this distinction does not help us identify them.

Hawthorne (2005) goes on to suggest that a different group of probabilities are degrees of belief: the probabilities of uncertain evidential statements. In section 2 I considered several possible bridge principles from the degree to which a proposition is supported by one's evidence to one's degree of belief in that proposition. All these principles imply that if E is part of one's evidence, then E is certain, in the sense that the rational credence to have in E is 1. This is because $P(E|E\&K) = 1$, for any E and background K. But one might hold, with Jeffrey (1983b), that we sometimes have *uncertain evidence*, in the sense that our experience sometimes directly changes our credences over a partition of evidential propositions without making any of these propositions certain. Our credences in other hypotheses should then be updated by "Jeffrey conditionalization" to conform with these new credences. As Hawthorne (2005: 310) puts it, "To handle uncertain evidence … the agent's belief strength for a hypothesis should be a weighted sum of the degrees to which each possible evidence sequence supports the hypothesis, weighted by the agent's belief strengths for each of those possible evidence sequences." Here the probabilities of the hypothesis on each evidence sequence are degrees of support, and the probabilities of the evidence sequences are degrees of belief.

The universal degree-of-support theorist will maintain, contra Hawthorne, that the probabilities of the uncertain "evidential" propositions here are in fact the degree to which these propositions are supported by some further proposition, and identify that further proposition with our actual evidence. It is possible to do this while still being fairly concessive to proponents of Jeffrey conditionalization.[34] Schwan and Stern (2017) note that while most discussions of Jeffrey conditionalizing assume that the agent becomes certain of nothing *that she can express*, they allow that she becomes certain of a "dummy proposition"—that is, a proposition about her experience

[34] Even in cases where we do become certain of something expressible, McGrew (2010) observes that the Jeffrey conditionalization formula can still be useful because (when it applies—see the discussion of when Jeffrey conditionalization is warranted below) it captures the *way* in which our foundational evidence F impacts a hypothesis H, namely through impacting the probabilities of the uncertain statements $\{E_1, \ldots, E_n\}$. (See also McGrew & McGrew 2008.) In the language of section 5 above, the formula will be a "proper" theorem to apply in that it breaks down the probability of H given F into probabilities that are easier to determine the value of, namely the probabilities of H given each of $\{E_1, \ldots, E_n\}$ and the probabilities of each of $\{E_1, \ldots, E_n\}$ given F.

that she cannot express. They argue that cases that call for Jeffrey conditionalization are not really instances of uncertain learning, but of ineffable learning:

> [N]early every case that appears in the literature seems to be describable in terms of learning something ineffable with certainty. In Jeffrey's classic candlelight cases, for example, the agent's credence that some cloth is a particular color changes because of how the cloth appears in candlelight. The agent plausibly learns that the cloth appears *that way* with certainty (even though she cannot describe what she sees). (Schwan & Stern 2017: 5n8)

The universal degree-of-support theorist can plausibly maintain that it is *because* the agent learns the ineffable fact that the cloth appears that way, and because that fact supports the cloth's being red to (say) degree 0.7, that the agent's credence that the cloth is red ought to be 0.7. More generally, we can still say that, in general, an agent's degree of belief in any proposition ought to be equal to the degree to which that proposition is supported by her evidence—just now allowing that that evidence may include ineffable propositions. If we are unable to have any credence in an ineffable proposition,[35] and unable to tell how much an ineffable proposition supports or is supported by other propositions,[36] then this just implies: first, that our bridge principle should allow for our agent to not have any credence in some propositions, so that she is not required to have any credence in her ineffable evidence; and second, that the facts that determine what degrees of belief our agent should have are in some sense inaccessible to her, so that a norm like Jeffrey conditionalization might still be useful in guiding conscious reasoning in cases where it delivers the same result as would conditionalizing on the dummy proposition (cf. Schwan & Stern 2017: 5, 13).

Indeed, Schwan and Stern argue that utilizing the dummy proposition lets us characterize the cases in which Jeffrey conditionalization is warranted. (Most proponents of Jeffrey conditionalization, including Jeffrey himself, acknowledge that it delivers counterintuitive results in some cases, and accordingly needs to be circumscribed in some way.) Roughly, their idea is that after learning a dummy proposition D that leads to an update over a partition, agents should update their credence in a hypothesis H using Jeffrey conditionalization just in case the already-updated partition screens off the impact that D has on H, where the relation of screening off is formalized using causal graphs.

There is much more to be said about Jeffrey conditionalization and uncertain/ineffable learning. But the foregoing shows that it is possible for the universal degree-of-support theorist to interpret the probabilities of uncertain "evidential" propositions as the degree to which these propositions are supported by a dummy proposition, a proposition that arguably needs to be posited anyway to characterize when Jeffrey conditionalization is rational. This discussion also illustrates that while the universal degree-of-support interpretation of epistemic probability may be committed to our evidence being "known with certainty" in some sense, it is a fairly weak sense. What this interpretation is really committed to is the claim that if X is part of our evidence, then we ought to either have a credence of 1 in X or no credence at all in X.

[35] This is most plausible if we understand an ineffable proposition to be one we cannot even think. We might instead understand an ineffable proposition to be one we can think, but not put into words, in which case it may be possible to have a credence in an ineffable proposition.

[36] As Schwan and Stern (2017: 6) note, this is compatible with our being able to make qualitative judgments about how dummy propositions causally relate to other propositions.

In this section I have argued that not only can the degree-of-support interpretation best account for probabilistic reasoning with old evidence of a particular kind, but old evidence of this kind is quite common, and the degree-of-support interpretation can best explain the propriety of our reasoning even in cases not involving old evidence. I also briefly considered two pluralist proposals on which some epistemic probabilities are degrees of support and others are degrees of belief. The main argument of this paper rules out the first proposal (that prior probabilities are degrees of belief), while the second proposal (that the probabilities of uncertain evidence are degrees of belief) arguably does a worse job of making sense of probabilistic reasoning in cases where we learn something but do not become certain of anything we can express. This suggests that it will be difficult to develop plausible pluralist theories that accommodate the lessons of the OldEvidence cases while still allowing for some probabilities to be degrees of (rational) belief. While further work on the possibility of pluralism about epistemic probabilities is welcome, we can tentatively conclude that all epistemic probabilities are degrees of support, and not degrees of (rational) belief.[37]

ACKNOWLEDGEMENTS

Open access publishing facilitated by Australian Catholic University, as part of the Wiley - Australian Catholic University agreement via the Council of Australian University Librarians.

ORCID

Nevin Climenhaga https://orcid.org/0000-0002-7376-8788

REFERENCES

Bird, Alexander & James Ladyman, ed. (2013). *Arguing about Science*. London: Routledge.

Black, Max (1952). "The Identity of Indiscernibles." *Mind* 61: 153–164. https://doi.org/10.1093/mind/LXI.242.153

Briggs, Rachael (2009). "Distorted Reflection." *Philosophical Review* 118: 59–85. https://doi.org/10.1215/00318108-2008-029

Carnap, Rudolf (1950). *Logical Foundations of Probability*. Chicago: University of Chicago Press.

Carr, Jennifer Rose (2016). "Don't Stop Believing." *Canadian Journal of Philosophy* 45: 744–766. https://doi.org/10.1080/00455091.2015.1123454

Chihara, Charles (1987). "Some Unresolved Problems in Bayesian Confirmation Theory." *British Journal for the Philosophy of Science* 38: 551–560. https://doi.org/10.1093/bjps/38.4.551

Christensen, David (1999). "Measuring confirmation." *Journal of Philosophy* 96: 437–461. https://doi.org/10.2307/2564707

Christensen, David (2007). "Epistemology of Disagreement: The Good News," *Philosophical Review* 116: 187–217. https://doi.org/10.1215/00318108-2006-035

Climenhaga, Nevin (2013). "A Problem for the Alternative Difference Measure of Confirmation." *Philosophical Studies* 164: 643–651. https://doi.org/10.1007/s11098-012-9872-0

Climenhaga, Nevin (2020). "The Structure of Epistemic Probabilities." *Philosophical Studies* 77: 3213–3242. https://doi.org/10.1007/s11098-019-01367-0

Climenhaga, Nevin (forthcoming). "If We Can't Tell What Theism Predicts, We Can't Tell Whether God Exists: Skeptical Theism and Bayesian Arguments from Evil." *Oxford Studies in Philosophy of Religion* 11, ed. Lara Buchak and Dean Zimmerman.

[37] I am grateful to Thomas Mulligan, Francesco Nappo, Jeremy Steeger, Liam Kofi Bright, Elizabeth Jackson, Branden Fitelson, David Thorstad, John Hawthorne, Simon Goldstein, Dmitri Gallow, Lydia McGrew, and an anonymous reviewer and editor for comments on earlier drafts, and to audiences at the Central States Philosophical Association, the Central APA, the Australasian Association of Philosophy, the University of Notre Dame, Northwestern University, Western Michigan University, Florida International University, the University of Kent, and the Dianoia Institute of Philosophy for feedback on oral presentations of this work.

de Finetti, Bruno (1931/1989). “Probabilism: A Critical Essay on the Theory of Probability and the Value of Science,” *Erkenntnis* 31: 169–223. Translation of “Probabilismo,” *Logos* 14: 163–219. https://doi.org/10.1007/bf01236563

de Finetti, Bruno (1970/2017). Theory of Probability: a Critical Introductory Treatment. *Translation of* Teoria Delle Probabilita (1970) by Antonio Machi and Adrian Smith. Chichester, UK: Wiley.

Douven, Igor (2021). “Abduction,” in *The Stanford Encyclopedia of Philosophy* (Summer 2021 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/sum2021/entries/abduction/

Eder, Anna-Maria Asunta (forthcoming). “Evidential Probabilities and Credences.” *British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axz043

Eells, Ellery & Fitelson, Branden (2000). “Measuring confirmation and evidence.” *Journal of Philosophy* 97: 663–672. https://doi.org/10.2307/2678462

Fienberg, Stephen E. (2006). “When Did Bayesian Inference Become “Bayesian”?” *Bayesian Analysis* 1: 1–40. https://doi.org/10.1214/06-BA101

Franklin, James (2001). “Resurrecting Logical Probability.” *Erkenntnis* 55: 277–305. https://doi.org/10.1023/a:1012918016159

Garber, Daniel (1983). “Old Evidence and Logical Omniscience in Bayesian Confirmation Theory.” In Minnesota Studies in the Philosophy of Science *10:* Testing Scientific Theories, ed. John Earman: 99–132. Minneapolis: University of Minnesota Press.

Glymour, Clark (1980). *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Hacking, Ian (2001). *An Introduction to Probability and Inductive Logic*. Cambridge University Press.

Hájek, Alan (2003). “What Conditional Probability could not be.” *Synthese* 137: 273–323. https://doi.org/10.1023/b:synt.0000004904.91112.16

Hartmann, Stephan & Jan Sprenger (2010). “Bayesian Epistemology,” in S. Bernecker and D. Pritchard (eds.), *Routledge Companion to Epistemology* (London: Routledge), 609–620.

Hawthorne, James (2005). “Degree-of-Belief and Degree-of-Support: Why Bayesians Need Both Notions.” *Mind* 114: 277–320. https://doi.org/10.1093/mind/fzi277

Hedden, Brian (2015). “Time-Slice Rationality.” *Mind* 124: 449–491. https://doi.org/10.1093/mind/fzu181

Howson, Colin (1991). “The ‘Old Evidence’ Problem.” *British Journal for the Philosophy of Science* 42: 547–555. https://doi.org/10.1093/bjps/42.4.547

Howson, Colin & Peter Urbach (2006). *Scientific Reasoning: The Bayesian Approach*, Third Edition. Open Court.

Jackson, Elizabeth (2021). “A Defense of Intrapersonal Belief Permissivism.” *Episteme* 18: 313–327. https://doi.org/10.1017/epi.2019.19

Jaynes, E.T. (2003). *Probability Theory: the Logic of Science*. Cambridge: Cambridge University Press, edited by G.L. Bretthorst.

Jeffrey, Richard (1983a). “Bayesianism with a Human Face.” In Minnesota Studies in the Philosophy of Science *10:* Testing Scientific Theories, ed. John Earman: 133–156. Minneapolis: University of Minnesota Press.

Jeffrey, Richard (1983b). *The Logic of Decision*, 2nd edition. University of Chicago Press.

Jeffreys, Harold (1939/1998). Theory of Probability*, reprinted in Oxford Classics in the Physical Sciences series*. Oxford: Oxford University Press.

Joyce, James (2021). “Bayes’ Theorem,” in *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/fall2021/entries/bayes-theorem/.

Kelly, Thomas (2013). “Evidence Can Be Permissive.” In Matthias Steup and John Turri (Eds.), *Contemporary Debates in Epistemology*: 298–311. Blackwell.

Keynes, John Maynard (1921). *A Treatise on Probability*. London: Macmillan and Co.

Maher, Patrick (1996). “Subjective and Objective Confirmation.” *Philosophy of Science* 63: 149–174. https://doi.org/10.1086/289906

Maher, Patrick (2006). “The Concept of Inductive Probability.” *Erkenntnis* 65: 185–206. https://doi.org/10.1007/s10670-005-5087-5

McGrew, Lydia (2010). “Probability Kinematics and Probability Dynamics.” *Journal of Philosophical Research* 35: 89–105.

McGrew, Timothy and Lydia McGrew (2008). “Foundationalism, Probability, and Mutual Support.” *Erkenntnis* 68: 55–77.

Meacham, Christopher (2014). “Impermissive Bayesianism.” *Erkenntnis* 79: 1–33. https://doi.org/10.1007/s10670-013-9547-z

Meacham, Christopher (2016). “Ur-Priors, Conditionalization, and Ur-Prior Conditionalization.” *Ergo* 3: 444–492. https://doi.org/10.3998/ergo.12405314.0003.017

Monton, Bradley (2006). “God, Fine-Tuning, and the Problem of Old Evidence.” *British Journal for the Philosophy of Science* 57: 405–424. https://doi.org/10.1093/bjps/axl008

Moss, Sarah (2015a). “Credal Dilemmas.” *Noûs* 49: 665–683. https://doi.org/10.1111/nous.12073

Moss, Sarah (2015b): “Time-Slice Epistemology and Action Under Indeterminacy.” In Gendler and Hawthorne (ed.), *Oxford Studies in Epistemology: Volume* 5. https://doi.org/10.1093/acprof:oso/9780198722762.003.0006

Podgorski, Abelard (2016). “A Reply to the Synchronist.” *Mind* 125: 859–871. https://doi.org/10.1093/mind/fzv153

Pust, Joel (2007). “Cartesian knowledge and confirmation.” *Journal of Philosophy* 104: 269–289. https://doi.org/10.5840/jphil2007104619

Ramsey, Frank (1926/1990), “Truth and Probability,” in his *Philosophical Papers*. Cambridge: Cambridge University Press.

Romeijn, Jan-Willem (2022). “Philosophy of Statistics.” *The Stanford Encyclopedia of Philosophy* (Spring 2022 Edition), Edward N. Zalta (ed.), https://plato.stanford.edu/archives/spr2022/entries/statistics/.

Rosenkrantz, Roger (1983). “Why Glymour is a Bayesian.” In Minnesota Studies in the Philosophy of Science *10:* Testing Scientific Theories, ed. John Earman: 69–98. Minneapolis: University of Minnesota Press.

Rowbottom, Darrell P. (2008). “On the Proximity of the Logical and ‘Objective Bayesian’ Interpretations of Probability.” *Erkenntnis* 69: 335–349. https://doi.org/10.1007/s10670-008-9117-y

Salmon, Wesley (1990). “Rationality and Objectivity in Science or Tom Kuhn Meets Tom Bayes.” In Minnesota Studies in the Philosophy of Science *14:* Scientific Theories, ed. C.W. Savage: 175–204. Minneapolis: University of Minnesota Press.

Schoenfield, Miriam (2014). “Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief.” *Noûs* 48: 193–218. https://doi.org/10.1111/nous.12006

Schwan, Ben & Reuben Stern (2017). “A Causal Understanding of When and When Not to Jeffrey Conditionalize.” *Philosophers’ Imprint* 17: 1–21. https://doi.org/2027/spo.3521354.0017.008

Sprenger, Jan (2015). “A Novel Solution to the Problem of Old Evidence.” *Philosophy of Science* 82: 383–401. https://doi.org/10.1086/681767

Strevens, Michael (2013). “Notes on Bayesian Confirmation Theory,” in Bird & Ladyman (ed.), ch. 19, pp. 293–328.

Swinburne, Richard (2001). *Epistemic Justification*. Oxford: Oxford University Press.

Talbott, William (2016). “Bayesian Epistemology,” in *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), Edward N. Zalta (ed.), http://plato.stanford.edu/archives/win2016/entries/epistemology-bayesian/

Titelbaum, Michael (2022). *Fundamentals of Bayesian Epistemology*. Oxford University Press.

Venn, John (1866). The Logic of Chance*. London and Cambridge*: Macmillan and Co.

Weisberg, Jonathan (2021). “Formal Epistemology,” in *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), edited by Edward N. Zalta, http://plato.stanford.edu/archives/spr2021/entries/formal-epistemology/.

White, Roger (2005). “Epistemic Permissiveness.” *Philosophical Perspectives* 19: 445–459. https://doi.org/10.1111/j.1520-8583.2005.00069.x

White, Roger (2013). “Evidence Cannot Be Permissive.” In Matthias Steup and John Turri (Eds.), *Contemporary Debates in Epistemology*: 312–324. Blackwell.

Williamson, Jon (2010). *In Defence of Objective Bayesianism*. Oxford: Oxford University Press.

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

How to cite this article: Nevin Climenhaga. Epistemic probabilities are degrees of support, not degrees of (rational) belief. *Philosophy and Phenomenological Research*. 2023;1-24. https://doi.org/10.1111/phpr.12947