PRECISE CREDENCES

Michael G. Titelbaum

I am more confident than not that I will go in to my office tomorrow. I'm not *certain* that I will go, and I haven't even hit the point of *believing* that I will: it is the summer, I have no courses to teach or students to meet, I may wake up tomorrow and decide it's not worth the effort. But I'm more confident that I will go than I am that I won't. If I had to place my confidence on a scale of 0 to 100, I'd put it somewhere above 50.

Credences are numerical degrees of confidence. While they could be expressed as percentages—between 0 to 100, inclusive—it has become customary to measure them on a scale from 0 to 1. Credences are also often called "degrees of belief," though that name may hold the connotation that they are a species of ordinary, qualitative belief.

It's better to think of credence not as a kind of qualitative belief, but instead as a member of the same family as qualitative belief. That family—the family of doxastic attitudes—also includes certainty, disbelief, suspension of belief, and probably comparative confidence as well. The members of this family have a variety of commonalities. For example, we tend to think of credences as taking the same sorts of *objects* as outright beliefs. Many authors take these objects to be propositions, and so classify both credences and beliefs as propositional attitudes. I will follow that trend here, but if you think beliefs are adopted towards something other than propositions (sentences, perhaps?), you will be inclined to the same view about credences.

The theory of credences was developed to address a number of philosophical problems. One was the proper interpretation of "probability" locutions. If I say, "The probability that I'll go to the office tomorrow is over 50%," what does this mean, and what are the truth-conditions for my utterance? A number of interpretations of probability have been offered and defended (some of which we will discuss in Section 1.6), and it's not clear that every use of the term "probability" should be interpreted the same way. But one prominent suggestion, the "subjective interpretation of probability," is that probability statements express the speaker's degree of confidence in a proposition. So my utterance expresses a confidence over 0.5 that I shall go to the office.

Yet even if "probability" statements rarely—or never—express an agent's degrees of confidence, such degrees of confidence may still exist, and have philosophical work to do. Degrees of belief play a prominent role in traditional decision theory, the classic formal approach to rational choice

(about which more in Section 2.2). Credences also figure in Bayesian confirmation theory (Section 2.1), an account of evidential support rivaling other statistical approaches such as frequentism and likelihoodism. And they can be applied to such further topics as coherentism, Inference to the Best Explanation, and social epistemology (Section 2.3).

So if we grant that credences exist, what exactly does it take to possess one? In line with contemporary behaviorist approaches in psychology, de Finetti (1937/1964) *defined* the degree of belief assigned to an event by an individual as the rate at which she'd bet that it would occur (more about the details in Section 2.2). But as was typical with operationalism, this definition ran into problems when, say, an agent displayed inconstant betting behaviors over time, and so was difficult to assign a particular credence to. Nowadays we may grant than an agent with a particular degree of belief will, *if rational*, display particular betting behavior (Christensen, 2004). But we also tend to think of this normative connection less as a *definition* of credence and more as one aspect of what it is to possess a degree of confidence. Just as our account of qualitative belief has progressed beyond behaviorism to a broader functionalism, we think of credence as a multi-faceted mental state with descriptive and normative connections to a wide variety of behaviors and other attitudes.

Besides their connections to desires, intentions, and decisions contemplated in action theory and decision theory, credences are connected to other varieties of doxastic attitudes (not to mention emotions, sensations, and memories). If comparative confidence is a distinct type of mental state, it clearly is connected to credence: I am more confident of P than Q just in case my credence in P is higher than my credence in Q. As for qualitative attitudes, certainty is often identified with credence 1 in a proposition (though see Section 1.7 below). There must also be links between credence and outright belief: if I believe P, my credence in P should be higher than my credence in $\sim P$.

Can we find a fully general connection between credence and outright belief? Some authors (e.g., Holton, 2014) maintain that to the extent there are any credences, to possess credence x in P is just to hold an outright belief that the probability of P is x. Yet it's difficult to find a single concept of probability that applies to every proposition to which an agent might assign a degree of belief. And it seems agents (such as children) can be more or less confident of propositions without possessing a concept of probability. Moreover, whatever concept of probability we select, it seems conceivable for an agent to adopt a degree of confidence in the proposition that P has probability x. (We'll see a further technical difficulty with the credence-as-outright-belief theory in Section 1.2.) Most theorists now hold that the numerical value of a credence is an attribute of the attitude

adopted *towards* a proposition, not part of the *content* of the proposition towards which that attitude is adopted.[1]

Going in the other direction, the "Lockean Thesis"[2] takes outright belief just to be credence above a particular threshold. The threshold credence is usually lower than 1 (belief need not be certainty) but well above 1/2, and may depend on contextual parameters. The main objection to the Lockean Thesis is that one can describe rationally acceptable credence distributions which, by way of the thesis, generate rationally unacceptable patterns of belief. In the Lottery Paradox (Kyburg, 1961) an agent assigns to each ticket in a lottery a low credence that it will win, while assigning a high credence (perhaps certainty) that some ticket will win. For any Lockean threshold less than 1, we can arrange the numbers so that the agent winds up believing of each ticket that it will lose, while believing that some ticket will win—a logically inconsistent overall set of beliefs. Similarly, in the Preface Paradox (Makinson, 1965), an author has high confidence in each claim made in her book while also being confident that at least one of those claims is false. Via the Lockean Thesis this becomes belief in each conjunct of a conjunction coupled with disbelief in that conjunction.

How, then, to relate credence and outright belief in general? The most radical possibility is to deny either the existence of beliefs or the existence of credences. More conservatively, one could offer a reduction of one category to the other, or at least a principle of descriptive supervenience. Alternatively, one could grant that while beliefs and credences appear in a variety of configurations in *actual* agents, normative principles specify how they'd align in a *rational* agent. The current consensus is that something beyond just the Lockean Thesis would be required to make either of these approaches work; recent attempts to articulate belief-credence principles can be found in Leitgeb (2017), Douven (2012), and Lin and Kelly (2012).

On the other hand, one could concede that beliefs and credences are both genuine kinds of mental states an agent can possess, there are some ways in which they interact (or interact if one is rational), but no systematic general principles are available. While this stance is available to strong realists about beliefs and credences, it is especially attractive to theorists who read belief and credence ascriptions as convenient, simplifying models of a highly complex cognitive system. The belief-model and the credence-model are each effective and efficient in different circumstances, and may be applied toward different ends. In that case, it would be unsurprising if no universal translation from one to the other were available.

---

1 Moss (2018) takes the numerical value to be part of a credence's content, but takes credal objects to be more complicated than simple propositions.

2 Locke (1689/1975, Bk. IV). See also Foley (1993) for discussion.

1   RATIONAL CONSTRAINTS ON CREDENCE

Once we understand what a credence is, the next question is what it takes for a set of credences to be rational.

1.1   *The Probability Axioms*

The most generally-accepted rational credence norms are Kolmogorov's (1933/1950) axioms. Suppose we have a language $\mathcal{L}$ of propositions, which starts with a finite set of atomic propositions and then closes them under the standard truth-functional connectives. Define a real-valued function c over $\mathcal{L}$ representing the credence values an agent assigns the propositions in $\mathcal{L}$.[3] The precise, real-number values that c assigns each proposition are the "precise credences" of this entry's title; I'll discuss alternative formal approaches in Section 5 below.

Given this setup, Kolmogorov's axioms become the following.

NON-NEGATIVITY. For any $X \in \mathcal{L}$, $c(X) \geq 0$.

NORMALITY. For any tautology $\mathsf{T} \in \mathcal{L}$, $c(\mathsf{T}) = 1$.

FINITE ADDITIVITY. For any mutually exclusive $X, Y \in \mathcal{L}$,
$$c(X \vee Y) = c(X) + c(Y).$$

Mathematicians often call these *the probability axioms*, and call any distribution satisfying them a *probability function*. Probabilism is the position that rational credences form a probability function; in other words, rational credences satisfy the Kolmogorov axioms.[4]

The probability axioms set $0 \leq c(X) \leq 1$ for every $X \in \mathcal{L}$. Probabilism also entails a number of intuitive constraints on rational credence. Here's one example.

○ For any $X \in \mathcal{L}$, $c(\sim X) = 1 - c(X)$.

Suppose you assign a high confidence that anthropogenic global warming has occurred. This constraint requires you to assign a low confidence that no anthropogenic warming has occurred. And should you become more confident that anthropogenic warming has occurred, this constraint

---

3 While I will consider languages containing propositions, other authors describe credences as distributed over sentences, or sets of possible worlds, or sets of events, etc.

4 Probabilism is often described as the doctrine that rational *agents* have credences satisfying the probability axioms, or (if that's considered too unrealistic) that *ideally rational agents* have probabilistic credences. Both of these formulations make agents (real or ideal) the targets of evaluation. Strictly speaking, I prefer to evaluate credences (or sets of credences) for rationality, rather than agents. But for ease of locution I will largely treat the two as interchangeable here.

will require your confidence in that proposition's negation to decrease accordingly.

Some other intuitive constraints following from the Kolmogorov axioms.

- ◦ For any contradiction $\mathsf{F} \in \mathcal{L}$, $c(\mathsf{F}) = 0$.

- ◦ For any $X, Y \in \mathcal{L}$ (mutually exclusive or otherwise),

$$c(X \vee Y) = c(X) + c(Y) - c(X \,\&\, Y).$$

- ◦ For any $X, Y \in \mathcal{L}$, if $X \vDash Y$ then $c(Y) \geq c(X)$.

- ◦ For any logically equivalent $X, Y \in \mathcal{L}$, $c(X) = c(Y)$.

- ◦ For any finite set of mutually exclusive $X_1, \ldots, X_n \in \mathcal{L}$,

$$c(X_1 \vee \ldots \vee X_n) = c(X_1) + \ldots + c(X_n).$$

The last bulleted constraint has an important consequence when an agent considers a *partition*—a set of propositions whose members are mutually exclusive and jointly exhaustive. Because the disjunction of a partition's elements is a tautology, probabilism demands that the credences assigned to elements of a partition sum to 1.

A further important consequence of probabilism is that credences are strongly extensional. If an agent is certain that two propositions X and Y have the same truth-value (that is, if $c(X \equiv Y) = 1$), then for the sake of calculating credences X and Y might as well be logically equivalent. For instance, any credence equation or inequality in which X appears would remain true were any of its Xs replaced with Ys. Any difference in meaning, modal profile, etc. is irrelevant to probability once truth-values are established to be identical.

We can illustrate probabilism with Kyburg's Lottery example from page 3. Given a lottery with, say, 100 tickets, introduce a language whose atomic propositions are W_1 through W_{100} (with W_i indicating that ticket i wins the lottery). If the lottery is fair, an agent might assign $c(W_i) = 1/100$ for each W_i. From our first intuitive consequence of the probability axioms, we then have $c(\sim W_i) = 99/100$; the agent is highly confident of each ticket that it will not win. However, assuming no more than one ticket can win, our final intuitive consequence listed above yields:

$$c(W_1 \vee \ldots \vee W_{100}) = c(W_1) + \ldots + c(W_{100}) = 1. \tag{1}$$

So our agent is certain some ticket will win, as intuitively she ought to be.[5]

---

[5] Notice that none of this solves the Lottery *Paradox*, which brings full beliefs into the lottery picture. My goal is just to illustrate how probabilism is compatible with and supportive of a natural account of rational credences in the lottery case. A similar illustration could be given for Makinson's Preface example.

While proofs in the probability calculus usually proceed from Kolmogorov's axioms, practical problem-solving is often made easier by working with state-descriptions. Define a literal to be an atomic proposition of $\mathcal{L}$ or its negation, then define a state-description in $\mathcal{L}$ to be a maximal consistent conjunction of its literals. Every noncontradictory $X \in \mathcal{L}$ then has a unique disjunctive normal form, a disjunction of state-descriptions logically equivalent to X.[6]

Carnap (1950) makes repeated use of the fact that a distribution c over $\mathcal{L}$ satisfies the probability axioms just in case it assigns: (1) non-negative values to $\mathcal{L}$'s state-descriptions summing to 1; (2) for every noncontradictory X, a value equal to the sum of the values assigned to the state-descriptions in X's disjunctive normal form; and (3) a value of 0 to every contradictory proposition.[7]

This result is handy in two ways. First, we can completely characterize any probability distribution over $\mathcal{L}$ by specifying the values it assigns to $\mathcal{L}$'s state-descriptions. Second, given partial information about a probability distribution, we can determine what this information says about the values assigned to state-descriptions, then from there work out the values of (or constraints on the values of) other propositions.

For example, suppose I tell you that Bob is certain of $P \supset Q$, and is twice as confident of P as $\sim P$. It immediately follows that Bob's confidence in $\sim Q$ is less than or equal to $1/3$. Why? Well, the disjunctive normal form equivalent of $\sim Q$ is $(P \& \sim Q) \vee (\sim P \& \sim Q)$. Since Bob is certain of $P \supset Q$, the first disjunct receives credence 0, so for Bob $c(\sim Q) = c(\sim P \& \sim Q)$. But since $c(P) + c(\sim P) = 1$, and $c(P) = 2 \cdot c(\sim P)$, we have $c(\sim P) = 1/3$. The disjunctive normal form equivalent of $\sim P$ is $(\sim P \& Q) \vee (\sim P \& \sim Q)$. By Non-Negativity Bob's credence in the first disjunct must be greater than or equal to 0, so the second disjunct receives a credence less than or equal to $1/3$.[8]

Finally, with the notion of a probability function in hand we can define the notion of an *expectation*. Suppose we have a numerical quantity for which many values are possible. To calculate an agent's expectation for that quantity, we multiply each value times the agent's credence that the quantity will take that value, then sum over all the values available. For example, if I'm 10% confident that I'll go into my office two days this

---

6 To make the disjunctive normal form unique, we require literals to appear in a state-description in some canonical order (perhaps alphabetical, if the propositions are designated by letters), and then we require state-descriptions to appear in disjunctive normal forms in a canonical order as well.

7 I have never been able to discover whether this result was original to Carnap or not. I would sincerely welcome any e-mails demonstrating its historical provenance!

8 For more on the mathematical theory underlying this approach, and for a Mathematica routine that will solve many probability problems once they are reduced to algebra using state-descriptions, see Fitelson (2008).

week, 60% confident that I'll go in just one day, and 30% confident that I won't go in at all, then my expectation for the numbers of days I'll go into my office this week is:

$$0.10 \cdot 2 \text{ days} + 0.60 \cdot 1 \text{ day} + 0.30 \cdot 0 \text{ days} = 0.8 \text{ days}. \qquad (2)$$

1.2 *The Ratio Formula*

So far we have discussed unconditional credence—an agent's degree of confidence that a particular proposition is true in light of her current understanding of what the world is like. We may also inquire after an agent's *conditional* credence in proposition X given Y; this is the agent's credence in X upon making the additional assumption that Y. Notice that Y may be a proposition in which the agent currently has low unconditional credence. In asking for her credence in X given Y, we ask her to set aside her current actual opinion about Y, temporarily add Y to the stock of propositions she takes to be true, then assess X in light of this enhanced suppositional set.[9]

An agent's conditional credence in X given Y is denoted $c(X \mid Y)$, and is usually taken to be governed by the Ratio Formula.

RATIO FORMULA. For any $X, Y \in \mathcal{L}$ with $c(Y) > 0$,

$$c(X \mid Y) = \frac{c(X \& Y)}{c(Y)}.$$

The Ratio Formula can be read as either a descriptive truth or as a normative requirement. On the former approach, an agent's conditional credence X given Y takes a particular value just in case her unconditional credences in $X \& Y$ and Y stand in that ratio. This reading is most natural if one wants to reduce one type of credence to the other: one could hold that to have a conditional credence just is to have unconditional credences standing in a particular ratio; or one could hold that conditional credences are basic and unconditional credences are a proper subset of those.[10] Alternatively, one could see conditional credence as just another type of doxastic attitude on equal footing with unconditional credences, then read the Ratio Formula

---

9 Notice that we are discussing indicative, not subjunctive, conditional credences. The supposition Y is to be *added* to the agent's current set of assumptions about the world, with the resulting suppositional set assumed to be consistent. Most discussions of conditional credence concern the indicative form. For a treatment of subjunctive conditional credences, see Joyce (1999).

10 From the Kolmogorov axioms and Ratio Formula, it follows that for any $X \in \mathcal{L}$, $c(X) = c(X \mid \top)$. So unconditional credences can be thought of as conditional credences conditional on a tautology. See Easwaran (this volume) for more.

as a rational requirement on how conditional and unconditional credences should align.[11]

Note that as I've defined the Ratio Formula, it remains silent when the agent assigns the condition (proposition Y) a credence of 0. We will return to credences conditional on credence-0 propositions in Section 1.7.

Combining the Ratio Formula and Kolmogorov's Axioms yields the handy Law of Total Probability.

LAW OF TOTAL PROBABILITY. For any $X, Y_1, \ldots, Y_n \in \mathcal{L}$ such that the $Y_1, \ldots, Y_n$ form a finite partition,

$$c(X) = c(X \mid Y_1) \cdot c(Y_1) + \ldots + c(X \mid Y_n) \cdot c(Y_n).$$

The Law of Total Probability calculates the unconditional credence of X as a weighted average of X's credences conditional on members of the Y-partition, weighted by the unconditional credences in the Ys.[12]

To illustrate once more with our lottery scenario, suppose B is the proposition that our agent will benefit from the outcome of the lottery. She holds tickets 1 through 3, so is sure to benefit if they win. Also, her sister holds the very last ticket (ticket 100), and the agent is 1/2 confident that her sister will share the winnings should that ticket come in. Applying the Law of Total Probability (and recalling that W_i is the proposition that ticket i will win), the agent's credence that she will benefit is

$$
\begin{aligned}
c(B) &= c(B \mid W_1) \cdot c(W_1) + c(B \mid W_2) \cdot c(W_2) + c(B \mid W_3) \cdot c(W_3) \\
&\quad + c(B \mid W_4) \cdot c(W_4) + \ldots + c(B \mid W_{100}) \cdot c(W_{100}) \\
&= 1 \cdot 1/100 + 1 \cdot 1/100 + 1 \cdot 1/100 \\
&\quad + 0 \cdot 1/100 + \ldots + 1/2 \cdot 1/100 \\
&= 0.035.
\end{aligned}
\tag{3}
$$

Conditional credence also plays a crucial role in the notion of credal relevance. When $0 < c(Y) < 1$, all of the following inequalities are equivalent:

$$c(X \mid Y) > c(X), \tag{4}$$
$$c(X) > c(X \mid {\sim}Y), \tag{5}$$
$$c(Y \mid X) > c(Y), \tag{6}$$
$$c(Y) > c(Y \mid {\sim}X), \tag{7}$$
$$c(X \& Y) > c(X) \cdot c(Y). \tag{8}$$

---

11 For a discussion of how conditional credences interact with an agent's credences in conditionals, see Briggs (this volume).

12 Put another way, the Law of Total Probability requires an agent's unconditional credence in X to equal her expectation of her credence in X conditional on the true element of the Y-partition.

When these inequalities hold, we say that Y is positively relevant to X on the agent's credence function. (Since positive relevance is a symmetric relation, we may also say that X is positively relevant to Y.) Another way to put this is that the agent takes X and Y to be positively correlated. Replacing the greater-thans with less-thans describes when Y is negatively relevant to X (or negatively correlated with X) on an agent's credences. On the other hand, when $c(X \& Y) = c(X) \cdot c(Y)$ (or any of the other inequalities above becomes equality), we say that X is irrelevant to Y for the agent, or probabilistically independent of Y.

These relevance relations are relative to an agent's credences; they reflect which propositions she assesses as relevant to each other given her current understanding of the world. But we can also temporarily enhance her current set of suppositions about the world, and see whether any relevance relations change. This takes us from a notion of unconditional relevance to conditional relevance. Y is relevant to X conditional on Z just in case

$$c(X \mid Y \& Z) > c(X \mid Z). \tag{9}$$

For each of the inequalities above, a corresponding characterization of conditional relevance can be given by adding Z as a condition to the expressions on each side.

The notion of conditional relevance underlies a crucial notion in the philosophy of science: screening off. We say that Z screens off X from Y when X and Y are unconditionally dependent but the following two equalities hold:

$$c(X \mid Y \& Z) = c(X \mid Z), \tag{10}$$
$$c(X \mid Y \& {\sim}Z) = c(X \mid {\sim}Z). \tag{11}$$

In other words, X and Y are independent conditional on each of Z and ${\sim}Z$. In a screening-off situation, supposing either Z or ${\sim}Z$ makes the correlation between X and Y disappear.[13]

To illustrate one application of this concept, Reichenbach (1956) argues that a common cause screens off its effects from each other. Suppose X is the proposition that my newspaper reports that the Yankees won last night, Y is the proposition that your newspaper reports that the Yankees won last night, and Z is the proposition that the Yankees actually won. On the one hand, while I remain ignorant of Z it would be rational for me to treat X as relevant to Y. X provides information about Z, and therefore also provides information about Y. But once the truth-value of Z is established, X and Y lose the ability to say anything about each other; X and Y become

---

13 This definition generalizes to the case in which Z is a random variable capable of taking a variety of values z_i. Screening off then occurs when X and Y are unconditionally correlated, but become independent conditional on each proposition of the form $Z = z_i$.

independent conditional on any supposition about Z. Thus Z will screen off X from Y on my credence function.

A proximal cause will also screen off its effect from a distal cause. (Imagine Y states the final score of last night's Yankees game, Z is the proposition that the Yankees won, and X is the proposition that my newspaper reports that they won.) In general, probabilistic correlations (conditional and unconditional) can provide useful evidence about the causal relations among a set of variables. Some philosophers have even *defined* causality in terms of probabilistic relations. For more on all of this, see Hitchcock (2012).

One final point about conditional credences. Earlier (p. 2) I mentioned the theory that a credence of x in P is just the outright belief that the probability of P is x. There I noted a number of problems for that theory; now we can add that the theory seems to lack a good way of understanding conditional credence. A conditional credence $c(P \mid Q)$ of x cannot be read as a qualitative belief in the proposition "If Q, then the probability of P is x," nor can it be read as the belief that "The probability of 'If Q, then P' is x." This was established by a series of triviality results initiated by Lewis (1976).[14] For instance, Lewis' work shows that if we assume $c(P \mid Q) = x$ just in case $p(Q \rightarrow P) = x$ for some suitable notion of probability p and some indicative conditional $\rightarrow$, then it follows that every proposition is probabilistically independent from every other! This is obviously absurd. A conditional credence just isn't a credence—or a belief—about a conditional.

1.3 *Updating by Conditionalization*

The rational constraints on credence listed to this point have been synchronic—when they relate multiple credences, all the credences related are held at the same time. The degree of belief literature has also proposed a number of diachronic constraints, governing relations among credences assigned at different times.

Suppose we have two times, t_i and t_j, with the latter occurring after the former. Let c_i and c_j be the agent's credence functions at these two times. The most traditional, well-established, and well-known diachronic credal constraint is Conditionalization.

CONDITIONALIZATION. If $E \in \mathcal{L}$ represents everything the agent learns between t_i and t_j, then for any $X \in \mathcal{L}$, $c_j(X) = c_i(X \mid E)$.

The intuitive idea of Conditionalization is simple. Suppose that at t_i you don't know whether E is true. I ask you to hypothetically suppose E (temporarily add it to your stock of assumptions about what the world is like), then ask for your conditional credence in X given this supposition.

---

14 For the recent state of the art in this area, see Hájek (2011) and Fitelson (2015).

You offer some number. Then, between t_i and t_j, you learn that E is actually true (and learn nothing else besides). If I now ask you at t_j for your unconditional credence in X, it seems you should offer the same number you reported as a conditional credence before. After all, the set of real-world conditions against which you're assessing X is the same at both times; it's just that at t_i you were supposing E as a fact about the world, while at t_j you know E to be true.

Conditionalization integrates nicely with our other credal constraints. For instance, if c_i satisfies the Kolmogorov axioms and $c_i(E) > 0$, then conditionalizing yields a c_j distribution that satisfies the axioms as well. So if an agent begins with a probability distribution and repeatedly updates by conditionalizing, she is guaranteed to respect probabilism on an ongoing basis. The probability axioms and Ratio Formula also make updating by conditionalization cumulative and commutative. If you conditionalize successively on E and then E', this yields the same result as conditionalizing just once on $E \,\&\, E'$, which means it also yields the same result as conditionalizing on E' followed by E.

For a conditionalizing agent, current credences interact in an interesting way with predictions about future credences. Suppose an agent is certain at t_i that her t_j credences will be formed by conditionalizing on a proposition she will learn from some particular finite partition. (Perhaps she will conduct an experiment between t_i and t_j, and the propositions in the partition represent all of its possible outcomes.) Assuming she meets a few other plausible side-conditions, such an agent will satisfy the Reflection Principle.

REFLECTION PRINCIPLE.   For any $X \in \mathcal{L}$, $c_i(X \mid c_j(X) = r) = r$.

This principle, introduced by van Fraassen (1984), sets the agent's t_i unconditional credence in X equal to her t_i expectation of her unconditional t_j credence in X.[15] Notice that although a c_j appears in the righthand expression, the principle governs *synchronic* credal interactions: it relates the agent's c_i credences in X to her c_i credences about her future credences in X. Given (again) a few side-conditions, Reflection may be derived from the Kolmogorov axioms, the Ratio Formula, and the agent's certainty that she will update by conditionalizing on some member of a particular partition. Van Fraassen, however, argues in the opposite direction: he provides independent motivation for Reflection, then views Conditionalization as a derivable consequence. For more on the arguments in each direction, and the specific side-conditions required, see Weisberg (2007) and Briggs (2009).

---

15 To see why, return to our formulation of the Law of Total Probability on page 8, and let each Y_i there assert that the agent's unconditional t_j credence in X will take some particular real value r.

When an agent repeatedly updates by Conditionalization, she often finds herself calculating the value of $c(X \mid E)$. This calculation can be streamlined by a famous theorem.

BAYES' THEOREM. For any $X, E \in \mathcal{L}$ with non-zero c-values,

$$c(X \mid E) = \frac{c(E \mid X) \cdot c(X)}{c(E)}.$$

Bayes' Theorem has proved so central to the application of Conditionalization that theorists who work with degrees of belief are often called "Bayesians" (or "subjective Bayesians," or "Bayesian epistemologists"). In a moment I'll describe why Bayes' Theorem is so useful. But first, it's worth noting that Bayes' Theorem is indeed a *theorem*, easily derivable from the Kolmogorov Axioms and Ratio Formula.[16] Bayesianism has generated a great deal of controversy, especially among statisticians. But the controversial claim in Bayesianism isn't that Bayes' Theorem is true. Everyone agrees that the theorem follows from the Kolmogorov Axioms, and that *if* an agent is going to generate new credences over time by conditionalizing, *then* the theorem provides a handy tool for calculating post-update credences from pre-update credences. The controversy is whether agents should really update their credences by conditionalizing, and whether scientific inference is best understood as a series of conditionalizations.

Setting this controversy aside, why is the particular analysis of $c(X \mid E)$ in Bayes' Theorem so useful? Consider a scientific context, in which a theorist has a finite partition of hypotheses $H_1, \ldots, H_n$ about what's going on with some phenomenon. The theorist plans to run an experiment that she hopes will discriminate among the hypotheses. At time t_i, before she has run the experiment, the theorist has a set of unconditional credences c_i, which we call her *priors*. The theorist runs the experiment between t_i and t_j, and let's suppose the observation she makes is represented by proposition E. Given this new evidence, Conditionalization helps her calculate her credences at t_j, which we call her *posteriors*.

Suppose we're interested in the theorist's confidence in some particular hypothesis H_m after the experimental results come in. Applying Conditionalization, Bayes' Theorem, and then the Law of Total Probability to the denominator of Bayes' Theorem, we derive:

$$c_j(H_m) = \frac{c_i(E \mid H_m) \cdot c_i(H_m)}{c_i(E \mid H_1) \cdot c_i(H_1) + \ldots + c_i(E \mid H_n) \cdot c_i(H_n)}. \tag{12}$$

---

16 The theorem is traditionally attributed to the Reverend Thomas Bayes. Though Bayes never published the theorem, Richard Price found it in his notes and published it after Bayes' death in 1761. Pierre-Simon Laplace rediscovered the theorem independently later on, and was responsible for much of its early popularization.

Consider the components of the right-hand fraction one at a time. First, we have a number of expressions of the form $c_i(H_x)$. These are the theorist's priors in the various hypotheses. Presumably going into the experiment she has some unconditional levels of confidence in the hypotheses she is considering; these supply the priors in question. Then we have expressions of the form $c_i(E \mid H_x)$. An agent's conditional credence in an experimental result E given some hypothesis H_x is called her *likelihood* for that evidence on that hypothesis. A well-defined scientific hypothesis should make a prediction for how the theorist's experiment will come out, or at least should assign probabilities to various possible outcomes. These inform the theorist's likelihoods for various experimental outcomes (such as E) on the various hypotheses she entertains. Thus Bayes' Theorem allows the theorist to form a posterior opinion about each hypothesis H_m that she entertains, based on the evidence she's received, her unconditional priors in the hypotheses, and her t_i likelihoods—elements that are plausibly all easily to hand.

1.4   *Jeffrey Conditionalization*

Statisticians and philosophers of science often worry that Conditionalization allows a scientist's final verdict on a hypothesis to be influenced by her initial credence in that hypothesis—her personal degree of belief in the hypothesis before any evidence came in. Epistemologists worry about Conditionalization's conception of evidence. It seems that for Conditionalization to work, it must be possible to identify some proposition E representing *everything* the agent learns between t_i and t_j. Moreover, the agent must become *certain* of E between t_i and t_j, because updating the agent's credence in E itself using Conditionalization yields $c_j(E) = 1$. Finally, once an agent becomes certain of some proposition, subsequent updates by Conditionalization will retain that certainty forever.[17]

Conditionalization therefore seems to embody a conception of learning on which what is learned is explicitly summarizable in propositional form, becomes certain, and is retained ever after. To epistemologists, this is reminiscent of foundationalist approaches to evidence abandoned decades ago. It also violates the Regularity Principle, which deems it irrational for an agent to assign absolute certainty to an empirical proposition. (After all, what evidence could ever make you *entirely certain* that some empirical claim was true?)

To address these problems, Richard C. Jeffrey offers an updating rule that generalizes Conditionalization to allow for learning experiences in

---

17 It's easy to show that if an agent conditionalizes on E between t_i and t_j, she will have $c_j(E) = 1$, and then if she conditionalizes on some other evidence between t_j and t_k, she will still have $c_k(E) = 1$ as well.

which no certainties are gained. He introduces his rule using the following example.

> The agent inspects a piece of cloth by candlelight, and gets the impression that it is green, although he concedes that it might be blue or even (but very improbably) violet. If G, B, and V are the propositions that the cloth is green, blue, and violet, respectively, then the outcome of the observation might be that, whereas originally his degrees of belief in G, B, and V were .30, .30, and .40, his degrees of belief in those same propositions after the observation are .70, .25, and .05. (Jeffrey, 1965, p. 154)

Discussing the example, Jeffrey writes:

> If there were a proposition E in [the agent's] preference ranking which described the precise quality of his visual experience in looking at the cloth, one would say that what the agent had learned from the observation was that E is true.... But there need be no such proposition E in his preference ranking; nor need any such proposition be expressible in the English language.... The description 'The cloth looked green or possibly blue or conceivably violet,' would be too vague to convey the precise quality of the experience.... It seems that the best we can do is to describe, not the quality of the visual experience itself, but rather its effects on the observer, by saying, "After the observation, the agent's degrees of belief in G, B, and V were .70, .25, and .05." (Jeffrey, 1965, pp. 154–5)

Jeffrey proposed an updating rule he called "probability kinematics"; nowadays everyone calls it "Jeffrey Conditionalization." The rule applies when an agent's experience impinges on her credences by altering her degree of belief distribution across a particular finite partition in $\mathcal{L}$; any other changes in her credences are caused by the changes to this partition. If the originating partition is $B_1, \ldots, B_n$, then Jeffrey's rule is as follows.

JEFFREY CONDITIONALIZATION. For any $A \in \mathcal{L}$,

$$c_j(A) = c_i(A \mid B_1) \cdot c_j(B_1) + \ldots + c_i(A \mid B_n) \cdot c_j(B_n).$$

Jeffrey did not mean to rule out the possibility that some learning occurs by certainty acquisition. He just wanted to allow for the possibility of other types of learning experiences as well. So in the case where one of the B_m goes to certainty (and therefore every other member of the partition goes to credence-0), Jeffrey Conditionalization reduces to traditional Conditionalization.

Let's see how Jeffrey Conditionalization applies to Jeffrey's cloth by candlelight example. Suppose the agent is interested in the proposition M, that the selected piece of cloth will match her couch. She's certain that anything violet will match, she's certain anything green will not, and she's 50% confident that a blue cloth will match. (The match depends on the specific shade of blue.) Let t_i be the time before she inspects the cloth by candlelight. Using the Law of Total Probability and the initial unconditional credences Jeffrey provides, we have

$$c_i(M) = c_i(M \mid G) \cdot c_i(G) + c_i(M \mid B) \cdot c_i(B) + c_i(M \mid V) \cdot c_i(V)$$
$$= 0 \cdot .30 + 0.5 \cdot .30 + 1 \cdot .40 = 0.55. \tag{13}$$

Jeffrey also provides the agent's unconditional credences in G, B, and V at t_j, after the inspection. With these values, Jeffrey Conditionalization yields

$$c_j(M) = c_i(M \mid G) \cdot c_j(G) + c_i(M \mid B) \cdot c_j(B) + c_i(M \mid V) \cdot c_j(V)$$
$$= 0 \cdot .70 + 0.5 \cdot .25 + 1 \cdot .05 = 0.175. \tag{14}$$

The glimpse by candlelight increases the agent's confidence that the cloth is green and decreases her confidence that the cloth is violet, so the Jeffrey-prescribed posterior that the cloth will match decreases.

Notice how this change in credence is effected. The agent's visual experience changes her credences by directly altering her distribution across the cloth-color partition. Any changes to other propositions in the agent's language (such as M) are downstream effects of this direct alteration. Yet the dependencies between these downstream propositions and the color propositions remain unaltered: changing the agent's opinions about the color of the cloth doesn't change how confident she is that particular colors will match the couch. This is why the same conditional credences appear in both the $c_i(M)$ and the $c_j(M)$ calculations.

Against the background of the Kolmogorov axioms and Ratio Formula, Jeffrey Conditionalization is equivalent to the following condition.

RIGIDITY. For any $A \in \mathcal{L}$ and any B_m, $c_j(A \mid B_m) = c_i(A \mid B_m)$.

In a Jeffrey Conditionalization, experience alters an agent's credences across the B-partition. The agent's credences in other propositions conditional on the B_ms don't change. So the agent sets her posteriors by adopting unconditional credences in the B_ms from experience, copying over her old conditional credences, then applying the Law of Total Probability to calculate her unconditional credences in non-B propositions.

1.5 *Further Rational Requirements*

We have now seen a variety of putative rational constraints on credence: the probability axioms, the Ratio Formula, the Reflection Principle, Regularity,

and the diachronic rules of Conditionalization and Jeffrey Conditionalization. Yet there are infinitely many credence distributions (and sequences of credence distributions over time) compatible with these constraints. Are all of those distributions rationally permissible? Some of them are quite strange, and unintuitive—for instance, some assign very high credence to skeptical scenarios; some will lead agents to reason counter-inductively.

One extreme position about the strength of rational constraints is sometimes called "Objective Bayesianism." This position endorses the Uniqueness Thesis (Feldman, 2007; White, 2005) that given any body of evidence, there is exactly one credence distribution rationally permitted to any agent with that body of total evidence. At the other extreme, what we might call "Extreme Subjective Bayesians" hold that any probabilistic credence distribution is rationally permissible. In between are "Moderate Subjective Bayesians," who hold that there are some rational constraints beyond the ones we've described, but not enough to generate a unique permissible distribution in every case.

What might these further rational constraints be? A constraint that might considerably narrow the field of what's rationally permissible is the

PRINCIPLE OF INDIFFERENCE. If an agent has no evidence favoring any possibility in a partition over any other, then she should assign equal credence to each element of the partition.[18]

The traditional objection to this principle is that it seems to give conflicting advice when we repartition the same space of possibilities. Following van Fraassen (1989), suppose I tell you that a cube has been produced from a factory, and its side length is between 0 and 1 meter. Given the paucity of further evidence, if I ask how confident you are that the side length is less than 0.5 meters, the Principle of Indifference seems to require a credence of 1/2. But if I now ask how confident you are that the volume (which must be between 0 and 1 cubic meter) is less than 0.5 cubic meters, the Principle of Indifference also seems to require a credence of 1/2. Since a side length of 0.5 meters corresponds to a volume of 0.125 cubic meters, the only way to assign both these credences consistently with the probability axioms is to be absolutely certain that the volume in cubic meters is not between 0.125 and 0.5![19]

Another family of putative rational constraints has a member we've already seen. The Reflection Principle directs us to set our current uncon-

18 The basic idea here dates back at least to Laplace (1814/1995), who saw it as an application of what Bernoulli (1713) called the "principle of insufficient reason."

19 A more technically-sophisticated cousin of the Principle of Indifference is Jaynes' (1957a, 1957b) Maximum Entropy Principle. This principle applies more naturally over infinite partitions, and adapts well to a variety of forms of evidence. Yet it still succumbs to partition variance problems, and also conflicts with updating by conditionalization in particular cases. See Seidenfeld (1986).

ditional credence in a proposition equal to what we're certain it will be in the future—or if we're not certain of our future credences, equal to our expectation of what they will be. This principle directs us to defer to the opinions of our future self as if she were some sort of expert. But of course there are other experts in the world, such as contemporaries who we think have better judgment or information than ourselves. Following the lead of the Reflection Principle, Elga (2007) suggests that if c_e is the credence distribution of an agent we consider an expert, then for any $X \in \mathcal{L}$ (or at least any X in the expert's area of expertise) we should assign

$$c(X \mid c_e(X) = r) = r. \tag{15}$$

Thinking more metaphorically, an "expert" distribution worthy of our deference need not even be an agent. It may be rational to align our credences with certain objective numerical values in the universe. This brings us to the topic of direct inference principles.

1.6 *Direct Inference Principles*

Page 1 briefly mentioned interpretations of probability—proposals for the meaning of "probability" locutions. For example, the classical interpretation, dating back at least to Laplace (1814/1995), defined probability as the number of favorable outcomes of a process divided by the total number of outcomes possible. Later, the frequency theory of probability (associated most closely with von Mises, 1928/1957), read probability as the frequency with which an outcome would occur were a particular process repeated many times.[20]

My task here is not to assess these notions of probability as proposals in the theory of meaning, or in the theory of probability. Instead, I want to ask what these notions have to do with rational credence. Many Bayesians have endorsed principles of direct inference: principles carrying the agent from information about some notion of probability to specific credences in specific events. For example, it might be that if I'm certain a particular type of experimental setup produces a particular type of outcome with frequency x, then when an experiment of that type is to be run, I should have credence x that it will yield an outcome of that type. This would be a principle of direct inference from frequency facts to credences in outcomes.

Frequency-to-credence principles face notorious difficulties, even when sketched out as roughly as I've just done. For one, a single event (I go

---

20 The previous section introduced one usage of "Objective/Subjective Bayesian" terminology. That usage should be carefully distinguished from another usage that often comes up in the literature about interpretations of probability. In that literature, "Subjective Bayesianism" describes the position that in everyday talk, "probability" always refers to or expresses subjective credences. "Objective Bayesianism," on the other hand, holds that probability talk refers to something beyond the subject, such as frequencies or chances.

in to my office tomorrow) can be classed as the outcome of a variety of experiment types (choosing whether to go in on a summer day, choosing whether to go in on a Tuesday, etc.), which may yield different frequencies and therefore different credal recommendations. (This is one version of the "reference class problem."[21]) Also, if we tried to use this principle as a general credence-setting strategy, we'd have trouble with experiments that look to be unrepeatable. Before the Large Hadron Collider was switched on, newspapers prominently reported physicists' degrees of belief that doing so would destroy the Earth. It's difficult to align such credences with the frequency with which switching on the collider would cause global destruction; in the event of such destruction, the switching-on only occurs once.

It may therefore be preferable to link rational credence with "objective chance." As a notion of probability, chance is objective, in the sense that its value is determined by the physical makeup of an experimental apparatus. Chance may also be applied to events that occur only once. A frequency-to-credence principle recommends credence 1/6 that a fair die roll will come up 3 on the grounds that repeating the roll will yield 3 one-sixth of the time. The objective chance theorist recommends 1/6 on the grounds that a fair die is physically constituted in a particular manner (equally weighted on each side, etc.). This would remain true even if the die had never been rolled before, and was guaranteed to be destroyed after the roll in question.

The most famous direct inference principle linking credence and chance is Lewis' (1980) Principal Principle. *Very* roughly, and skipping over a great many details,[22] the Principal Principle directs an agent to set

$$c(A \mid Ch(A) = x) = x, \tag{16}$$

unless she possesses inadmissible evidence relevant to A. Here $Ch(A) = x$ is the proposition that the objective chance of A is x. So—setting aside the matter of inadmissible evidence for a moment—if the agent is certain that, say, a particular die has a 1/6 chance of coming up 3, the Principal Principle will set her credence in 3 at 1/6. If, on the other hand, the agent knows the die is biased, but splits her credence evenly between the number 3's having a 1/10 chance and a 1/5 chance of coming up, the Law of Total Probability will combine with the Principal Principle to yield:

$$
\begin{aligned}
c(3) &= c(Ch(3) = 1/10) \cdot c(3 \mid Ch(3) = 1/10) \\
&\quad + c(Ch(3) = 1/5) \cdot c(3 \mid Ch(3) = 1/5) \\
&= 1/2 \cdot 1/10 + 1/2 \cdot 1/5 \\
&= 0.15.
\end{aligned}
\tag{17}
$$

---

21 See Hájek (2007) for many more versions.
22 See Meacham (2010) for some of those details.

In other words, her credence that the die will come up 3 is her expectation of the objective chance of getting a 3. We can therefore think of the Principal Principle as an expert deference principle in which the expert is objective chance.

The key innovation of Lewis' Principal Principle is its treatment of evidence the agent takes to be relevant to the outcome of a chance event. Lewis divides such evidence into two sorts: admissible evidence is evidence that the agent takes to be relevant to the outcome because it affects her opinion of the objective chance of the event. For example, information about the weighting of the die is admissible with respect to the outcome of the roll—it affects how the agent thinks the roll will come out *by way of* affecting what the agent thinks are the chances of a 3. Inadmissible evidence affects the agent's opinion in some other way. For instance, if a confederate tells her how the roll came out, this affects the agent's opinion of whether it came out 3, but not by making her think the chances of a 3 were any different going in. Lewis' insight was that chance facts about an outcome screen off admissible information relevant to that outcome. So if E is admissible, the Principal Principle also gives us:

$$c(A \mid Ch(A) = x \,\&\, E) = c(A \mid Ch(A) = x) = x. \tag{18}$$

Admissible evidence relates to chances much the way a distal cause relates to the proximal cause of an event.

1.7 *Countable Additivity*

Up to this point the examples we've considered have typically involved only finitely many possibilities. But what if an agent considers a partition of infinitely many possible outcomes, and distributes her credence equally among all of them? How can this be modeled in our Bayesian epistemology?

To have a concrete example, let's suppose that a positive integer has been selected by some process, and our agent wants to assign equal credence to each integer's having been selected. Presumably that should be possible. But what numerical value might that credence take? It's easy to show that the probability axioms prevent its being a positive real. For suppose the agent assigns

$$r = c(1) = c(2) = c(3) = \dots. \tag{19}$$

(Where $c(1)$ is her credence that 1 was selected.[23]) For any positive real r, there will exist a positive integer n such that $r > 1/n$. Now consider the

---

23 Notice we are now dealing with a language containing infinitely many atomic propositions. While this is a change from our earlier setup, it's not too difficult to manage, and is fairly common in formal models.

agent's credence that the selected integer is between 1 and n (inclusive). If you look back at the list of intuitive constraints following from the Kolmogorov axioms (Section 1.1), the last principle on the bulleted list will give us

$$c(1 \lor 2 \lor \ldots \lor n) = c(1) + c(2) + \ldots + c(n) = r \cdot n > 1, \qquad (20)$$

which violates the axioms.

What other options are available? One popular suggestion is that when an agent assigns equal confidence to infinitely many possibilities, we represent that level of confidence as a credence of 0. So we would say that $c(1) = c(2) = \ldots = 0$.

Using credence 0 in this way introduces a few problems. First, up until this point we've conceived credence 1 as representing certainty in a proposition, and credence 0 as certainty that the proposition is false. Now we'll have to allow an agent to assign $c(P) = 0$ even if the agent admits P might be true, and $c(\sim P) = 1$ even if the agent isn't certain P is false. And we'll have to phrase the Regularity principle carefully: we may still prohibit agents from assigning certainty to empirical propositions, but no longer ban credences of 1 and 0 in such propositions.

Second, the Ratio Formula we've provided only relates the conditional credence $c(X \mid Y)$ to unconditional credences when $c(Y) > 0$. We'll need to expand this principle to handle cases in which $c(Y) = 0$ yet the agent doesn't rule Y out. For instance, our agent assigning equal credence to the selection of each positive integer might assign $c(2 \mid 2 \lor 4) = 1/2$, even though $c(2 \lor 4) = c(2) + c(4) = 0$.[24]

Third and most importantly, we'll want a way to sum credences over infinite disjunctions. Finite Additivity only covers disjunctions with finitely many disjuncts—what if we want to calculate our agent's credence that the selected integer is even? A natural extension of Finite Additivity is the following.

COUNTABLE ADDITIVITY. For any countable partition $\{Q_1, Q_2, Q_3, \ldots\} \subset \mathcal{L}$,
$$c(Q_1 \lor Q_2 \lor Q_3 \lor \ldots) = c(Q_1) + c(Q_2) + c(Q_3) + \ldots.$$

Countable Additivity is not only natural; it also allows us to establish a very important constraint on credences.

CONGLOMERABILITY. For any proposition $P \in \mathcal{L}$ and partition $\{Q_1, Q_2, Q_3, \ldots\} \subset \mathcal{L}$, $c(P)$ is no greater than the largest $c(P \mid Q_i)$ and no less than the least $c(P \mid Q_i)$.

---

24 One way to manage this situation is to take conditional credences as basic. See footnote 10 for more information.

Given Conglomerability, the $c(P \mid Q_i)$ establish upper and lower bounds on the value of $c(P)$. This makes sense if you think of $c(P)$ as a weighted average of the credences the agent would assign to P conditional on all the different possible Q_i. And it's especially important when the agent has a partition $\{E_1, E_2, E_3, \ldots\}$ of possible new pieces of evidence she might receive before her next update. Assuming she plans to update by Conditionalization, she knows that her future credence in P will be one of her current $c(P \mid E_i)$; Reflection then demands she satisfy Conglomerability.[25]

The Conglomerability/Countable Additivity package is attractive. But it's inconsistent with assigning a credence of 0 to each positive integer in our example. The reason is simple: given Countable Additivity, the agent's credence that any positive integer will be selected at all is the sum of her credences in each individual integer. But the former value should be 1, while the latter individual values are each 0. So advocates of Countable Additivity have suggested instead that in this situation the agent assign an infinitesimal value to each integer's being selected. The infinitesimals are an extension of the set of real numbers, defined to be greater than 0 but less than any given real number. Thus they don't fall prey to the problem of our Equation 20. At the same time, adding up infinitely many infinitesimals can yield a real number, so we can maintain both Countable Additivity and a credence of 1 that any integer will be selected at all.

Yet infinitesimals introduce difficulties of their own; for some of the difficulties, and many of the mathematical details, see Hájek (2003, Section 5), Williamson (2007), Easwaran (2014), and Wenmackers (this volume).

2   APPLICATIONS OF CREDENCE

I've presented the Bayesian study of credence as the study of a doxastic attitude type, and what it takes to make such attitudes rational. This study is valuable in its own right, as a contribution to epistemology and the philosophy of mind. But historically it's also been pursued to enhance our understanding of other topics, some of which we'll discuss in this section.

2.1   *Confirmation Theory*

A Bayesian epistemologist or philosopher of science studies justification and evidential support by thinking about "confirmation." The type of con-

---

25 Notice that my statement of Conglomerability doesn't specify the cardinality of the Q_i partition. For finite partitions, Conglomerability can be proven from the standard probability axioms. Adopting Countable Additivity extends Conglomerability to countable partitions. For an agent who entertains larger disjunctions than that, Seidenfeld, Schervish, and Kadane (manuscript) show that at each cardinality we need the relevant Additivity principle to secure Conglomerability for partitions of that size.

firmation studied is usually incremental, rather than all-things-considered; when we say that "evidence E confirms hypothesis H," we mean that E provides at least some positive evidential support for H, not that it settles the matter of H or even pushes H past some crucial threshold.[26] For a Bayesian, confirmation is also always relative to a probability distribution, and to a background corpus of propositions. Most commonly, the probability distribution will be some agent's credence function, and the background corpus will be the total evidence informing that credence function. (On a Conditionalization regime, the corpus is represented formally by the set of all propositions X such that $c(X) = 1$.[27]) So we take a given agent at a given time, and ask whether E confirms H for her, relative to her credences and background corpus at that time.

Letting K represent a background corpus, and c_k represent a probability distribution informed by that corpus, Bayesian confirmation theory posits that

> E confirms H relative to c_k just in case $c_k(H \mid E) > c_k(H)$.

Bayesian confirmation is just positive probabilistic relevance relative to c_k. (Similarly, disconfirmation is usually defined as negative relevance relative to c_k.)

Though fairly simple, this theory of confirmation turns out to be surprisingly subtle, powerful, and convincing. To illustrate—and fix the intended notion of evidential support in the reader's mind—suppose a fair die has just been tossed, and you know nothing of the outcome. Perhaps in accordance with the Principal Principle, some frequency principle, or even the Principle of Indifference, you assign equal credence to each of the six possible outcomes. Relative to your credence distribution and background corpus, if you received evidence that the toss came up with a prime number, this would confirm for you that the toss came up odd. Why? Because if you satisfy the Kolmogorov axioms and Ratio Formula, then you assign

$$2/3 = c(\text{odd} \mid \text{prime}) > c(\text{odd}) = 1/2. \qquad (21)$$

This doesn't mean that prime evidence should make you *certain* the toss came up odd, or even that it would justify you in *believing* the toss came up odd. But if you update by Conditionalization, learning that the toss came up prime would make you at least somewhat more confident that the toss came up odd. Again, the confirmation here is incremental.

---

26 This contrasts with the way "confirms" is sometimes used in English, as when we speak of a nominee's being confirmed, or even a dinner reservation.

27 Notice that despite our suggestion in Section 1.7 that it might sometimes be interpreted otherwise, I have gone back to treating credence 1 as representing certainty. To simplify discussion, I will continue to do this going forward.

This Bayesian theory of confirmation gives the confirmation relation some interesting and intuitive formal properties.[28]

○ If $E \dashv\vDash E'$ and $H \dashv\vDash H'$, then E confirms H just in case E' confirms H'.

○ E confirms H just in case E disconfirms $\sim H$.

○ If $E \& K \vDash H$ but $K \nvDash H$, then E confirms H.

○ If $H \& K \vDash E$ but $K \nvDash H$, then E confirms H.

The first of these properties ensures that logical equivalents behave the same within the confirmation relation. The second relates confirmation to disconfirmation. The third and fourth properties[29] specify how confirmation relates to entailment. The third property tells us that entailment is a form of confirmation; if E entails H jointly with K while K didn't entail H on its own, then E confirms H. As for the fourth property, it captures the idea[30] that a hypothesis which predicts an evidential observation (in concert with one's background corpus) is confirmed by that observation.

On the other hand, the Bayesian theory withholds from the confirmation relation certain properties that are sometimes mistakenly ascribed to it. Here are two examples.

○ If E confirms both H and H', then the set H, H', K is logically consistent.

○ If X confirms Y and Y confirms Z, then X confirms Z.

The first of these properties is important to reject because we're talking about incremental confirmation. For example, in Jeffrey's example in which an agent inspects a piece of cloth by candlelight, his brief glimpse may confirm that the cloth is green, while also confirming that it's blue or even that it's violet. (Perhaps the glimpse disconfirms that the cloth is red and disconfirms that it's orange.) This is perfectly reasonable, despite the fact that green, blue, and velvet are inconsistent hypotheses about the color of the cloth. Similarly, in scientific settings the same observation may confirm mutually exclusive theories from a partition, while at the same time (perhaps) ruling others out.

The latter property is the supposed property of confirmation transitivity. This is one of the most common mistakes made about confirmation, support, justification, and other related notions.[31] Just because X confirms Y

---

28 In every one of these properties, the expressions "E confirms H" and "E disconfirms H" should be followed by the phrase "relative to c_k." Going forward I'll simplify locutions by leaving the relativization to c_k implicit whenever possible.

29 Both of which require a side-condition that the set $\{E, K, H\}$ is logically consistent.

30 Familiar from hypothetico-deductivism (Crupi, 2016, Section 2).

31 Correcting this mistake has been a theme of the epistemology literature about epistemic and justificatory closure. See, e.g., Dretske (1970), Davies (1998) and Wright (2003).

and Y confirms Z does not mean that X confirms Z—even in the special case when Y *entails* Z! To see why, imagine a card has been drawn at random from a standard playing card deck. Information that the card is a spade confirms (incrementally!) that the card is the Jack of Spades. But information that the card is a spade does not even incrementally confirm that the card is a jack.

Another common mistake is to conflate what Carnap (1962) called "firmness" and "increase in firmness" accounts of confirmation.[32] The Bayesian account we've been discussing is an increase in firmness account. A firmness account, on the other hand, says that E confirms H relative to c_k just in case $c_k(H \mid E)$ is high (where the necessary height may be influenced by, say, contextual parameters). Among many other problems, the firmness account errs by maintaining that E confirms H in cases when $c_k(H \mid E)$ is high simply because the prior $c_k(H)$ is high. In fact, a firmness account may say that E confirms H relative to c_k even though $c_k(H \mid E)$ is lower than $c_k(H)$ (as long as $c_k(H \mid E)$ is nevertheless high)! The Bayesian account focuses on the relation between E and H—how E would *alter* the agent's opinion of H—rather than just on where that opinion would land were E taken into account.

We can provide more information about E's effect on the agent's opinion of H by measuring the degree of incremental confirmation. The simplest way to measure confirmation is to calculate $c_k(H \mid E) - c_k(H)$; this measure simply asks how much conditionalizing on E would increase the agent's confidence in H. Yet as a measure of E's bearing on H, this simple difference has some drawbacks. For example, the degree to which E can confirm H will be limited by the value of $c_k(H)$. If, say, $c_k(H) = 0.99$, then even if E entails H, the maximal degree to which it can confirm H will be 0.01. Bayesian confirmation theory thus has a considerable literature proposing and assessing alternative measures of confirmational strength; see Crupi (2016, Section 3.4) for a recent summary and references.

One upshot of the literature on measuring confirmation is a new approach to "solving" traditional paradoxes of confirmation. For example, we usually think that universal generalizations are confirmed by their positive instances. The hypothesis that all ravens are black is typically confirmed by the evidence that a particular raven is black.[33] In symbols, $(\forall x)(Rx \supset Bx)$ is confirmed by $Ra \& Ba$. But now suppose we discover an item that is a non-black non-raven. The evidence $\sim Ba \& \sim Ra$ is a positive

---

32 Carnap was well-acquainted with this mistake, having made it himself in the first 1950 edition of his *Logical Foundations of Probability*.

33 I say "typically" because it is possible to generate a deviant background corpus against which it would be reasonable for the observation of a black raven to *dis*confirm that all ravens are black. (For examples, see Swinburne, 1971, and Rosenkrantz, 1977, Chapter 2.) The generation of the paradox doesn't rely on such deviant corpora, so we will set them aside for the rest of the discussion.

instance of the generalization $(\forall x)(\sim Bx \supset \sim Rx)$, so it should confirm that generalization. Yet the latter generalization is (by contraposition) logically equivalent to our former one. So by the first property of confirmation I endorsed above, $\sim Ba \,\&\, \sim Ra$ should confirm that all ravens are black. This is Hempel's (1945) famous "Paradox of the Ravens," which seems to generate the absurd conclusion that a hypothesis about the color of ravens may be confirmed by the observation of a white shoe.

Recently, a number of Bayesian confirmation theorists have conceded that perhaps a white shoe does confirm that all ravens are black—it's just that observing a white shoe confirms this hypothesis much *less* than observing a black raven would.[34] Fitelson and Hawthorne (2010), for instance, specify conditions on c_k such that as long as these conditions are met, evidence of a black raven will confirm the ravens hypothesis much more strongly than evidence of a non-black non-raven, on virtually every proposed measure of confirmation in the literature. It's highly plausible that most of us in the real world have credence distributions satisfying Fitelson and Hawthorne's conditions, accounting for our intuitions about the asymmetry of favoring in this case. Similar approaches have been taken to the problem of irrelevant conjunction (Hawthorne & Fitelson, 2004) and Goodman's (1955) grue paradox (Chihara, 1981; Eells, 1982).

2.2   *Decision Theory*

Since this handbook contains an extensive article on decision theory (Thoma, this volume), I will give only a brief sketch here. In formal decision theory, an agent is confronted with a decision problem, represented by a partition of acts she may perform. Once she performs an act, some outcome will occur, and the agent values different outcomes to different degrees. These valuations are represented by a utility function, which assigns real-number utilities to each possible outcome. (The key assumption about utilities is that they measure value *uniformly*—the agent takes each added unit of utility to be as valuable as the next. The same is not true of money; your first dollar may be much more valuable to you than your billionth.)

So what's difficult about that—shouldn't the agent just choose the act leading to the most valuable outcome? The trouble is that the agent may be uncertain which acts will lead to which outcomes. Put another way, the agent may be unsure what state the world is in, and the outcome that follows her decision may depend both on the act she chooses and on the remaining state of the world. For example, suppose I'm trying to decide whether to go into my office tomorrow. I know that if I go, it *may* be quiet

---

34 Though the idea dates all the way back to Hosiasson-Lindenbaum (1940).

and peaceful there, in which case I'll get a great deal of writing done, which is an outcome I highly value. On the other hand, there may be loud construction happening outside my office window, in which case I'll dally on the internet and get no writing done, an outcome to which I assign little utility. Since I don't know the state of construction around my building tomorrow, it's unclear to me which available act (go into the office, stay home) correlates with which outcomes, complicating my decision.

The standard solution to this problem is to have the agent assign an expected value to each available act. An agent's expected value for an act is her expectation for the amount of utility that will accrue if she performs the act—calculated using her credences that various states of the world obtain. Given a decision between two acts, a rational agent prefers the act to which she assigns the higher expected value (and is indifferent in case of ties). We can thus use her credence and utility assignments to develop a preference ordering over the acts available to her in any decision problem.

For example, suppose I assign a utility of 100 to a day of peaceful writing at my office, but a utility of 0 to spending the day there with construction going on. If I'm 40% confident there'll be no construction tomorrow, my expected utility of going into the office is

$$
\begin{aligned}
EU(\text{go to office}) = {} & c(\text{no construction}) \cdot u(\text{peaceful writing}) \\
& + c(\text{construction}) \cdot u(\text{wasted day}) \\
& = 0.40 \cdot 100 + 0.60 \cdot 0 \\
& = 40,
\end{aligned}
\tag{22}
$$

where the function u designates the amount of utility I assign to a given outcome. Given this expected utility for going to the office, I should prefer to stay home only if I expect doing so to yield me a utility greater than 40.

We can prove that if an agent sets her preferences by maximizing expected utility, her preference ordering over acts will satisfy various intuitive conditions, commonly known as the "preference axioms." For example, her preferences will be asymmetric (she never prefers both A to B and B to A) and transitive (if she prefers A to B and B to C, then she prefers A to C).

As I said, I'm going to avoid the many subtleties of developing a full-blown decision theory. One crucial concern is cases in which the agent's act may be correlated with the state of the world. Evidential decision theorists (Jeffrey, 1965) respond by working with the agent's credence in a state *conditional* on her performing a particular act, while causal decision theorists (Gibbard & Harper, 1978; Lewis, 1981; Joyce, 1999; Weirich, 2012) consider the agent's credence that her act will *cause* a particular state to obtain. Another concern is modeling risk-averse agents—such as an agent who prefers a guaranteed payout with utility 1 to a fair coin flip on which heads yields a prize with utility 3 (Allais, 1953; Buchak, 2013).

There is, however, one more notion from decision theory that we'll need in what follows: fair betting price. Consider a proposition P and a betting slip that guarantees its possessor $1 if P turns out to be true. How much is that betting slip worth to you? That depends how confident you are that P obtains. If you're certain of P, that slip is worth $1 to you. If you're certain P is false, the slip is worth nothing. In between, the more confident you are of P, the more value you assign to the betting slip.

To be more precise, your expected value in dollars of the fair betting slip is $c(P) \cdot \$1$. We call this your *fair betting price* for this gamble on P. In general, if a bet pays out $\$X$ dollars when P is true, your fair betting price for the bet is

$$c(P) \cdot \$X. \tag{23}$$

What does it mean to say this is your fair betting price? Suppose someone offers to sell you a betting slip that pays off on P. Your fair betting price is the price at which you'd expect to break even on such an investment. Assuming you value money linearly (so that each additional cent confers the same amount of additional utility on you), decision theory says that you should be willing to purchase the betting slip for any amount lower than your fair betting price, and indifferent about buying it at exactly your fair betting price. Conversely, if you possess such a slip, you should be willing to *sell* it for any amount *above* your fair betting price.

2.3 *Other Applications*

Historically, confirmation and decision theory have been major drivers of Bayesianism's development and the two most common applications to which the approach has been put. But the Bayesian theory of credences has been applied to many other philosophically significant topics as well. Here are a few examples.

○ Probabilities have been used to measure when the propositions in a set cohere. Coherentism about justification has then been evaluated by asking whether coherence among propositions makes it rational to invest a higher credence in each of them. See Shogenji (1999), Bovens and Hartmann (2003), Huemer (2011), and Olsson (2017).

○ It's been debated whether an agent who updates by conditionalization will thereby increase her credence in the hypothesis that best explains evidence observed. Van Fraassen (1989) argues that Bayesianism is incompatible with Inference to the Best Explanation. Replies have been offered by, *inter alia*, Okasha (2000), Lipton (2004), Weisberg (2009), and Henderson (2013).

○ Elga (2007) argues that when an agent discovers that an epistemic peer has assigned different credences than her based on the same evidence, that agent should move her credences closer to her peer's. A great deal of debate has ensued about whether such conciliationism is the rational response to peer disagreement. Christensen (2009) presents a useful survey that is unfortunately now outdated; Christensen and Lackey (2013) is a more recent collection. (Though plenty has been published on the subject since then!)

○ The peer disagreement controversy intersects with broader questions about the rational response to higher-order evidence—evidence concerning whether one has responded rationally to one's evidence. New essays on higher-order evidence and its connection to disagreement may be found in Rasmussen and Steglich-Petersen (forthcoming).

○ Peer disagreement is also an aspect of social epistemology, which has considered for decades how groups and individuals should combine the opinions of multiple experts to form a coherent single view. The literature on probabilistic opinion pooling dates back at least to Boole (1952). More recent discussions, with copious additional references, include Bradley (2007), Russell, Hawthorne, and Buchak (2015), and Easwaran, Fenton-Glynn, Hitchcock, and Velasco (2016).

3   ARGUMENTS FOR CREDAL CONSTRAINTS

Many of the constraints on credences presented in Section 1 have an intuitive claim on being rationally required. It's just plausible that the more confident you are it will rain tomorrow, the less confident you should be that it won't rain. But can we provide *arguments* for the various rational constraints? Here I'll survey three historically-significant approaches to arguing for rational constraints on credence.

3.1   *Representation Theorem Arguments*

In Section 2.2 I suggested that if an agent has credence and utility functions, decision theory can combine these to determine her rational preferences among acts. But decision theory can also work in the opposite direction. Suppose I observe an agent make a number of decisions over her lifetime. Assuming these choices express her preferences among acts, I can construct credence and utility functions for her that would rationalize such preferences if she is an expected utility maximizer. I might then use

these credence and utility functions to predict choices she'll make in the future.[35]

We can prove that as long as an agent's preferences are rational, she can be represented as maximizing expected utility by combining credence and utility functions. More precisely, a representation theorem shows that given a preference ordering over acts satisfying certain preference axioms, there exists a utility function and a probabilistic credence function on which those preferences maximize expected utility. Since there are many different versions of decision theory, there are many sets of preference axioms, and so many different representation theorems.[36] But typically the preference axioms can be divided up into two sorts: substantive constraints such as the asymmetry and transitivity requirements I mentioned earlier; and what Suppes (1974) calls "structure axioms" specifying that the preference ordering is complete, has acts available at a variety of levels of preference, etc. (Structure axioms are usually considered a convenience to make the theorems cleaner and the proofs easier.)

Representation theorems can be highly useful. For instance, economists engaged in rational choice theory often model market participants as maximizing expected utility based on a utility function and a probabilistic credence function. A representation theorem assures us that as long as an agent remains rational—in the sense of making choices that satisfy the preference axioms—her behavior will continue to conform to such a model.

Yet there's a big step from arguing that rational agents can be *modeled* as employing a probabilistic credence function to arguing that rational agents actually *possess* probabilistic credence functions (Hájek, 2009; Meacham & Weisberg, 2011). We can begin to see the problem by noting that an agent's preferences will often underdetermine her utility and credence distributions. That is, if all we know is an agent's preferences, there are (infinitely) many different pairs of utility and credence functions that will generate that preference ordering by maximizing expected utility. Moreover, many of those pairs feature credence functions that don't satisfy the probability axioms. Standard representation theorems prove only that if an agent's preferences satisfy the axioms, *there exists* a corresponding credence/utility pair in which the credence function satisfies the probability rules. This hardly shows that rationality requires probabilistic credences.

---

35 We can think of this as a formalization of the folk deployment of a "theory of mind." I watch what you do, I surmise what you want and what you believe, then I let that information guide my interactions with you going forward.

36 Representation theorems were inspired by early, suggestive results in Ramsey (1931). The first rigorous representation theorem of the type we're discussing is in Savage (1954). (Though see also von Neumann and Morgenstern, 1947.) A representation theorem for evidential decision theory appears in Jeffrey (1965), while Joyce (1999) proves one for causal decision theory.

Matters can be improved with a representation theorem based on some ideas Lara Buchak and I came up with together. (A sketch of a proof appears in the Appendix.) This theorem shows that if an agent's preferences satisfy various preference axioms, and she maximizes expected utility, then her credence function must be a *positive scalar transformation* of a probability distribution. In other words, her credences will be nonnegative, they will be finitely additive, they will assign the same value to every tautology, and that value will be greater than the value assigned to contradictions. A credence function like this will have all the same properties as a probabilistic function, except that the maximal value it assigns to tautologies may be some positive number other than 1. Yet nothing substantive hangs on whether we measure credence on a 0 to 1 scale or instead, say, a percentage scale from 0 to 100.

Still, even the improved theorem assumes that the agent's credences and utilities interact with preferences through the maximization of expected utility. Zynda (2000) notes that there are many other mathematical quantities combining credence and utility that an agent could choose to maximize. So to argue for probabilism (or something close to it) using one of these representation theorems, we need to assume not only that rationality requires satisfying the preference axioms, but also that it requires maximizing expected utility.

3.2   *Dutch Book Arguments*

As with representation theorems, an inspiration for Dutch Book arguments can be found in Ramsey's (1931), in which he commented,

> These are the laws of probability, which we have proved to be necessarily true of any consistent set of degrees of belief. . . .
> If anyone's mental condition violated these laws, his choice would depend on the precise form in which the options were offered him, which would be absurd. He could have a book made against him by a cunning better and would then stand to lose in any event. (p. 84)

Suppose, for instance, that I am both 0.7 confident that I will go to my office tomorrow and 0.7 confident that I will not. Now consider two betting slips—one that pays a dollar if I go to the office, and another that pays a dollar if I don't go to the office. Given my credences, my fair betting price for each of these slips is $0.70. That means I'm willing to pay up to $0.70 for each of them. So suppose I buy both, at a price of $0.70 each. I've now spent a total of $1.40, and no matter what happens tomorrow, I will only make $1. My non-probabilistic credence distribution has made me

susceptible to a combination of bets on which I will lose $0.40, come what may!

De Finetti (1937/1964) proved that if an agent's credences violate the probability axioms, a set of bets exists such that if the agent purchases each of them at her fair betting price, she will lose money in every possible world. For unknown reasons, such a set of bets is called a "Dutch Book." The proof works by going through each of the axioms one at a time, and showing how to construct a Dutch Book against an agent who violates the relevant axiom. Moreover, we can establish what Hájek (2009) calls a "Converse Dutch Book Theorem," showing that if an agent *satisfies* the probability axioms, no Dutch Book of the types described in de Finetti's proof can be constructed against that agent.

Other proofs show how to construct Dutch Books against agents who violate the Reflection Principle (van Fraassen, 1984), the Principal Principle (Howson, 1992), Regularity (Kemeny, 1955; Shimony, 1955), and Countable Additivity (Adams, 1962). We can also construct what is known as a "Dutch Strategy" against any agent who violates Conditionalization (Teller, 1973, reporting a result of David Lewis') or Jeffrey Conditionalization (Armendt, 1980; Skyrms, 1987b). A Dutch Strategy is not strictly speaking a particular set of bets guaranteed to give the agent a sure loss; instead, it's a strategy for placing bets with the agent in which certain bets are placed at an initial time, then future bets are placed depending on what the agent learns after that time. Still, the idea of a Dutch Strategy is that no matter what happens (and no matter what the agent learns), if she purchases the bets at her fair betting prices when they're offered, she'll face a net loss come what may.

Avoiding Dutch Books and Dutch Strategies seems an important advantage for the probabilistic agent. Still, can we *argue* that rationality forbids susceptibility to Dutch Strategies and Books? One problem is that the negative effects of violating probabilism highlighted by Dutch Books seem oddly practical. We might have thought that the Kolmogorov axioms provided constraints of *theoretical* (rather than *practical*) rationality on agents' credences. Yet here we're arguing for those axioms by pointing to financial consequences of violating them. Moreover, it's unclear how seriously we should take those potential consequences. Are non-probabilistic agents ever really going to face the precise set of bets that would expose them to a Dutch Book? And what if the non-probabilistic agent has read about Dutch Books, and decides that instead of changing her credences, she'll just be more cautious in her betting behavior? In the example above concerning my going to the office, I might pay $0.70 for the bet that pays off if I go into the office, but then refuse to buy the second bet because I see a Dutch Book coming. In that case I'll still have non-probabilistic credences, but will manage by practical strategizing to avoid the prospect of a sure loss.

Taking a cue from the second sentence of the Ramsey quote above, a number of authors have tried to "depragmatize" Dutch Book arguments. Skyrms writes that "For Ramsey, the cunning bettor is a *dramatic* device and the possibility of a dutch book a striking *symptom* of a deeper incoherence" (Skyrms, 1987a, p. 227, emphases mine). For these authors,[37] susceptibility to Dutch Book merely brings out an underlying inconsistency in the agent's credences—the inconsistency of evaluating the same thing different ways depending on how it's presented.

Return to my bets on whether I'll go into the office tomorrow. Given my 0.7 confidence that I'll go, my fair betting price for a bet that pays $1 if I go and nothing otherwise is $0.70. So I value that bet at $0.70; if I'm offered the opportunity to purchase that bet at any lower amount—say, $0.50—I'd consider that a favorable deal. On the other hand, my 0.7 confidence that I won't go gives me a fair betting price of $0.70 for a bet that pays $1 if I *don't* go and nothing if I do. So I would consider it unfavorable to sell that bet at any price less than $0.70—for instance, $0.50. Yet buying the first bet at $0.50 and selling the second bet at $0.50 *are the exact same transaction*; each one would net me $0.50 if I go to the office and lose me $0.50 if I don't. So do I view that transaction favorably or not? One of my credences suggests I view it favorably, while the other demands I don't. How those credences evaluate those bets reveals the conflict between them.[38]

Still, even depragmatized Dutch Book arguments make potentially controversial assumptions. First, we're assuming that a rational agent's fair betting prices equal her expected payouts—an assumption that might fail for risk-averse agents. And second, to construct a Dutch Book against some violations of Finite Additivity, we need to assume a "package principle"— that a rational agent's fair betting price for a combination of two bets equals the sum of her betting prices for each bet considered singly. Each of these assumptions would follow easily if we assumed that rational agents always choose to maximize expected utility. But if we could assume *that*, we'd already have a representation-theorem argument for something very close to probabilism (Section 3.1).[39] So it's unclear why the detour through cunning bettors would be required.

---

37 See also Armendt (1992), Christensen (2004), and Howson and Urbach (2006).
38 Notice that I wouldn't have this problem if I satisfied the probability calculus by, say, assigning credence 0.7 that I'll go and credence 0.3 that I won't. In that case I'll look favorably on buying the first bet at $0.50 and also look favorably on selling the second one at $0.50, so my evaluations will harmonize.
39 In fact, the representation theorem proof in the appendix closely mirrors the structure of traditional Dutch Book theorems for probabilism.

3.3 *Accuracy Arguments*

In his 1998, James M. Joyce sets out to provide a "nonpragmatic vin-
dication of probabilism" that would explicitly avoid invoking practical
consequences in its defense of the probability axioms as rational con-
straints on credence. His work builds on mathematical results from de
Finetti (1974) and Rosenkrantz (1981), but uses those results to construct a
new kind of argument.

   Joyce's key idea is that from a point of view of pure theoretical rationality,
agents should aim to make their credences as accurate as possible. How
might we measure the accuracy of a credence function? Historically, one
option had been to focus on calibration. Function c is perfectly calibrated
if, for every $0 \leq x \leq 1$, when we look at all the propositions in $\mathcal{L}$ to which
c assigns credence x, the fraction of those propositions that are true is
exactly x. If I'm perfectly calibrated, exactly half of the propositions to
which I assign credence $1/2$ are true, exactly a third of the propositions to
which I assign credence $1/3$ are true, etc.

   Van Fraassen (1983) and Shimony (1988) argue for probabilism by show-
ing that in order for a credence distribution to be embeddable in larger and
larger systems approaching perfect calibration, that credence distribution
must satisfy the probability axioms. This might stand as a good argument
for probabilism, except that calibration has some intuitively undesirable
features as a measure of accuracy. For example, consider two agents who
assign credences to four propositions as in Table 1. I hope you'll agree

	A	B	C	D
AGENT 1	0.5	0.5	0.5	0.5
AGENT 2	1	1	0.01	0
TRUTH-VALUES	T	T	F	F

Table 1: Two credence assignments

that intuitively, Agent 2's credences are much more accurate (close to the
truth) than Agent 1's. Yet Agent 1 is perfectly calibrated—exactly half the
propositions to which she assigns credence $1/2$ are true—while Agent 2 is
not.

   Our intuitions about accuracy work by looking at each credence assign-
ment one at a time, assessing how accurate that credence is given the
truth-value of the proposition, and then aggregating those local accuracy
assessments across all the propositions. Yet calibration works with *global*
features of a probability distribution, which (as we've just seen) can lead
to distorting effects.

So Joyce uses a gradational accuracy approach instead. On this approach, we select a scoring rule to measure how far each individual credence assignment to a proposition is from the truth about that proposition. Intuitively, when proposition P is true, higher credences in P are more accurate; when P is false, lower credences are better. We can formalize this by having a function I that assigns 1 to P if it's true and 0 if it's false, then measuring how far $c(P)$ is from $I(P)$. Historically, it's been popular to measure this distance as

$$(I(P) - c(P))^2. \tag{24}$$

Notice that this measurement *increases* the *farther* you are from the truth; so it's a measure of credal *in*accuracy. A rational agent aiming to be as accurate as possible should look to minimize this quantity for each proposition. Globally, she should look to minimize the sum of this quantity across all the propositions she entertains. (This sum is commonly known as the *Brier score*, named for meteorologist George Brier's discussion of it in his 1950.)

Joyce shows that if we use the Brier score to measure accuracy, then any non-probabilistic credence distribution will be accuracy-dominated by another, probabilistic distribution over the same set of propositions. That is, if you take an agent whose credences over some language violate the probability axioms, there will be another, probabilistic credence distribution over the same language that has a more accurate Brier score than hers *in every possible world*. When the nonprobabilistic agent considers that alternative distribution, she will know that it's more accurate than hers, even without knowing anything about which possible world is actual. Joyce argued that for an agent to maintain her nonprobabilistic distribution, despite this information that another distribution was certainly more accurate, would be irrational. And since the same situation will confront any agent whose credences violate the probability axioms, this constitutes an argument for probabilism.[40]

Related accuracy arguments have been offered for a variety of other Bayesian norms: Conditionalization (Greaves & Wallace, 2006; Briggs & Pettigrew, forthcoming), the Principal Principle (Pettigrew, 2013), the Principle of Indifference (Pettigrew, 2014), Reflection (Easwaran, 2013), and Conglomerability (Easwaran, 2013).

There are two main concerns in the literature about these accuracy arguments. First, there's a general concern about assessing the rationality of credences by measuring their distance to the truth. The gradational accuracy approach evinces a sort of epistemic consequentialism, in which

---

40 Importantly, the same kind of argument cannot be run *against* probabilism. A credence function that satisfies the probability axioms will not be accuracy-dominated in the manner Joyce describes by any other function (probabilistic or otherwise).

attitudes aim for some outcome (in this case, truth), and are evaluated by how well they approximate that goal. Just as teleological approaches to normativity have aroused suspicion in ethics and other areas of philosophy, the gradational accuracy program has been criticized by such authors as Greaves (2013), Berker (2013), and Carr (2017).

Second, among those who accept the gradational accuracy program, there's a concern about how to select an appropriate scoring rule for measuring accuracy. Maher (2002) suggests that instead of using the Brier score, we might gauge the distance between an individual credence $c(P)$ and a truth-value $I(P)$ by calculating

$$|I(P) - c(P)|. \tag{25}$$

Historically, the Brier score was favored over this absolute-value score because the former is a "proper" scoring rule while the latter is not. To understand the difference, suppose a six-sided die has just been rolled, and we have two characters who do not yet know the outcome. Our first character, Chancey, assigns credence 1/6 to each of the possible outcomes. Our second character, Pessimist, assigns credence 0 to each outcome. Chancey's credence function satisfies the probability axioms, while Pessimist's does not.

Now suppose each of our characters calculates an expected inaccuracy value for herself and for the other person. To give an example of how this works, suppose Chancey calculates an expected inaccuracy value for her own distribution using the Brier score. To do so, Chancey considers each of the six possible worlds available (that is, each of the six possible outcomes of the die roll), evaluates what her Brier score would be in that possible world, multiplies by her credence that that possible world is actual, then sums across all the possibilities. If, for instance, the die roll comes up 3, Chancey's Brier score will be

$$
\begin{aligned}
& (I(1) - c(1))^2 + (I(2) - c(2))^2 + (I(3) - c(3))^2 \\
& \qquad + (I(4) - c(4))^2 + (I(5) - c(5))^2 + (I(6) - c(6))^2 \\
= \ & (0 - 1/6)^2 + (0 - 1/6)^2 + (1 - 1/6)^2 \\
& \qquad + (0 - 1/6)^2 + (0 - 1/6)^2 + (0 - 1/6)^2 \\
= \ & 1/36 + 1/36 + 25/36 + 1/36 + 1/36 + 1/36 \\
= \ & 30/36 \\
= \ & 5/6.
\end{aligned}
\tag{26}
$$

A bit of reflection will show that this is Chancey's Brier score in each of the six possible worlds. So her expected Brier score across all those worlds is also 5/6. In the meantime, I'll leave it to the reader to calculate that Pessimist's expected Brier score is 1. Since higher scores mean more

inaccuracy—and less accuracy—Chancey expects her credences to be more accurate than Pessimist's when the Brier score is used to calculate accuracy.

Exactly the opposite happens if we use the absolute-value measure. Again, I'll leave it to the reader to calculate that Chancey's expected absolute-value score is 5/3, while Pessimist's is again 1. So by the lights of the absolute-value score, the nonprobabilistic Pessimist is expected to be more accurate than the probabilistic Chancey.

Proper scoring rules are rules on which a probabilistic agent will never expect some other agent to be more accurate than herself. The Brier score is one of many proper scoring rules, while the absolute-value score is improper. In general, it seems irrational for an agent to hold onto a credence distribution when she expects some other agent's credences to be more accurate than her own (Lewis, 1971). So a theorist who has already accepted that probabilistic distributions are rational has good reason to work with proper scoring rules rather than improper ones. The accuracy-based arguments for Conditionalization, the Principal Principle, the Indifference Principle, etc. mentioned above all confine themselves to working with proper scoring rules.

Predd et al. (2009) show that Joyce's accuracy-dominance argument for probabilism could be run using any proper scoring rule. Yet in the context of an argument for probabilism, favoring proper scoring rules over improper ones seems question-begging. Proper scoring rules are defined as those on which probabilistic distributions are rated more expectedly accurate than the alternatives. Unless you have an antecedent reason to think probabilistic distributions *should* come out looking better than the alternatives, this is no reason to prefer a proper score.[41]

4   ARGUMENTS AGAINST CREDAL CONSTRAINTS

Having surveyed some arguments in *favor* of various rational constraints on credences, what are the arguments *against* these constraints? Of course there are many, and they multiply over time. Here I will focus on a handful that have generated insightful discussion and interesting positive responses.

4.1   *The Problem of Logical Omniscience*

Savage (1967) famously considered the plight of "a person required to risk money on a remote digit of π." His concern was that according to the Normality axiom, an agent is required to assign certainty to every tautology in her language $\mathcal{L}$. Arguably, the fact that a given digit of π

---

41 Though there may be other reasons. See, e.g., Joyce (2009) and Pettigrew (2016).

PRECISE CREDENCES   37

takes a particular value is a tautology.[42] So according to probabilism, a rational agent should be certain of all the digits of π. Yet this seems too much for rationality to demand of any real agent.

Savage's discussion initiated a literature on what is known as "the problem of logical omniscience." I actually think there are multiple, related problems here, which we might label as follows.[43]

CREDAL COMPLETENESS. Probabilism requires an agent to assign a credence to each proposition in her language.

LOGICAL DISCERNMENT. Probabilism forbids an agent from assigning a credence other than 1 to any tautology.

LOGICAL LEARNING. A probabilistic agent will never pass from a lower credence in a tautology to a higher credence.

The problem of Credal Completeness is that the probability axioms require an agent to assign a credence to every proposition in her language. For instance, Non-Negativity says that every $X \in \mathcal{L}$ receives some non-negative credence value. Even in a language with finitely many atomic propositions, closure under truth-functional connectives will generate a language of infinite size. Yet it seems not only impossible for a finite agent to assign that many credences, but also inadvisable under Harman's (1986) principle of Clutter Avoidance.

CLUTTER AVOIDANCE. One should not clutter one's mind with trivialities.

Yet we can slightly alter our formalism so that it no longer demands credal completeness and evades clutter avoidance concerns. The idea is to require not that an agent's credence distribution actually satisfy the probability axioms, but only that it be *extendable* to a distribution that does. In other words, we permit an agent to adopt a partial credence distribution that assigns numerical values to only some of the propositions in $\mathcal{L}$, but we require that there be some possible way of assigning values to the rest of $\mathcal{L}$ so that the resulting full distribution satisfies the axioms. This approach recovers intuitive results such as the stricture that if an agent assigns credences to both P and $\sim P$, those credences must sum to 1. But it will not fault an agent if she fails to adopt attitudes towards P, $\sim P$, or both.

Moving to partial distributions avoids the problem of Credal Completeness, but leaves the problem of Logical Discernment intact. It seems

---

42 If your views about logicism in the philosophy of mathematics entail that facts about digits of π are not tautologies, we can always substitute in a conditional whose antecedent is various arithmetic axioms and whose consequent reports a digit of π. Or we can work instead with some highly complex logical truths.

43 The "Logical Learning" label is common in the literature; I invented the other two labels for our discussion here.

perfectly rational for me to assign credence 1/10 that the trillionth digit of π is a 2. Yet any credence distribution—partial or complete—containing that assignment is not extendable to a probabilistic distribution. It's either a tautology that the billionth digit is a 2, or it's a tautology that the billionth digit isn't, so probabilism either demands that I assign that proposition a credence of 1 or demands that I assign it a credence of 0. Whichever is the true demand, it seems a bit too demanding, since I don't have any good way to figure out which demand it is.

Before considering responses to this problem of Logical Discernment, let's quickly consider Logical Learning. The following credal sequence seems quite reasonable: I assign credence 1/10 that the trillionth digit of π is a 2, Talbott (1991) tells me that it is indeed a 2, so my credence that it is dramatically increases (perhaps all the way to 1). It seems in this case that I have learned a logical truth, and my credal increase is a rational response to that learning episode. Yet a traditional Bayesian system will not approve of this response, or be able to usefully model it, since a probabilistic system countenances only credence distributions (at any time) that assign that proposition a value of 1.

If we solved the Logical Discernment problem by building a Bayesian theory that allowed rational credences in tautologies other than 1, presumably that theory would also allow increases and decreases in such credences. So there's hope that a solution to Logical Discernment would open up a solution to Logical Learning.

How, then, might we model a Bayesian agent without perfect logical discernment? Responding to Savage, Hacking (1967) suggests we identify a proposition as "personally possible" for an agent if the agent doesn't know it's false. We then adjust Normality to demand certainty only in propositions whose negations are personally impossible, and Finite Additivity to apply only when $P \& Q$ is personally impossible. This allows an agent to be ignorant of arbitrarily many logical truths, and therefore less-than-certain of those truths.

Yet this approach creates three problems. The first is formal. Hacking works with credence distributions over sentences, and he's free to treat whatever sentences he wants as personally possible or impossible. But if we think of those sentences as representing underlying propositions, and those propositions in turn as representing underlying sets of possibilities, it seems natural to ask what possibilities an agent entertains when she entertains as personally possible that which is logically impossible. To address this sort of gap, Hintikka (1975) constructs a semantics admitting of logically impossible worlds, which can enter into the content of propositions in just the manner of classical possible worlds.

A second, intuitive problem is that Hacking's approach allows for arbitrarily large amounts of logical non-omniscience—nothing in Hacking's

formalism indicts an agent who assigns less-than-certainty to $P \vee \sim P$, as long as that agent doesn't know the proposition is true. Bjerring and Skipper (manuscript) complain Hacking's formalism is so permissive that in sacrificing logical omniscience, it fails to capture any rational requirement of basic logical competence. They make similar complaints about a framework from Garber (1983), and various formalisms developed using Hintikka's semantics.

Finally, it's important to see what a Bayesian system loses when it's redefined in terms of personal rather than logical possibility. If an agent fails to know that $P \& \sim P$ is impossible, then by Hacking's lights she need not apply Finite Additivity to P and $\sim P$. As a result, such an agent may assign P and $\sim P$ credences summing to more than 1. She may increase her credence in P without decreasing her credence in $\sim P$. In our relevance-based theory of confirmation, she may not see P as disconfirming $\sim P$. And when she selects actions by maximizing expected epistemic utility, she may violate the preference axioms in a variety of ways. In other words, the very features and applications that make Bayesianism a plausible picture of rationality begin to dissolve once logical discernment requirements are loosened.

So perhaps we should go in the other direction? A number of theorists have begun to wonder if logical omniscience requirements are not an annoying side-effect of our epistemic formalisms, but instead a hint from those formalisms about the underlying normative domain. Smithies (2015) argues that certainty in logical truths is in fact a requirement of rationality; Titelbaum (2015) and Littlejohn (2018) advocate related positions.

4.2 *The Problem of Old Evidence*

Clark Glymour initiated the Old Evidence debate with a famous example.

> Scientists commonly argue for their theories from evidence known long before the theories were introduced.... The argument that Einstein gave in 1915 for his gravitational field equations was that they explained the anomalous advance of the perihelion of Mercury, established more than half a century earlier. Other physicists found the argument enormously forceful, and it is a fair conjecture that without it the British would not have mounted the famous eclipse expedition of 1919. Old evidence can in fact confirm new theory, but according to Bayesian kinematics, it cannot. (Glymour, 1980, pp. 306–7)

We've already seen (Section 1.3 and Section 1.4) that a traditional Bayesian models evidence acquisition as the gaining of certainties, which are then retained. At the same time (Section 2.1), confirmation is understood as

positive relevance. Combining these two approaches, we have a problem: once an evidential proposition has been learned, it receives credence 1. When $c(E) = 1$, $c(H \mid E) = c(H)$ for any $H \in \mathcal{L}$. So once an agent learns something, that piece of information is confirmationally inert ever after.

Given these basic facts about Bayesianism, we can identify two challenges in Glymour's story about Einstein. Christensen (1999) calls them the "synchronic" and "diachronic" problems of old evidence.[44] The diachronic problem is about changes in credence. Over the course of 1915, Einstein increased his confidence in the General Theory of Relativity (GTR), and we think this had something to do with the perihelion of Mercury. Yet it can't be that Einstein increased his confidence because he learned of the anomalous advance—he already knew about that well before 1915. So what changed his opinion, and how can we reflect it in a Bayesian system?

The synchronic problem of old evidence comes up after 1915, when the perihelion of Mercury has already had its effect on Einstein's attitudes toward GTR. Presumably even after 1915, Einstein would have cited the perihelion advance of Mercury as a crucial piece of evidence supporting GTR. Yet relative to Einstein's credence function at that time—which assigns 1 to the perihelion facts—those facts are not positively relevant to GTR. So how can a Bayesian about confirmation interpret that evidential support?

Proposals to solve the synchronic problem usually work by relativizing confirmation to some probability function other than the agent's current credence distribution. Since the agent currently assigns $c(E) = 1$, E can't confirm anything relative to that current distribution. So we look for some other relevant distribution that doesn't assign 1 to E. For instance, we might adopt a "historical backtracking" approach on which we look back to some time when the agent wasn't yet certain of E, and ask whether E was positively relevant to H in her credence distribution at that time. But this approach is limited for a number of reasons. For instance, Einstein probably knew about the perihelion of Mercury long before he ever considered GTR. So if we backtrack to a time well before 1915 when he wasn't yet certain of E, we won't be able to find any conditional or unconditional credences he assigned to the relevant H at that time. And so we won't be able to say that E confirms H for Einstein now because at some time in the past he assigned $c(H \mid E) > c(H)$.

In light of this and other difficulties, Howson and Urbach (2006) advocate a "counterfactual backtracking" approach. Instead of looking to a time in the past when the agent didn't know E, we look to a close possible world in which the agent knows everything she knows now except E. Well,

44 I'm using Christensen's terminology because I find it the most helpful. But earlier, related disambiguations of the problem of old evidence can be found in Garber (1983), Eells (1985), and Zynda (1995).

not quite *everything*—we will probably also want a world in which she doesn't know logical equivalents to *E*, immediate entailments of *E*, etc. But Howson and Urbach (p. 300) have a technical proposal for identifying the propositions that should be subtracted out. Setting the technical details aside, Earman (1992, p. 123) worries this counterfactual approach will suffer from similar defects to other counterfactual analyses; moving to a non-actual world may have side-effects that spoil the analysis. For example, the historical record suggests that Einstein was motivated to formulate GTR in part to explain Mercury's anomalous advance. So the closest possible world in which Einstein doesn't know *E* yet still assigns credences to *H* may be very far—and very different from our own—indeed.

Perhaps the best approach is to say that when an agent explains the evidence supporting some hypothesis, the support she's describing may be relative not to her own personal credences but to some other probabilistic distribution. That distribution may be one assumed pertinent by her audience, or by a particular scientific community. Or if we are Objective Bayesians (Section 1.5), it may be the objective distribution that determines how all rational agents should set their credences. Maher (1996), for instance, develops a proposal of the latter sort. Yet many details remain to be resolved. For example, how does either a scientific community or an objective rational distribution assign a prior probability to the proposition that GTR expresses the physical laws of our universe?[45]

As for the diachronic problem of old evidence, the typical response is to identify something *other* than learning of Mercury's perihelion advance that gave Einstein new confidence in GTR over the course of 1915. For one, Einstein might have discovered sometime in 1915 not that Mercury's perihelion advances anomalously, but that GTR predicts such an anomalous advance. Since it's a logical fact that GTR (along with other empirical information of which Einstein was already aware) entails the details of the advance, this would be an instance of logical learning. So a Bayesian implementation of this explanation will depend on the logical omniscience issues discussed in Section 4.1.

Another possibility is that Einstein's high confidence in GTR at the end of 1915 was new because he hadn't had *any* attitude towards GTR at the beginning of 1915. Perhaps Einstein hadn't yet conceived of GTR at the beginning of 1915, so the language over which he assigned credences at that time didn't contain a proposition expressing GTR's truth. This approach would certainly explain why Einstein had a new, high credence at the end of the year that he didn't have at the beginning. But it probably doesn't generalize to all cases of confirmation by old evidence (and may not

---

45 Even if we wanted to use an Indifference Principle (Section 1.5) here, we'd need a partition to divide our credence evenly across, and it's difficult to determine what alternative sets of physical laws should go into such a partition.

even be historically accurate in Einstein's case). Moreover, cases in which agents add new propositions to their cognitive language pose another challenge for Bayesianism. All of the updating norms we have considered (Conditionalization, Jeffrey Conditionalization) work over a language that remains fixed over time. The so-called "problem of new theories" challenges us to build a formalism that allows an agent's language to change over time, and that places reasonable constraints on how the agent's credences should evolve across such changes.

Finally, we might focus on the fact that both versions of the problem of old evidence seem to arise because Conditionalization treats acquiring evidence as gaining certainties. If newly-acquired evidence didn't go to (and remain at) a credence of 1, then we wouldn't have the problem that old evidence always has credence 1 and therefore can't be positively relevant to anything. Suppose we adopt the Regularity principle (forbidding certainty in empirical propositions), and mandate Jeffrey Conditionalization as the rational updating scheme. Then evidence acquisition will increase credence in particular propositions, but never send it to 1, and the problem of old evidence will never arise.

Christensen (1999) pursues this approach and finds much to recommend it, but eventually encounters a new difficulty. The problem of old evidence is that acquiring a piece of evidence shouldn't rob it of its ability to confirm hypotheses. Generalizing this idea, we should agree that becoming more *confident* in a piece of evidence shouldn't affect the *degree* to which it confirms a hypothesis. So Christensen seeks a confirmation measure (Section 2.1) on which Jeffrey Conditionalizations that change $c(E)$ don't affect E's level of confirmation of H. He is unable to find a measure that satisfies this constraint, meets other plausible formal conditions, and works intuitively in examples.

4.3   *Memory Loss and Context-sensitivity*

Certainty acquisition and retention also pose other problems for a Conditionalization-based updating framework. For instance, many of us have the experience of gaining a piece of evidence one day and then forgetting it a short time later. Yet if we are constant conditionalizers, a proposition that achieves credence 1 at some time may never sink to a lower credence later. So Conditionalization deems memory loss irrational.[46]

---

46 Or at least, the version of Conditionalization we've been discussing deems memory loss irrational, because it governs an agent's updating across any arbitrary interval of times t_i to t_j. One might embrace a more limited version of Conditionalization (compare Titelbaum, 2013a, Chapter 6) that applies only across intervals during which the agent's information

While this problem was recognized at least as far back as Levi (1987), Talbott (1991) puts it particularly forcefully. He considers the response that Bayesian rules are meant to model ideally rational agents—not everyday agents—"and an ideally rational agent would not be subject to the fallibility of human memory." (p. 141) For what it's worth, I don't see why elephantine recall should make one agent more *rational* than another (though see Carr, 2015), but the whole question may be sidestepped by an ingenious example due to Arntzenius (2003). While I won't work through the details here, the upshot of Arntzenius's example is that Conditionalization indicts not only agents who actually forget evidence, but also agents who suspect they might have forgotten evidence (even if they actually haven't). Surely we can't require of ideally rational agents certainty in the empirical proposition that they have never forgotten anything in their lives!

Can we alter Conditionalization to allow for certainty loss? One popular approach is to take advantage of a feature traditional Conditionalization already displays. Suppose we have an agent who conditionalizes throughout her entire life. As she gains evidence, she will accumulate certainties; the total set of certainties she possesses at any time will represent her total evidence at that time. Let's refer to the proposition expressing the conjunction of all the agent's evidence/certainties at time t_i as E_i. If the agent is a faithful conditionalizer, there will exist at least one regular[47] probability distribution p_h such that for any time t_i at which that agent assigns credences, and any proposition X in her language $\mathcal{L}$, $c_i(X) = p_h(X \mid E_i)$. In other words, there exists a single function p_h relating to every moment in the agent's life, such that her credence distribution at any moment can be recovered by conditionalizing p_h on her total evidence at that moment.

I'll refer to this distribution p_h as the agent's *hypothetical prior*; it is sometimes also called an "*ur*-prior" or an "initial credence distribution." This last moniker comes from thinking of p_h as representing the agent's credences at some earliest moment in her life when she lacked any empirical certainties. Because conditionalization is cumulative and commutative, if an agent *did* have such an initial moment in her life—before her first update by Conditionalization—the credences she assigned at that time *would* relate to her later opinions in the way that p_h relates to c_i. Yet it's difficult to imagine that any actual agent has ever had a moment when she entirely lacked empirical information.

So I prefer to think of an agent's hypothetical prior as a convenient tool for separating out two influences on her credences. On the one hand,

---

strictly increases. In that case the problem would be that rather than deeming memory loss irrational, the limited updating rule fails to give any guidance in memory loss cases at all.

47 By saying the distribution is "regular," we mean that it assigns credence 1 only to logical truths.

there's her evidence; on the other, there are her epistemic standards, which encapsulate her principles and tendencies for interpreting evidence. The agent's total evidence changes over time, and is represented at time t_i by E_i. Yet as her evidence changes, she may retain a constant set of standards for interpreting evidence, represented by her hypothetical prior p_h. Applying these standards to the agent's total evidence at t_i—by conditionalizing p_h on E_i—yields her credence distribution c_i.[48]

This generally attractive picture is entailed by Conditionalization: *if* an agent conditionalizes at every update, *then* her credences throughout her life will be representable as faithful to a constant hypothetical prior. Yet interestingly, the entailment does not run in the opposite direction. That is, an agent may maintain fealty to a constant hypothetical prior even if her updates do not always satisfy Conditionalization. For instance, it's possible that an agent could both gain and lose certainties between two times t_i and t_j, and yet there still exists a single hypothetical prior p_h such that for every $X \in \mathcal{L}$, $c_i(X) = p_h(X \mid E_i)$ and $c_j(X) = c_j(X \mid E_j)$.

We can therefore achieve a plausible diachronic model of agents who both gain and lose certainties by generalizing Conditionalization not to demand that an agent conditionalize between each earlier time t_i and later time t_j, but instead to demand (whatever happens to her certainties) that she set her credences in line with a constant hypothetical prior throughout her life. This new diachronic norm generates plausible results for a number of forgetting stories, such as those featured by Talbott. In cases where an agent does strictly gain certainties between two times, it mimics the effects of traditional Conditionalization. And in cases where an agent strictly *loses* certainties between times, it gives us reverse-temporal Conditionalization. That is, the agent's earlier unconditional credences will equal her later credences conditional on the information she lost. Thus forgetting becomes like learning backwards in time.

Unfortunately, shifting to this new diachronic norm does not suffice alone to address another problem with Conditionalization: the way it treats context-sensitive information. Here I refer to "self-locating" claims that change their truth-values across times, persons, and locations—such as "Today is Tuesday," "I am a sailor," and "We are in Detroit." For one thing, to model these sorts of claims in our formalism we will need to add to our language $\mathcal{L}$ something like what Lewis (1979) called "centered propositions." But even then, Conditionalization will face challenges. It may be rational right now to be certain that it's Tuesday, but that certainty will not remain rational into perpetuity.

The context-sensitivity challenge is sometimes described as yet another problem with Conditionalization's certainty-retention. But even when we shift to a diachronic norm that requires fealty only to a constant

---

48 Compare Schoenfield (2014) and Meacham (2016).

hypothetical prior (and therefore allows for certainty loss), problems still remain. This is because the Bayesian system was designed to model agents whose evidence changed over time, but who used that evidence to evaluate hypotheses with truth-values that were fixed targets.[49] Adding in another level of shiftiness generates complications for Conditionalization, Jeffrey Conditionalization, and hypothetical priors.

A number of formal frameworks have been proposed to model credence updates in context-sensitive propositions. Some retain Conditionalization, some make use of hypothetical priors, but in every case new, additional norms are required to capture the full range of phenomena. There isn't space to survey the various approaches here.[50] But I will note that solving the problem of updating self-locating beliefs may have important consequences beyond fun philosophical thought-experiments like the Sleeping Beauty Problem (Elga, 2000). For instance, fine-tuning arguments for the existence of the multiverse, and debates about the proper interpretation of quantum mechanics, may both turn on how agents should manage credences in context-sensitive propositions.[51]

5 OTHER CONFIDENCE FORMALISMS

In closing, I should note that there are a number of alternative formalisms for modeling agents' varying levels of confidence in claims. First, we can think simply about whether an agent is more confident in one proposition than another. Composing these comparisons together yields a confidence ordering that may float free of any numerical assignments (see Konek, this volume). A second approach, called "ranking theory" (Spohn, 2012; Huber, this volume), attaches numbers to the confidence ranking but works only with the structure of non-negative integers. Third, we can employ a formal structure even richer than the reals. For instance, instead of representing an agent's levels of confidence at a given time with a single probability distribution, we may represent them with a *set* of such distributions (Mahtani, this volume). Or we may have one real-valued function to track the agent's attitudes and a separate (though related) one to track her evidence. This yields a fourth approach, commonly called "Dempster-Shafer Theory" (Dempster, 1966; Shafer, 1976).

Each of these approaches may be supported by some of the argument-types described above, and each is plagued by some of the problems above as well. Some allow formal structures more flexible and expressive than

---

49 In philosophy of science applications, for instance, scientific hypotheses about the physical laws of the universe or the evolutionary origins of hominids do not typically change their truth-values over time.

50 Titelbaum (2016) provides a big-picture summary with copious references.

51 For these applications and others, see Titelbaum (2013b).

Bayesianism, while some trade expressive power for added psychological plausibility. I will not attempt to choose a favorite here. But it's worth noting that among all the formalisms for representing disparate confidence levels, none is currently more studied or more often applied than the real-valued credal approach.[52]

6 APPENDIX

Here's a proof sketch for the representation theorem mentioned in Section 3.1. We will assume that in the decision theory of interest, the following hold.

- ○ Structural axioms ensuring that betting acts with various structures (as described in the proof below) are always available to the agent.

- ○ Weak dominance principle: when acts are independent of states, if there is no state in which act A yields a greater utility than act B, then A is not preferred to B.

- ○ Strong dominance principle: when acts are independent of states, if act A yields a greater utility than act B in every state, then A is preferred to B.

- ○ For any acts A and B, the agent prefers A to B just in case $EU(A) > EU(B)$, where EU is calculated as described in the main text.

The dominance principles above employ a notion of act/state independence, and the relevant notion will vary depending on which decision theory (evidential, causal, etc.) is in play. So fleshing out the proof below for a given decision theory will require showing that the acts and states appearing in each step of the proof are independent in the relevant sense. Given the types of acts involved, that should be fairly straightforward.

Notice that the following is a corollary of the weak dominance principle.

- ○ Equivalence principle: when acts are independent of states, if two acts yield the same utility as each other in every possible state, the agent is indifferent between them.

The argument is simply that if A and B yield the same utility in every possible state, then by weak dominance A is not preferred to B and B is not preferred to A. So the agent is indifferent between them, and $EU(A) = EU(B)$.

To show that any credence function c appearing in a decision theory with the features above must be a positive scalar transform of a probability function, we need to prove that it satisfies four conditions.

1. Every tautology in $\mathcal{L}$ receives the same c-value.

   Proof. Suppose for reductio we have two tautologies $T_1, T_2 \in \mathcal{L}$ such that the agent assigns a credence of x to the first and a *different* credence y to the second. Consider an act that pays 1 util on T_1 and 0 utils otherwise, and an act that pays 1 util on T_2 and 0 utils otherwise. The agent will assign the first act an expected utility of x, the second act an expected utility of y. Since x and y are different, the agent will prefer one act to the other. Yet the two acts each yield the same payout (1 util) in every possible state, so we've violated the equivalence principle (and therefore weak dominance).

2. For any tautology and contradiction $T, F \in \mathcal{L}$, $c(T) > c(F)$.

   Proof. Suppose for reductio we have a T and F such that $c(T) \leq c(F)$. Now consider an act that pays 1 util on T and 0 utils otherwise, and another act that pays 1 util on F and 0 utils otherwise. Given the supposition, the agent will assign the first act an expected utility no greater than the second. Yet the first act yields a greater utility than the second in every possible state, so by strong dominance the first act must receive a higher expected utility.

3. For any mutually exclusive $X, Y \in \mathcal{L}$, $c(X \vee Y) = c(X) + c(Y)$.

   Proof. First consider the act of purchasing a bet that pays 1 util on X, 1 util on Y, and 0 utils otherwise. Since X and Y are mutually exclusive, we may partition the possible states into X, Y, and $\sim X \& \sim Y$. Using this partition, the expected utility of this act is

$$
\begin{aligned}
& c(X) \cdot u(X) + c(Y) \cdot u(Y) + c(\sim X \& \sim Y) \cdot u(\sim X \& \sim Y) \\
= & \ c(X) \cdot 1 + c(Y) \cdot 1 + c(\sim X \& \sim Y) \cdot 0 \\
= & \ c(X) + c(Y).
\end{aligned}
\tag{27}
$$

   Now consider the act of purchasing a bet that pays 1 util on $X \vee Y$, and 0 utils otherwise. Partitioning the states into $X \vee Y$ and $\sim(X \vee Y)$, the expected utility of this act is

$$
\begin{aligned}
& c(X \vee Y) \cdot u(X \vee Y) + c(\sim[X \vee Y]) \cdot u(\sim[X \vee Y]) \\
= & \ c(X \vee Y) \cdot 1 + c(\sim[X \vee Y]) \cdot 0 \\
= & \ c(X \vee Y).
\end{aligned}
\tag{28}
$$

   These two acts have the same payout in every possible state, so to satisfy the equivalence principle the agent must be indifferent

between them. This means that their expected utilities are equal, so $c(X \vee Y) = c(X) + c(Y)$.

4. For any $X \in \mathcal{L}$, $c(X) \geq 0$.

Proof. First, we show that there can be no $Y, \top \in \mathcal{L}$ such that $\top$ is a tautology and $c(Y) > c(\top)$. Suppose for reductio that we had such two such propositions Y and $\top$. Now consider an act that pays 1 util if Y is true, and 0 utils otherwise, and an act that pays 1 util if $\top$ is true, and 0 utils otherwise. The first act has expected utility $c(Y)$, while the second has expected utility $c(\top)$. By our supposition, the agent prefers the first act. But since $\top$ is true in every state, there is no state in which the first act yields a greater utility than the second. So we have violated weak dominance.

Now to the main result. Assume for reductio that there exists an $X \in \mathcal{L}$ such that $c(X) < 0$. Since X and $\sim X$ are mutually exclusive, $c(X \vee \sim X) = c(X) + c(\sim X)$ by (3) above. If $c(X) < 0$, then $c(X \vee \sim X) < c(\sim X)$. But $X \vee \sim X$ is a tautology, so this is impossible.

REFERENCES

Adams, E. (1962). On rational betting systems. *Archiv für mathematische Logik und Grundlagenforschung, 6,* 7–29.

Allais, M. (1953). Le Comportement de l'homme rationnel devant le risque: Critique des postulates et axiomes de l'ecole Américaine. *Econometrica, 21,* 503–46.

Armendt, B. (1980). Is there a Dutch Book argument for probability kinematics? *Philosophy of Science, 47,* 583–588.

Armendt, B. (1992). Dutch strategies for diachronic rules: When believers see the sure loss coming. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, 1,* 217–229.

Arntzenius, F. (2003). Some problems for conditionalization and reflection. *The Journal of Philosophy, 100,* 356–370.

Berker, S. (2013). Epistemic teleology and the separateness of propositions. *Philosophical Review, 122,* 337–93.

Bernoulli, J. (1713). *Ars conjectandi.* Basiliae.

Bjerring, J. C. & Skipper, M. (manuscript). Bayesianism for average Joe.

Boole, G. (1952). On the application of the theory of probabilities to the question of the combination of testimonies or judgments. *Studies in Logic and Probability,* 308–85.

Bovens, L. & Hartmann, S. (2003). *Bayesian epistemology.* Oxford: Oxford University Press.

Bradley, R. (2007). Reaching a consensus. *Social Choice and Welfare, 29,* 609–32.

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review*, *78*, 1–3.

Briggs, R. A. (2009). Distorted reflection. *The Philosophical Review*, *118*, 59–85.

Briggs, R. A. (2019). Conditionals. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Briggs, R. A. & Pettigrew, R. (forthcoming). An accuracy-dominance argument for Conditionalization. *Noûs*.

Buchak, L. (2013). *Risk and rationality*. Oxford: Oxford University Press.

Carnap, R. (1950). *Logical foundations of probability*. Chicago: University of Chicago Press.

Carnap, R. (1962). *Logical foundations of probability* (2nd). Chicago: University of Chicago Press.

Carr, J. R. (2015). Don't stop believing. *Canadian Journal of Philosophy*, *45*.

Carr, J. R. (2017). Epistemic utility theory and the aim of belief. *Philosophy and Phenomenological Research*, *95*, 511–34.

Chihara, C. (1981). Quine and the confirmational paradoxes. In P. French, H. Wettstein, & T. Uehling (Eds.), *Midwest studies in philosophy 6: Foundations of analytic philosophy* (pp. 425–52). University of Minnesota Press.

Christensen, D. (1999). Measuring confirmation. *The Journal of Philosophy*, *96*, 437–61.

Christensen, D. (2004). *Putting logic in its place*. Oxford: Oxford University Press.

Christensen, D. (2009). Disagreement as evidence: The epistemology of controversy. *Philosophy Compass*, *4*, 756–67.

Christensen, D. & Lackey, J. (Eds.). (2013). *The epistemology of disagreement: New essays*. Oxford: Oxford University Press.

Crupi, V. (2016). Confirmation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2016). Metaphysics Research Lab, Stanford University.

Davies, M. (1998). Exernalism, architecturalism, and epistemic warrant. In C. Wright, B. Smith, & C. Macdonald (Eds.), *Knowing our own minds* (pp. 321–61). Oxford: Oxford University Press.

de Finetti, B. (1937/1964). Foresight: Its logical laws, its subjective sources. In H. E. Kyburg Jr & H. Smokler (Eds.), *Studies in subjective probability* (pp. 94–158). Originally published as "La prévision; ses lois logiques, ses sources subjectives" in *Annales de l'Institut Henri Poincaré*, Volume 7, 1–68. New York: Wiley.

de Finetti, B. (1974). *Theory of probability*. New York: Wiley.

Dempster, A. P. (1966). New methods for reasoning towards posterior distributions based on sample data. *Annals of Mathematical Statistics*, *37*, 355–74.

Douven, I. (2012). The Lottery Paradox and the pragmatics of belief. *Dialectica*, *66*, 351–73.

Dretske, F. I. (1970). Epistemic operators. *The Journal of Philosophy*, *67*, 1007–1023.

Earman, J. (1992). *Bayes or bust? A critical examination of Bayesian confirmation theory*. Cambridge, MA: The MIT Press.

Easwaran, K. (2013). Expected accuracy supports conditionalization—and conglomerability and reflection. *Philosophy of Science*, *80*, 119–142.

Easwaran, K. (2014). Regularity and hyperreal credences. *Philosophical Review*, *123*, 1–41.

Easwaran, K. (2019). Conditional probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Easwaran, K., Fenton-Glynn, L., Hitchcock, C., & Velasco, J. D. (2016). Updating on the credences of others: Disagreement, agreement, and synergy. *Philosophers' Imprint*, *16*, 1–39.

Eells, E. (1982). *Rational decision and causality*. Cambridge Studies in Philosophy. Cambridge: Cambridge University Press.

Eells, E. (1985). Problems of old evidence. *Pacific Philosophical Quarterly*, *66*, 283–302.

Elga, A. (2000). Self-locating belief and the Sleeping Beauty problem. *Analysis*, *60*, 143–7.

Elga, A. (2007). Reflection and disagreement. *Noûs*, *41*, 478–502.

Feldman, R. (2007). Reasonable religious disagreements. In L. M. Antony (Ed.), *Philosophers without gods: Meditations on atheism and the secular life*. Oxford: Oxford University Press.

Fitelson, B. (2008). A decision procedure for probability calculus with applications. *The Review of Symbolic Logic*, *1*, 111–125.

Fitelson, B. (2015). The strongest possible Lewisian triviality result. *Thought*, *4*, 69–74.

Fitelson, B. & Hawthorne, J. (2010). How Bayesian confirmation theory handles the Paradox of the Ravens. *Boston Studies in the Philosophy of Science*, *284*.

Foley, R. (1993). *Working without a net*. Oxford: Oxford University Press.

Garber, D. (1983). Old evidence and logical omniscience in Bayesian confirmation theory. In J. Earman (Ed.), *Testing scientific theories* (Vol. 10, pp. 99–132). Minnesota Studies in the Philosophy of Science. Minneapolis: University of Minnesota Press.

Gibbard, A. & Harper, W. (1978). Counterfactuals and two kinds of expected utility. In C. A. Hooker, J. L. Leach, & E. F. McClennan (Eds.), *Foundations and applications of decision theory* (Vol. 13a, pp. 125–62). University of Western Ontario Series in Philosophy of Science. Dordrecht: D. Reidel Publishing Company.

Glymour, C. (1980). *Theory and evidence*. Princeton, NJ: Princeton University Press.

Goodman, N. (1955). *Fact, fiction, and forecast*. Cambridge, MA: Harvard University Press.

Greaves, H. (2013). Epistemic decision theory. *Mind*, *122*, 915–52.

Greaves, H. & Wallace, D. (2006). Justifying conditionalization: Conditionalization maximizes expected epistemic utility. *Mind*, *115*, 607–632.

Hacking, I. (1967). Slightly more realistic personal probability. *Philosophy of Science*, *34*, 311–325.

Hájek, A. (2003). What conditional probability could not be. *Synthese*, *137*, 273–323.

Hájek, A. (2007). The reference class problem is your problem too. *Synthese*, *156*, 563–85.

Hájek, A. (2009). Arguments for—or against—probabilism? In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (Vol. 342, pp. 229–251). Synthese Library. Springer.

Hájek, A. (2011). Triviality pursuit. *Topoi*, *30*, 3–15.

Harman, G. (1986). *Change in view*. Boston: The MIT Press.

Hawthorne, J. & Fitelson, B. (2004). Re-solving irrelevant conjunction with probabilistic independence. *Philosophy of Science*, *71*, 505–514.

Hempel, C. G. (1945). Studies in the logic of confirmation (I). *Mind*, *54*, 1–26.

Henderson, L. (2013). Bayesianism and inference to the best explanation. *British Journal for the Philosophy of Science*, *65*, 687–715.

Hintikka, J. (1975). Impossible possible worlds vindicated. *Journal of Philosophical Logic*, *4*, 475–84.

Hitchcock, C. R. (2012). Probabilistic causation. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2012).

Holton, R. (2014). Intention as a model for belief. In M. Vargas & G. Yaffe (Eds.), *Rational and social agency: The philosophy of Michael Bratman* (pp. 12–37). Oxford: Oxford University Press.

Hosiasson-Lindenbaum, J. (1940). On confirmation. *The Journal of Symbolic Logic*, *5*, 133–48.

Howson, C. (1992). Dutch Book arguments and consistency. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association*, *2*, 161–8.

Howson, C. & Urbach, P. (2006). *Scientific reasoning: The Bayesian approach* (3rd). Chicago: Open Court.

Huber, F. (2019). Ranking theory. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Huemer, M. (2011). Does probability theory refute coherentism? *Journal of Philosophy*, *108*, 463–72.

Jaynes, E. T. (1957a). Information theory and statistical mechanics i. *Physical Review*, *106*, 620–30.

Jaynes, E. T. (1957b). Information theory and statistical mechanics ii. *Physical Review*, *108*, 171–90.

Jeffrey, R. C. (1965). *The logic of decision* (1st). McGraw-Hill series in probability and statistics. New York: McGraw-Hill.

Joyce, J. M. (1998). A nonpragmatic vindication of probabilism. *Philosophy of Science*, *65*, 575–603.

Joyce, J. M. (1999). *The foundations of causal decision theory*. Cambridge: Cambridge University Press.

Joyce, J. M. (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (Vol. 342, pp. 263–297). Synthese Library. Springer.

Kemeny, J. G. (1955). Fair bets and inductive probabilities. *The Journal of Symbolic Logic*, *20*, 263–273.

Kolmogorov, A. N. (1933/1950). *Foundations of the theory of probability*. Translation edited by Nathan Morrison. New York: Chelsea Publishing Company.

Konek, J. (2019). Comparative probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Kyburg, H. E., Jr. (1961). *Probability and the logic of rational belief*. Middletown: Wesleyan University Press.

Laplace, P.-S. (1814/1995). *Philosophical essay on probabilities*. Translated from the French by Andrew Dale. New York: Springer.

Leitgeb, H. (2017). *The stability of belief: How rational belief coheres with probability*. Oxford: Oxford University Press.

Levi, I. (1987). The demons of decision. *The Monist*, *70*, 193–211.

Lewis, D. (1971). Immodest inductive methods. *Philosophy of Science*, *38*, 54–63.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review*, *85*, 297–315.

Lewis, D. (1979). Atittudes *de dicto* and *de se*. *The Philosophical Review*, *88*, 513–543.

Lewis, D. (1980). A subjectivist's guide to objective chance. In R. C. Jeffrey (Ed.), *Studies in inductive logic and probability* (Vol. 2, pp. 263–294). Berkeley: University of California Press.

Lewis, D. (1981). Causal decision theory. *Australasian Journal of Philosophy*, *59*, 5–30.

Lin, H. & Kelly, K. T. (2012). A geo-logical solution to the Lottery Paradox. *Synthese*, *186*, 531–75.

Lipton, P. (2004). *Inference to the best explanation* (2nd). London: Routledge.

Littlejohn, C. (2018). Stop making sense? On a puzzle about rationality. *Philosophy and Phenomenological Research*, *96*(2), 257–272.

Locke, J. (1689/1975). *An essay concerning human understanding* (P. H. Nidditch, Ed.). Oxford: Oxford University Press.

Maher, P. (1996). Subjective and objective confirmation. *Philosophy of Science*, *63*, 149–74.

Maher, P. (2002). Joyce's argument for probabilism. *Philosophy of Science*, *96*, 73–81.

Mahtani, A. (2019). Imprecise probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Makinson, D. C. (1965). The paradox of the preface. *Analysis*, *25*, 205–7.

Meacham, C. J. (2010). Two mistakes regarding the Principal Principle. *British Journal for the Philosophy of Science*, *61*, 407–31.

Meacham, C. J. (2016). Ur-priors, conditionalization, and ur-prior conditionalization. *Ergo*, *3*(17), 444–492.

Meacham, C. J. & Weisberg, J. (2011). Representation theorems and the foundations of decision theory. *Australasian Journal of Philosophy*, *89*, 641–663.

Moss, S. (2018). *Probabilistic knowledge*. Oxford: Oxford University Press.

Okasha, S. (2000). Van Fraassen's critique of inference to the best explanation. *Studies in History and Philosophy of Science*, 691–710.

Olsson, E. (2017). Coherentist theories of epistemic justification. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2017). Metaphysics Research Lab, Stanford University.

Pettigrew, R. (2013). A new epistemic utility argument for the Principal Principle. *Episteme*, *10*, 19–35.

Pettigrew, R. (2014). Accuracy, risk, and the Principle of Indifference. *Philosophy and Phenomenological Research*, *90*.

Pettigrew, R. (2016). *Accuracy and the laws of credence*. Oxford: Oxford University Press.

Predd, J., Seiringer, R., Lieb, E., Osherson, D., Poor, V., & Kulkarni, S. (2009). Probabilistic coherence and proper scoring rules. *IEEE Transactions on Information Theory*, *55*, 4786–4792.

Ramsey, F. P. (1931). Truth and probability. In R. B. Braithwaite (Ed.), *The foundations of mathematics and other logic essays* (pp. 156–198). New York: Harcourt, Brace and Company.

Rasmussen, M. & Steglich-Petersen, A. (Eds.). (forthcoming). *Higher-order evidence: New essays*. Oxford: Oxford University Press.

Reichenbach, H. (1956). The principle of common cause. In *The direction of time* (pp. 157–160). University of California Press.

Rosenkrantz, R. (1977). *The paradoxes of confirmation*. Synthese Library. Dordrecht: D. Reidel Publishing Company.

Rosenkrantz, R. (1981). *Foundations and applications of inductive probability*. Atascadero, CA: Ridgeview Press.

Russell, J., Hawthorne, J., & Buchak, L. (2015). Groupthink. *Philosophical Studies*, *172*, 1287–1309.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Savage, L. J. (1967). Difficulties in the theory of personal probability. *Philosophy of Science*, *34*, 305–310.

Schoenfield, M. (2014). Permission to believe: Why permissivism is true and what it tells us about irrelevant influences on belief. *Noûs*, *48*, 193–218.

Seidenfeld, T. (1986). Entropy and uncertainty. *Philosophy of Science*, *53*, 467–491.

Seidenfeld, T., Schervish, M. J., & Kadane, J. (manuscript). Non-conglomerability for countably additive measures that are not κ-additive.

Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.

Shimony, A. (1955). Coherence and the axioms of confirmation. *Journal of Symbolic Logic*, *20*, 1–28.

Shimony, A. (1988). An Adamite derivation of the calculus of probability. In J. Fetzer (Ed.), *Probability and causality* (pp. 151–161). Dordrecht: Reidel.

Shogenji, T. (1999). Is coherence truth-conducive? *Analysis*, *59*, 338–45.

Skyrms, B. (1987a). Coherence. In N. Rescher (Ed.), *Scientific inquiry in philosophical perspective* (pp. 225–42). Pittsburgh: University of Pittsburgh Press.

Skyrms, B. (1987b). Dynamic coherence and probability kinematics. *Philosophy of Science*, *54*, 1–20.

Smithies, D. (2015). Ideal rationality and logical omniscience. *Synthese*, *192*, 2769–93.

Spohn, W. (2012). *The laws of belief: Ranking Theory & its philosophical applications*. Oxford: Oxford University Press.

Suppes, P. (1974). *Probabilistic metaphysics*. Uppsala: University of Uppsala Press.

Swinburne, R. (1971). The paradoxes of confirmation: A survey. *American Philosophical Quarterly*, *8*, 318–30.

Talbott, W. J. (1991). Two principles of Bayesian epistemology. *Philosophical Studies*, *62*, 135–150.

Teller, P. (1973). Conditionalization and observation. *Synthese*, *26*, 218–258.

Thoma, J. (2019). Decision theory. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

Titelbaum, M. G. (forthcoming). *Fundamentals of Bayesian epistemology*. Oxford: Oxford University Press.

Titelbaum, M. G. (2013a). *Quitting certainties: A Bayesian framework modeling degrees of belief*. Oxford: Oxford University Press.

Titelbaum, M. G. (2013b). Ten reasons to care about the Sleeping Beauty Problem. *Philosophy Compass*, *8*, 1003–17.

Titelbaum, M. G. (2015). Rationality's fixed point (or: In defense of right reason). In T. S. Gendler & J. Hawthorne (Eds.), *Oxford studies in epistemology* (Vol. 5, pp. 253–94). Oxford University Press.

Titelbaum, M. G. (2016). Self-locating credences. In A. Hájek & C. R. Hitchcock (Eds.), *The Oxford handbook of probability and philosophy* (pp. 666–680). Oxford: Oxford University Press.

van Fraassen, B. C. (1983). Calibration: A frequency justification for personal probability. In R. Cohen & L. Laudan (Eds.), *Physics philosophy and psychoanalysis* (pp. 295–319). Dordrecht: Reidel.

van Fraassen, B. C. (1984). Belief and the will. *The Journal of Philosophy*, *81*, 235–256.

van Fraassen, B. C. (1989). *Laws and symmetry*. Oxford: Clarendon Press.

von Mises, R. (1928/1957). *Probability, statistics and truth*. (English edition of the original German *Wahrscheinlichkeit, Statistik und Wahrheit*.) New York: Dover.

von Neumann, J. & Morgenstern, O. (1947). *Theory of games and economic behavior* (2nd). Princeton, NJ: Princeton University Press.

Weirich, P. (2012). Causal decision theory. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2012).

Weisberg, J. (2007). Conditionalization, reflection, and self-knowledge. *Philosophical Studies*, *135*, 179–197.

Weisberg, J. (2009). Locating IBE in the Bayesian framework. *Synthese*, *167*, 125–44.

Wenmackers, S. (2019). Infinitesimal probabilities. In R. Pettigrew & J. Weisberg (Eds.), *The open handbook of formal epistemology*. PhilPapers.

White, R. (2005). Epistemic permissiveness. *Philosophical Perspectives*, *19*, 445–459.

Williamson, T. (2007). How probable is an infinite sequence of heads? *Analysis*, *67*, 173–80.

Wright, C. (2003). Some reflections on the acquisition of warrant by inference. In S. Nuccetelli (Ed.), *New essays on semantic externalism*. Cambridge, MA: MIT Press.

Zynda, L. (1995). Old evidence and new theories. *Philosophical Studies*, *77*, 67–95.

Zynda, L. (2000). Representation theorems and realism about degrees of belief. *Philosophy of Science*, *67*, 45–69.