

Accuracy, Chance, and the Principal Principle

Richard Pettigrew

University of Bristol

How should my degrees of credence in propositions about objective chances interact with my degrees of credence in other propositions? David Lewis (1980) formulated a precise answer. He called it the “Principal Principle.” And he took it to specify the role that a feature of the world must play if it is to count as chance. However, he also argued that the chances delivered by his best-system analysis of laws and chances could not possibly play the roles required of them by the Principal Principle: the Principal Principle is inconsistent with the best-system analysis. Later, Lewis (1994) came to accept a slightly different precise answer to the question. This is Michael Thau’s and Ned Hall’s amendment to the Principal Principle, which is known as the New Principle (Hall 1994; Thau 1994; Joyce 2007). The best-system analysis is consistent with the New Principle. Later still, Jenann Ismael (2008) proposed an alternative amendment to the Principal Principle, which she called the “Generalized Principal Principle.” Ismael’s principle is also consistent with the best-system analysis.

The large literature that has grown up around these proposals includes a number of different and illuminating arguments in their favor, as well as subtle improvements in their formulation. However, no attempt has been made to give them a formal justification. That is, there has been no attempt to deduce any chance-credence norm from more

I would like to thank Rachael Briggs, Branden Fitelson, Ben Levinstein, Samir Okasha, Barry Loewer, Michael Strevens, Jim Pryor, Chris Hitchcock, and the referees and editors for this journal for very helpful comments on earlier versions of this article. This work for this article was carried out during my time as a British Academy Postdoctoral Fellow.

fundamental epistemic principles that have been formulated with mathematical precision. This is notable since such justifications abound in the case of probabilism, the fundamental norm of degrees of credence: in that case, we have the Dutch Book Theorem, the Representation Theorems, the Calibration Theorems, and Cox's Theorem, to name the most prominent. In this article, I aim to fill this gap and give a formal justification of the chance-credence norms by adapting an argument for probabilism developed by Jim Joyce (1998). In Joyce's argument, the fundamental epistemic principles from which probabilism is deduced concern how we should measure the inaccuracy of a set of credences at a world, that is, how far those credences lie from the *correct* or *vindicated* credences. Probabilism is then deduced from the characterization of these inaccuracy measures together with the claim (roughly) that it is irrational to have a set of credences if there is another set that is more accurate however the world turns out. As we will see, using the adaptation of Joyce's argument that I will describe here, if we begin from different accounts of what chances are, the argument can be used to justify the Principal Principle, the New Principle, or the Generalized Principal Principle. I do not take a stand on which account of chance is more plausible in this article. My concern is with the norms that follow from different accounts. Thus, what is offered here is not a full justification of any single chance-credence norm. Rather, a strategy is offered that allows us to justify a particular chance-credence norm once the metaphysics of chance has been settled.

It is notable that the strategy offered here for justifying chance-credence norms is available equally to the nonreductionist and the reductionist about chance.¹ After all, Lewis maintained that only the reductionist could possibly establish chance-credence norms. He maintained that if a feature of the world could not be reduced to some nonmodal feature, then knowledge of that feature could not constrain our credences in the way demanded by the chance-credence norms (Lewis 1994, 484).² One of the consequences of the strategy for justifying chance-credence norms described here is that this is mistaken.

1. As I will use the terminology here, a *reductionist* holds that objective chances are reducible to a nonmodal feature of the world; a *nonreductionist* holds that they are not.

2. Lewis's argument is reminiscent of Mackie's famous argument that if values were to exist as an objective feature of the world, they could not be such that knowledge of them would motivate action in the way that knowledge of values does (Mackie 1977, 29ff.).

Another attraction of the present approach is that it reveals exactly which account of chance is required to justify which chance-credence norm. This is one of the strengths of any formal justification of an epistemic norm: it reveals the precise underlying assumptions required by different but closely related versions of a given norm, and thus it allows us to adjudicate between them.

In section 1, I describe the Principal Principle and its variants. I then describe Joyce's justification of the norm of probabilism (sections 3 and 4) and show how it may be adapted to justify the chance-credence norms discussed in the first section (section 5). In section 6, I consider objections to this argument, and in section 7, I conclude by considering one way in which the argument might be extended to justify other closely related norms. The appendix gives the proofs of the key theorems.

1. Chance-Credence Norms

Suppose I am at the beginning of my epistemic life. That is, I have accumulated no evidence. In this situation, how strongly should I believe that a given future coin toss will come up heads conditional on the proposition that the objective chance that it will come up heads is 0.7? Intuitively: I should believe it to degree 0.7. More generally, suppose C_{ch} is the proposition that the objective chances are given by the particular function ch —where ' ch ' rigidly designates a particular function. Then it seems plausible that my initial degree of credence in a given proposition A conditional on C_{ch} should be $ch(A)$. This gives us

Miller's Principle (MP): An agent ought to have an initial credence function b such that, for all possible initial chance functions ch and all propositions A ,

$$b(A | C_{ch}) = ch(A)$$

providing $b(C_{ch}) > 0$.

Following Lewis (1980, 267), and heeding the recent salutary reminder from Chris Meacham (2010), I will talk throughout of norms that govern an agent's *initial* or *ur*-credence function: that is, the agent's credence function at the beginning of his or her epistemic life, prior to accumulating any evidence; the function upon which he or she conditionalizes as he or she collects evidence. Moreover, I will state those norms in terms of so-called *initial* or *ur*-chance functions (see, for example, Hall 2004, 95): that is, the chance function at the beginning of a world's history; the

function upon which we conditionalize with entire histories up to time t in order to give the chance function at time t . As Meacham has decisively shown, only by doing this do we avoid inconsistency. I will drop the ‘initial’ and ‘ur-’ qualifications henceforth, except in the official statement of the norms, where it is worth reminding ourselves of precisely what they say. Interestingly, returning to Lewis’s original focus on initial credence functions also saves my strategy for justifying the chance-credence norms from a recent objection raised against Joyce’s justification of probabilism by Branden Fitelson and Kenny Easwaran. But more of this below (section 6).

(MP) captures an important intuition about the connection between initial chances and initial credences: we ought to defer to those chances in setting those credences. But it is not clear that it covers all situations in which we think such deference is appropriate. For instance, my degree of credence that the coin will land heads conditional on the proposition that its chance of landing heads is 0.7 *and* conditional on the proposition that it rained yesterday in Edinburgh should also be 0.7. That is, the evidence concerning the chance of heads “screens off” or “trumps” the evidence concerning yesterday’s rain in Edinburgh—that is, it renders it irrelevant. But it is clear that not all evidence has this feature. For instance, if E entails A , then the chances cannot “screen off” E : in this case, my degree of credence in A conditional on C_{ch} *and* conditional on E ought to be 1, not $ch(A)$. Lewis intended his Principal Principle to generalize (MP) by identifying all cases in which an agent ought to defer to the objective chances. To do this, he needed to identify which pieces of evidence are “screened off” or “trumped” by chance and which are not. Lewis’s answer is that it is the *admissible* evidence that is “screened off” by chance, where evidence E is admissible for a proposition A if it gives no information about the truth of A that does not go through the chance of A . Thus, A is clearly inadmissible for A , as is any E that entails A ; and a tautologous proposition is clearly admissible for any proposition. But there are other, stronger sorts of admissible evidence as well. For instance, Lewis (1980, 275) held that all propositions concerning the past are admissible for propositions concerning the present or future, as are so-called history-to-chance conditionals. Using this notion, he formulated his chance-credence norm:

Accuracy, Chance, and the Principal Principle

Principal Principle (PP): An agent ought to have an initial credence function b such that, for all possible initial chance functions ch and all propositions A and E , if E is admissible for A ,

$$b(A | C_{ch} \wedge E) = ch(A)$$

providing $b(C_{ch} \wedge E) > 0$.

Lewis never gave a mathematically precise account of the admissibility condition. But in this article, I seek a mathematically precise justification of chance-credence norms, so I require equally precise formulations of those norms. Fortunately, it turns out that Miller’s Principle already entails the requisite formulation (see, for example, Bogdan 1984):³

Levi’s Principal Principle (LPP): An agent ought to have a credence function b such that, for all possible chance functions ch and all propositions A and E , if E and A are stochastically independent according to ch (that is, $ch(A | E) = ch(A)$),

$$b(A | C_{ch} \wedge E) = ch(A)$$

providing $b(C_{ch} \wedge E) > 0$.⁴

(MP) and (LPP) both run into problems in the presence of two further assumptions. The first assumption is probabilism (Prob), the norm that says that an agent’s credence function ought to be a probability function. The second assumption involves an objective chance function ch that is *self-undermining* in the following sense: ch is not certain that it gives the

3. *Proof.* Suppose b satisfies (MP). Then

$$b(A | C_{ch} \wedge E) = \frac{b(A \wedge C_{ch} \wedge E)}{b(C_{ch} \wedge E)} = \frac{b(A \wedge E | C_{ch})b(C_{ch})}{b(E | C_{ch})b(C_{ch})} = \frac{b(A \wedge E | C_{ch})}{b(E | C_{ch})} = \frac{ch(A \wedge E)}{ch(E)} = ch(A | E).$$

4. Those familiar with Lewis 1980 may wonder how (MP), (PP), and (LPP) relate to Lewis’s second formulation of the Principal Principle in that paper: $b(A | T_w \wedge H_{tw}) = ch_{tw}(A)$, for any world w and time t , where T_w is the conjunction of history-to-chance conditionals that hold at w , H_{tw} is the history of w up to t , and ch_{tw} is the objective chance function of w at t . The answer is that, on one natural way of translating this formulation into the framework we adopt here, (MP) entails this chance-credence norm. In our framework, the conjunction of history-to-chance conditionals that hold at a world w —or what Lewis calls the “complete theory of chance” for w —is expressed by the proposition C_{ch_w} , where ch_w is the ur-chance function at w . C_{ch_w} says that the ur-chances are given by ch_w . Thus, it entails that the chances at time t are given by $ch_{tw}(\cdot) := ch_w(\cdot | H_{tw})$, if H_{tw} is the history up to t . Thus, Lewis’s reformulation of the Principal Principle becomes: $b(A | C_{ch_w} \wedge H_{tw}) = ch_{tw}(A) = ch_w(A | H_{tw})$, which follows from (MP) by the proof in note 3.

objective chances; in symbols, $ch(C_{ch}) < 1$. Now, suppose that b is the credence function of a rational agent and suppose that this agent has a nonzero credence in the proposition C_{ch} that the objective chances are given by ch , where ch is self-undermining: that is, $b(C_{ch}) > 0$ and $ch(C_{ch}) < 1$. Then, on the one hand, (MP) demands that $b(C_{ch} | C_{ch}) = ch(C_{ch}) < 1$, since ch is self-undermining. And on the other hand, (Prob) demands that $b(C_{ch} | C_{ch}) = 1$, since (Prob) demands $b(A | A) = 1$ for every A such that $b(A) > 0$. That is, (Prob) and (MP) make incompatible demands on any agent who does not completely rule out the possibility of self-undermining chance functions.

Of course, one might be tempted to respond to this problem by claiming that self-undermining chance functions are impossible in some sense, and thus that it is a rational constraint on an agent that he or she assign zero degree of credence to C_{ch} whenever ch is self-undermining. If we do this, then the conflict between (Prob) and (MP) is removed, since an agent on whom they make incompatible demands is already ruled out as irrational by this further rational constraint. Unfortunately, this response is available only on certain philosophical accounts of objective chance.

Take, for instance, an actual frequentist account of chance. On this account, roughly, the chance of a given event is the actual relative frequency with which events like that event occur. Now consider a world that contains nothing but a coin that is flipped exactly four times and that lands heads twice and tails twice. Then the actual frequentist must say that the chance of heads on any given toss is 0.5 since the relative frequency of heads among all tosses is 0.5. Then, if the chance function considers the four different coin tosses to be independent, then it will assign a chance of $0.5^4 = 0.0625$ to the proposition that the coin will land heads on every toss. But, for the actual frequentist, the proposition that the coin will land heads on every toss is equivalent to the proposition that the chance of heads on any given toss is 1, since the relative frequency of heads among all tosses is 1 just in case each toss lands heads. Thus, although the chance function gives 0.5 as the chance of heads, it gives a nonzero chance—namely, 0.0625—to the proposition that the chance of heads is 1. Thus, if ch is the actual frequentist's chance function, we have $ch(C_{ch}) < 1$, since ch must assign a nonzero probability to the chances being different from what ch says they are. Thus, actual frequentism is committed to the existence of self-undermining chances.

Similar considerations show that the chances given by Lewis's best-system analysis will also typically be self-undermining. On Lewis's ac-

count, as on the frequentist account, many worlds will have chance functions that assign nonzero chances to different ways the world might turn out such that the chances given by the best-system analysis would be different from the way they actually are if the world were to turn out that way. Indeed, for most worlds that contain chancy events, frequentism and the best-system analysis will give a nonzero chance to the proposition that the world is deterministic; that is, that all chances are 0 or 1.

This is not to say that all reductionist accounts of chance give rise to self-undermining chances. Jonathan Schaffer (2003) gives a recipe for transforming chances that undermine themselves into chances that do not: if ch is a self-undermining chance function, then $ch(\cdot | C_{ch})$ is not. And, if we have an account of chance on which chances supervene on nonmodal features of the world, as in Lewis's best-system analysis, then the chance functions obtained by applying Schaffer's recipe will also supervene on those features. Thus, there are reductionist accounts that avoid self-undermining chances.

A certain sort of nonreductionist account of chance also avoids self-undermining chances. Some nonreductionists hold that, at the beginning of the history of the world, the chances are fixed or determined; in other words, their chance is 1. Thus, if those fixed chances are given by ch , then the chance of C_{ch} is 1: determined events get chance 1. That is, $ch(C_{ch}) = 1$ and ch is not self-undermining.

Finally, just as there are reductionist accounts of chance with no self-undermining chances, so there are nonreductionist accounts that do have them. Some reductionists hold that at the beginning of the history of the world, the chances are *not* fixed or determined; in other words, the chance that they are given by some particular possible chance function is less than 1. This gives rise to self-undermining chances.

In response to the problem of self-undermining chances, a variety of philosophers have proposed chance-credence norms that are closely related to (MP) and (LPP) but that are consistent with (Prob) even in the presence of such chances. Each arises from a diagnosis of what gives rise to the inconsistency described above. The idea is that (MP) and (LPP) derive their plausibility from the more basic intuition that we ought always to defer to chances when it comes to assigning degrees of credence to propositions. The problems with these norms do not undermine this intuition. Rather, they undermine the way in which we have formalized the notion of epistemic deference that is used in stating the intuition. It cannot be that deferring to the objective chances amounts to setting our degree of credence in A conditional on C_{ch} to the objective chance of A

according to ch . For, as we have seen, in the presence of self-undermining chances, this will leave us less confident in C_{ch} conditional on C_{ch} than we can coherently be. Instead, some alternative must be sought.

Two alternative accounts of epistemic deference have been proposed: the first is due to Hall (1994) and Thau (1994), the second to Ismael (2008). According to Hall (1994) and Thau (1994), deferring to chance does not amount to setting our degree of credence in A conditional on C_{ch} to the objective chance of A according to ch . Rather, it amounts to setting our degree of credence in A conditional on C_{ch} to the objective chance of A conditional on C_{ch} according to ch . Roughly speaking, their point is this: We defer to chance not solely because chance has more information about the world than we have. (It does have more information, but unlike truth, it is not perfectly well informed.) We defer to chance also, they argue, because chance has a better ability to process whatever information we have about the world and give a probability on the basis of this. Thus, when we defer to chance, we must ensure that it has as much information as we have. We must defer to whatever the chances are, conditional on all the information we have. And of course, this information includes the fact that the chances are given by ch , when it is ch to which we defer. More precisely, Hall (1994) and Thau (1994) give the following amendment to the Principal Principle:

New Principle (NP): An agent ought to have an initial credence function b such that, for all possible initial chance functions ch and all propositions A and E ,

$$b(A | C_{ch} \wedge E) = ch(A | C_{ch} \wedge E)$$

providing $b(C_{ch} \wedge E), ch(C_{ch} \wedge E) > 0$.

Under the assumption that ch is not self-undermining, (NP) and (LPP) are equivalent. If ch is self-undermining, they are not. And, moreover, (NP) is consistent with (Prob) even for an agent for whom $b(C_{ch}) > 0$ for some such self-undermining ch .⁵

5. To see this, note that, for each assignment of probabilities $\lambda_1, \dots, \lambda_n$ to the partition $C_{ch_1}, \dots, C_{ch_n}$, there is a unique credence function b such that (i) b satisfies (NP) and (Prob) and (ii) $b(C_{ch_i}) = \lambda_i$. Given A in the algebra on which b and ch_i are defined, we calculate $b(A)$ as follows:

$$b(A) := \sum_i \lambda_i ch_i(A | C_{ch_i}).$$

Ismael offers an alternative account of deference to chance. To defer to chance, she thinks, is not to set our degree of credence in A conditional on C_{ch} to $ch(A)$, as (LPP) claims; nor is it to set our degree of credence in A conditional on C_{ch} to $ch(A | C_{ch})$, as (NP) claims. Rather, it is to set our *unconditional* degree of credence in A to $ch(A)$ if ch gives the actual objective chances. But if that is what it means to defer to chance, then the norm that says that we ought to defer to chance becomes an externalist norm: that is, it makes reference to a quantity to which we are not guaranteed internal access, namely, the actual objective chances. Thus, in order to produce from this account an internalist norm akin to the Principal Principle, Ismael instead demands that we set our degree of credence in A to what we *expect* the objective chances to be. Indeed, strangely, her proposal is perhaps best expressed by Hall’s informal (and inaccurate) statement of (NP): “Rationality requires conforming one’s subjective credence in a proposition to one’s estimate of the objective chance that the proposition will come true” (Hall 1994, 505). This gives us:

Generalized Principal Principle (GPP): An agent ought to have an initial credence function b such that, for all propositions A ,

$$b(A) = \sum_{ch} b(C_{ch})ch(A)$$

where the sum ranges over all possible initial chance functions ch .

Notice that (GPP) follows from (LPP), but fortunately (LPP) does not follow from (GPP) unless ch thinks that A and C_{ch} are independent, and this is something that will not typically hold if ch is self-undermining. Thus, we have another alternative chance-credence norm that seeks to capture what is right about (PP), while avoiding inconsistency.⁶

6. It is less straightforward to show that (GPP) is consistent with (Prob) than that (NP) is. We do so as follows: Given possible chance functions ch_1, \dots, ch_n , we define the following function F , which takes probabilistic credence functions to probabilistic credence functions:

$$F : b(\cdot) \mapsto \sum_i b(C_{ch_i})ch_i(\cdot)$$

Then it is straightforward to show that when we represent credence functions as points in Euclidean space whose dimension is the number of propositions in the algebra on which credence functions are defined, F is a continuous function. Moreover, on this representation, the set of probabilistic credence functions forms a closed, convex, and bounded region of the Euclidean space. Thus, F satisfies the hypotheses of Brouwer’s Fixed Point Theorem, which says that any continuous function from a closed, convex, bounded

Note that (NP) and (GPP) are inconsistent with one another. After all,

$$\begin{aligned} \text{(NP)} \Rightarrow b(A) &= \sum_{ch} b(C_{ch})ch(A | C_{ch}) \\ \text{(GPP)} \Rightarrow b(A) &= \sum_{ch} b(C_{ch})ch(A) \end{aligned}$$

and these will typically be different if ch is self-undermining. If ch is not undermining, (LPP), (NP), and (GPP) are all equivalent.

2. A Sketch of Joyce’s Strategy

How are we to justify the various chance-credence norms described above? In the remainder of the article, I will present a justification that is adapted from Jim Joyce’s argument for probabilism (Joyce 1998). In this section, I sketch the strategy of Joyce’s argument; in the next section, I give more details. I will present the argument as having two parts: the first involves a general strategy by which we might hope to establish various norms governing degrees of credence; the second part deploys this strategy in a particular way to justify probabilism. I will then show how we might use the general strategy that forms the first part of Joyce’s argument but deploy it in different ways in order to justify (Prob) + (PP) or (Prob) + (NP) or (Prob) + (GPP), depending on the account of chance with which we begin.

2.1. Joyce’s General Strategy

As I will present it, Joyce’s argument turns on the notion of *vindication* for credence functions. The idea is borrowed from the theory of full beliefs. Suppose we are given a set of propositions. Then, intuitively, for each possible world, there is a set of full beliefs concerning those propositions that is *vindicated* or *perfect* or *exactly correct* at that world: it is the set of full beliefs in each proposition in the set that is true at that world and full disbeliefs in each proposition in the set that is false. Similarly, as I am presenting it, Joyce’s argument for probabilism assumes that there is, for each world, a credence function that is vindicated or perfect or exactly

region of a Euclidean space to itself has a fixed point. Thus, F has a fixed point b^* : that is, $F(b^*) = b^*$. But the fixed points of F are exactly the probabilistic credence functions that satisfy (GPP). Therefore, for any possible initial chance functions ch_1, \dots, ch_n , there is at least one credence function b^* such that b^* satisfies (Prob) and (GPP).

correct at that world. As Alan Hájek (n.d.) puts it, we assume that there is something that stands to credence functions as truth stands to sets of full beliefs:

Truth: Set of full beliefs: ???: Credence function

I will say nothing now about the identity of the vindicated credence functions. It will turn out that this is the point at which my argument diverges from Joyce's: he takes one sort of credence function to be vindicated; I take another. And the result is two arguments with different conclusions: Joyce's conclusion is (Prob); mine is (Prob) + (NP) or (Prob) + (PP) or (Prob) + (GPP), depending on the account of objective chance with which we begin.

So let us assume for the moment that we have a notion of vindication for credence functions. The first part of the argument—what I called above the general strategy—is indifferent as to the identity of the vindicated credence functions. Rather, it proceeds by laying down conditions on what would count as a measure of the distance of a particular credence function from being vindicated at a given world—following Joyce, we call this an *inaccuracy measure*; and we talk of the *inaccuracy* of a credence function as another term for its distance from vindication.

To make things precise, let's introduce some notation. Let \mathcal{F} be the algebra of propositions over which our credence functions are defined. In this article, we assume that \mathcal{F} is finite, though I will ask in section 7 whether this can be relaxed. Let \mathcal{B} be the set of possible credence functions over \mathcal{F} —that is, $\mathcal{B} = \{b : \mathcal{F} \rightarrow \mathbb{R}\}$. Finally, let \mathcal{W} be the set of possible worlds that correspond to the atoms of the finite algebra \mathcal{F} —that is, each world corresponds to an atom that is true only at that world. Then an inaccuracy measure is a function $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$ that is intended to measure the distance of the credence function b from being vindicated at world w . Given a world w , we denote as v_w the credence function that is vindicated at w . Thus, $I(b, w)$ might be thought of as measuring the distance of b from v_w . We let $\mathcal{V} := \{v_w : w \in \mathcal{W}\}$. So $\mathcal{V} \subseteq \mathcal{B}$.

In the next section, we will consider Joyce's conditions on I in some detail. But let us first see how the argument continues after these conditions are in place. The next step is the central result of Joyce's article. It involves the following set of credence functions:

$$\mathcal{V}^+ := \left\{ \sum_{w \in \mathcal{W}} \lambda_w v_w : 0 \leq \lambda_w \leq 1 \text{ and } \sum_{w \in \mathcal{W}} \lambda_w = 1 \right\}$$

This is the so-called *convex hull* of \mathcal{V} . It is the set of all convex combinations of the vindicated credence functions v_w in \mathcal{V} . Or, equivalently, it is the smallest set of credence functions that (i) contains all vindicated credence functions and (ii) contains, for any two credence functions that it contains, every *mixture* of those two credence functions. That is, $\mathcal{V} \subseteq \mathcal{V}^+$ and if $b, b' \in \mathcal{V}^+$, then $\lambda b + (1 - \lambda)b' \in \mathcal{V}^+$ for any $0 \leq \lambda \leq 1$, and \mathcal{V}^+ is the smallest set with those two properties. Joyce's theorem has two parts:

- The first part: If b is a credence function and b is not in \mathcal{V}^+ , then there is another credence function c that is in \mathcal{V}^+ and that is closer to being vindicated than b at all worlds w . That is, $I(c, w) < I(b, w)$ for all $w \in \mathcal{W}$.

That is, any agent whose credence function falls outside the convex hull of \mathcal{V} would do better *no matter how the world turns out* by moving to a particular credence in that convex hull. Adopting the language of decision theory, any credence function outside \mathcal{V}^+ is *accuracy dominated* by some credence function inside \mathcal{V}^+ . Now, at first, this might seem to give a compelling reason to have a credence function in \mathcal{V}^+ . However, for all that has been said so far, it might be that each credence function in \mathcal{V}^+ is also accuracy dominated by some other credence function. We require the second half of Joyce's theorem to rule out this possibility:

- The second part: If c is in \mathcal{V}^+ , then there is no credence function b in \mathcal{B} such that, for all $w \in \mathcal{W}$, $I(b, w) \leq I(c, w)$.

That is, if an agent has a credence function in \mathcal{V}^+ , then, by moving to any other credence function, he or she will be sure to move further from vindication in at least one world. Together, these two parts of Joyce's theorem provide the basis for the following norm:

Joyce's norm: An agent ought to have a credence function in \mathcal{V}^+ , where \mathcal{V} is the set of vindicated credence functions.

2.2. Applying the General Strategy

Notice that, in my description of Joyce's general strategy, no mention has been made of the identity of the vindicated credence functions v_w in \mathcal{V} . This first part of Joyce's argument is schematic. We complete it by identifying the vindicated credence functions and characterizing \mathcal{V}^+ , the set of convex combinations of those functions.

Joyce identifies the vindicated credence functions as follows: for each world w , a credence function is vindicated at w if it gives maximal credence to propositions that are true at w and minimal credence to propositions that are false at w . That is, $v_w(A) = 1$ if A is true at w ; and $v_w(A) = 0$ if A is false at w . Thus, for Joyce:

Truth: Set of full beliefs:: Truth: Credence function

He then appeals to a theorem of de Finetti's to show that, defined thus, \mathcal{V}^+ is the set of probability functions on \mathcal{F} .

On the other hand, I will argue that it is not truth but objective chance, or something closely related, that stands to degrees of credence as truth stands to full beliefs. More precisely, I argue that if chances are not self-undermining, then $v_w(\cdot) = ch_w(\cdot)$, where ch_w is the objective chance function at w ; and if they are self-undermining, we must choose whether $v_w(\cdot) = ch_w(\cdot)$ or $v_w(\cdot) = ch_w(\cdot \mid C_{ch_w})$. I then establish the following three analogues of de Finetti's theorem—see section 5 for precise statements of these results with all assumptions made explicit:

- If $v_w(\cdot) = ch_w(\cdot)$ and no ch is self-undermining, then \mathcal{V}^+ is the set of probability functions that satisfy (LPP).
- If $v_w(\cdot) = ch_w(\cdot)$ and some ch are self-undermining, then \mathcal{V}^+ is the set of probability functions that satisfy (GPP).
- If $v_w(\cdot) = ch_w(\cdot \mid C_{ch_w})$ and some ch are self-undermining, then \mathcal{V}^+ is the set of probability functions that satisfy (NP).

This is the strategy by which I seek to justify the chance-credence norms. In the remainder of the article, I spell out the details and justify the assumptions.

3. Inaccuracy Measures

I begin by listing the conditions we impose on inaccuracy measures $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$, which are intended to measure the distance of a credence function from vindication at a world. In each case I offer a justification.

Our first postulate says simply that, according to any inaccuracy measure, a credence function that is vindicated at a possible world should be closer to vindication at that world than a credence function that is not

vindicated at that world. That is, every inaccuracy measure is *nontrivial* in the following sense.⁷

Definition 3.1 (Nontrivial): *I* is nontrivial if, for all worlds $w \in \mathcal{W}$ and $b \in \mathcal{B}$, if $b \neq v_w$, then $I(v_w, w) < I(b, w)$.

Our second postulate says that if *I* is an inaccuracy measure then, for each $w \in \mathcal{W}$, the function $I(\cdot, w)$ on \mathcal{B} should be “continuous” in a certain sense. This is intended to rule out “jumps” in inaccuracy as we move from one credence function b' to another b that are not reflected in “jumps” in the differences $|b(A) - b'(A)|$ between the degrees of belief assigned by b and b' to the propositions in the algebra $A \in \mathcal{F}$. In particular, we demand that every inaccuracy measure is *proposition-wise continuous* in the following sense.⁸

Definition 3.2 (Proposition-wise Continuity): *I* is proposition-wise continuous if, for all $w \in \mathcal{W}$ and $b \in \mathcal{B}$, we have

$$(\forall \varepsilon > 0)(\exists \delta > 0)(\forall b' \in \mathcal{B}) \\ [(\forall A \in \mathcal{F})(|b(A) - b'(A)| < \delta) \Rightarrow |I(b, w) - I(b', w)| < \varepsilon].$$

One might think that such a “jump” should occur if we cross some sort of threshold as we move from b' to b . For instance, perhaps the degree of belief assigned to proposition A by b' is insufficient for it to count as a full belief, while the degree of belief assigned by b is sufficient for this. In this case, we cross a threshold as we move from b' to b , and we might hope that this will be reflected in our inaccuracy measure *I*—perhaps having a full belief in a truth reduces the inaccuracy of a credence function by a fixed amount, resulting in a “jump” as we cross the threshold from nonbelief to belief. Our second postulate rules this out. So is it too strong? I think not. While such a threshold might exist—if we can circumvent lottery paradox concerns—it seems that it is not relevant to the inaccuracy of a credence function. Suppose I am estimating the height of a group of people. As my estimate for a particular individual gets higher and higher,

7. In Joyce 1998, this postulate follows from **Dominance**. In fact, **Dominance** is stronger than is needed for the final proof. It is sufficient to demand that *I* is nontrivial.

8. In Joyce 1998, the continuity of $I(\cdot, w)$ is stated in terms of the Euclidean metric on \mathbb{R}^n . This leaves Joyce vulnerable to the charge that he has arbitrarily adopted this metric without argument. The following formulation avoids this concern.

it will eventually cross a threshold beyond which I will be estimating that he or she is tall. But this does not mean that our measure of inaccuracy for my estimate should be discontinuous at that threshold. The point is that there is nothing significant for the inaccuracy of my height estimates about the move from nontall to tall that is not already captured by the increase in estimated height. Similarly, there is nothing epistemically significant for the inaccuracy of an agent's epistemic state about the move from nonbelief to belief that is not already captured by the increase in the degree of belief.

Our third postulate says that the order of two credence functions $b, b' \in \mathcal{B}$ with respect to their inaccuracy as measured from $w, w' \in \mathcal{W}$ can be determined, in certain extreme situations, by the order of the differences between $b(A)$ and $v_w(A)$, on the one hand, and $b'(A)$ and $v_{w'}(A)$, on the other hand, for $A \in \mathcal{F}$. More precisely, every inaccuracy measure is *weakly difference dominating* in the following sense.⁹

Definition 3.3 (Weak difference dominance): *I is weakly difference dominating if, for all $b, b' \in \mathcal{B}$ and $w, w' \in \mathcal{W}$, if $|b(A) - v_w(A)| \leq |b'(A) - v_{w'}(A)|$ for all $A \in \mathcal{F}$, then $I(b, w) \leq I(b', w')$.*

This seems to be just what we mean when we say that $I(b, w)$ measures the distance of b from being vindicated. For note that b and v_w consist of nothing more than the values they assign to each of the propositions A in \mathcal{F} . And we know when a particular pair of assignments $b(A)$ and $v_w(A)$ are at most as far apart as another pair $b'(A)$ and $v_{w'}(A)$. It is when $|b(A) - v_w(A)| \leq |b'(A) - v_{w'}(A)|$. Thus, it seems right that b is at most as inaccurate at w as b' is at w' if $b(A)$ is at most as far from $v_w(A)$ as $b'(A)$ is from $v_{w'}(A)$, for every proposition A in \mathcal{F} . This gives us our third postulate.

Our fourth postulate says simply that as $b(A)$ tends toward ∞ or $-\infty$, the inaccuracy of the whole credence function b must tend toward ∞ . That is, every inaccuracy measure is *unbounded* in the following sense.¹⁰

Definition 3.4 (Unbounded): *I is unbounded if, for all $w \in \mathcal{W}$ and $A \in \mathcal{F}$, $I(b, w) \rightarrow \infty$ as $b(A) \rightarrow \infty$ or $b(A) \rightarrow -\infty$.*

9. In Joyce 1998, this follows from **Normality** and **Dominance**. As noted above, **Dominance** is unnecessarily strong. This weaker version entails **Normality** but not **Dominance**.

10. In Joyce 1998, this is part of **Structure**.

This says that there is no upper bound on the inaccuracy of a credence function. For any putative upper bound, there is a credence function whose inaccuracy at some world exceeds it. This seems reasonable. After all, the difference $|b(A) - v_w(A)|$ between $b(A)$ and $v_w(A)$ is unbounded. And, as we have seen, I is a function of these differences. Thus, it seems fitting that I should be similarly unbounded.

Our final three postulates concern the behavior of I when a group of agents effect a compromise between their respective credence functions. The fifth and sixth postulates concerns the situation in which two agents with distinct credence functions are equally inaccurate at a particular world. The seventh concerns a more complicated situation involving four agents.

The fifth postulate says that, when two credence functions are equally inaccurate at a world, there is a compromise between them that is more accurate than either credence function at that world. A credence function counts as a compromise between two credence functions b and b' if it lies on the line segment $bb' = \{\lambda b + (1 - \lambda)b' : 0 \leq \lambda \leq 1\}$ that runs between them. As noted above, such a credence function is called a *mixture* of b and b' . If $\lambda < \frac{1}{2}$, then $\lambda b + (1 - \lambda)b'$ is said to be biased toward b' ; if $\lambda > \frac{1}{2}$, then $\lambda b + (1 - \lambda)b'$ is said to be biased toward b ; $\frac{1}{2}b + \frac{1}{2}b'$ is not biased toward either. Thus, our fifth postulate demands that I is *disagreement resolving* in the following sense.¹¹

Definition 3.5 (Disagreement resolution): I is disagreement resolving if, for all $b, b' \in \mathcal{B}$ and $w \in \mathcal{W}$, if $b \neq b'$ and $I(b, w) = I(b', w)$, then there is $0 \leq \lambda \leq 1$ such that

$$I(\lambda b + (1 - \lambda)b', w) < I(b, w), I(b', w).$$

There are many ways to argue that such a constraint or some close relative of it should be imposed upon our inaccuracy measures. Joyce (1998) gave an invalid argument for his slightly stronger version called **Convexity**; and this was subsequently criticized by Maher (2002). More recently, Joyce (2009) has given an argument that appeals to W. K. Clifford's position in his debate with William James over the correct norms for full beliefs. However, this assumes an extra postulate, known as **Separability** or **Addi-**

11. This is weaker than Joyce's **Convexity** postulate, though that postulate follows from this along with the other postulates.

tivity, to which I do not wish to appeal here. Instead, I offer a new argument for this postulate that appeals to the notion of *epistemic disagreement*.

The epistemology of disagreement has received much-deserved attention in the recent literature on epistemology. The central question is this: What should be the effect on an agent's epistemic state of learning that an epistemic peer disagrees with him or her? Should this piece of evidence lead him or her to change his or her epistemic state, perhaps by compromising between his or her epistemic state and that of his or her peer? In Christensen's terminology, those who take the *conciliatory approach* argue that it should, while those who take the *steadfast approach* argue that it shouldn't (Christensen 2009).

I favor the conciliatory approach, at least in certain circumstances, and this leads to my fifth postulate. Suppose there are two agents with distinct credence functions b and b' respectively. In the standard cases considered in the literature on disagreement, these agents would obtain two pieces of evidence: they would learn that b and b' are distinct, and they would each learn that the other agent is an "epistemic peer." Suppose that our two agents—with credence functions b and b' respectively—learn that their credence functions are distinct. Now to learn that another agent is an epistemic peer is to learn that his or her epistemic state demands respect. In our example, suppose our two agents learn that the credence functions b and b' are equally inaccurate at the actual world. So, each learns that the other's epistemic state demands respect. After all, relative to the chief epistemic goal—closeness to vindication—they are equals. Then, according to the conciliatory approach to epistemic disagreement, it is rationally required of our agents to compromise. However, if an inaccuracy measure is not disagreement resolving in the sense just defined, our agents will not be obliged to compromise between their credence functions since there is no such compromise that is strictly better than either of their original credence functions at the actual world. Thus, in order to make it the case that compromise is a rational requirement in cases of epistemic disagreement such as this, we must at least have it that our inaccuracy measure is disagreement resolving. This gives us our fifth postulate.

The sixth postulate says that when two agents—with distinct credence functions b and b' —are equally inaccurate at a possible world, then whatever compromise is made in this situation, if it is biased toward b by a certain amount, then the compromise that is biased by the same amount toward b' is equally inaccurate. That is, $\lambda b + (1 - \lambda)b'$ is exactly as

inaccurate as $(1 - \lambda)b + \lambda b'$. So I should be *symmetric* in the following sense.¹²

Definition 3.6 (Symmetry): I is symmetric if, for all $b, b' \in \mathcal{B}$ and $w \in \mathcal{W}$, if $b \neq b'$ and $I(b, w) = I(b', w)$, then for all $0 \leq \lambda \leq 1$,

$$I(\lambda b + (1 - \lambda)b', w) = I((1 - \lambda)b + \lambda b', w).$$

This says simply that, in a case of epistemic disagreement, if the agent with credence function b accords to the agent with b' the same weight as the agent with b' accords to the agent with b , and if they then compromise in accordance with these weightings, they will end up with equally inaccurate credence functions. That is, neither b nor b' has any advantage. This is what we want if we think that any advantage would have to be discernible already in the inaccuracies of b and b' themselves, which, by hypothesis, are equal.

Our seventh and final postulate does not appear in Joyce 1998, but it is required to prove the theorem that Joyce requires for the general or schematic part of his argument. It is best illustrated by an example. Suppose there are four agents, Barb, Bob, Carl, and Col. And suppose that Barb's credence function is at least as accurate as Carl's, while Bob's is at least as accurate as Col's. Moreover, we know that there is a natural sense in which Barb's credence function is exactly as far from Bob's as Carl's is from Col's. For we know that for each proposition A in \mathcal{F} , the difference between Barb's and Bob's degrees of credence in A is exactly the same as the difference between Carl's and Col's. Then our seventh postulate says that any compromise between Barb and Bob should be at least as accurate as the corresponding compromise between Carl and Col. That is, I should be *dominated-compromise respecting* in the following sense.

Definition 3.7 (Dominated-compromise respect): I is dominated-compromise respecting if, for all $w \in \mathcal{W}$ and $b, b', c, c' \in \mathcal{B}$, if

- (i) $I(b, w) \leq I(c, w)$ and $I(b', w) \leq I(c', w)$ and
- (ii) $|b(A) - b'(A)| = |c(A) - c'(A)|$ for all $A \in \mathcal{F}$,

then

$$I(\lambda b + (1 - \lambda)b', w) \leq I(\lambda c + (1 - \lambda)c', w).$$

12. This is exactly Joyce's **Symmetry** postulate.

After all, in this situation, any compromise between Barb and Bob involves Barb moving her credence function by a certain amount toward Bob's; and the corresponding compromise between Carl and Col involves Carl moving his credence function by exactly the same amount toward Col's. Thus, since Carl is at least as far from vindication as Barb, and Col at least as far as Bob, it seems plausible that when Carl moves a certain amount toward Col and Barb moves exactly the same amount toward Bob, the result of Carl's move should be no more accurate than the result of Barb's move.

This completes our list of postulates concerning inaccuracy measures. Can we be sure that they are consistent with one another? Yes, since each member of the following family of inaccuracy measures can be shown to satisfy all of the seven postulates:

Definition 3.8: Suppose $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$. We say that I is a quadratic scoring rule if there is a family $\{\lambda_A > 0 : A \in \mathcal{F}\}$ and a strictly increasing function $F : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$I(b, w) = F\left(\sum_{A \in \mathcal{F}} \lambda_A |b(A) - v_w(A)|^2\right).$$

We say that an inaccuracy measure that satisfies our seven postulates is *Joycean*.

Definition 3.9 (Joycean): Suppose $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$. Then I is a Joycean inaccuracy measure if I is nontrivial, proposition-wise continuous, weakly difference dominating, unbounded, disagreement resolving, symmetric, and dominated-compromise respecting.

Lemma 3.10: All quadratic scoring rules are Joycean.

We are now ready to state and prove the Fundamental Theorem of Joycean Inaccuracy Measures. As above, $\mathcal{V} = \{v_w : w \in \mathcal{W}\}$, and \mathcal{V}^+ is the convex hull of \mathcal{V} .

Theorem 3.11: Suppose $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$ is Joycean. Then

- (I) If $b \notin \mathcal{V}^+$, then there is $c \in \mathcal{V}^+$ such that, for all $w \in \mathcal{W}$, $I(c, w) < I(b, w)$.

- (II) If $c \in \mathcal{V}^+$, then there is no $b \in \mathcal{B}$ such that, for all $w \in \mathcal{W}$, $I(b, w) \leq I(c, w)$

We consign the proof of this theorem to our appendix (section 8).

4. Joyce's Theorem and Probabilism

To mobilize the schematic result that I have called the “Fundamental Theorem of Joycean Inaccuracy Measures,” we must specify which credence functions are to count as vindicated at various worlds. We know what counts as vindication for sets of full beliefs. What counts as vindication for credence functions? As I have already noted, for Joyce, $v_w(A) = 1$ if A is true at w , and $v_w(A) = 0$ if A is false at w . Given this, he mobilizes his theorem by appealing to the following result:

Theorem 4.1 (de Finetti): *If \mathcal{V} is the set of classically consistent assignments of truth values to the propositions in \mathcal{F} , then \mathcal{V}^+ is the set of all (finitely additive) probability functions on \mathcal{F} .*

Combining this theorem with Theorem 3.11, we have

- (I') If $b \in \mathcal{B}$ is not a probability function, then there is a probability function p that is closer to vindication than b at all possible worlds.
- (II') If $p \in \mathcal{B}$ is a probability function, then there is no distinct credence function $b \in \mathcal{B}$ that is at least as close to vindication as p at all possible worlds.

The conjunction of these two conclusions constitutes Joyce's argument for probabilism.¹³

In fact, if Joyce's argument works, it seems to establish something more than merely probabilism. Where probabilism says only that conforming to the axioms of the probability calculus is a *necessary* condition on rationality, Joyce's argument seems to establish that it is also a *sufficient* condition. That is, Joyce's argument seems to establish an extreme version of subjectivism that rules out all norms that are stronger than probabilism. After all, (II') says that whatever probability function is your

13. In fact, in Joyce 1998, Joyce proved only Theorem 3.11 (I). Thus, he established only (I'). As noted above, this is not enough to establish (Prob). After all, without (II'), we have no guarantee that probabilists are not in the same epistemically bad situation that nonprobabilists are shown to be in by (I'). Thus, Joyce requires (II') as well. This follows from Theorem 3.11 (II).

credence function, you will do worse in at least one world by changing to another. In light of this, it is hard to see how any probabilistic agent could be accused of irrationality. But of course there are many stronger conditions than probabilism that have been proposed as necessary for rationality—indeed, the chance-credence norms we have been discussing are prominent among them. It is a central motivation of this article to adapt Joyce’s strategy to avoid this criticism.

5. Joyce’s Theorem and the Chance-Credence Norms

5.1. A Different Account of Vindication

I propose to adapt Joyce’s argument by applying his general strategy to a different notion of vindication. I wish to argue that a credence function is vindicated not when it assigns maximal credence (that is, 1) to truths and minimal credence (that is, 0) to falsehoods but when it assigns to each proposition the objective chance that the proposition is true. Of course, in the case of many propositions, this will come to the same thing: if a proposition is determined to be true, then degree of credence 1 in that proposition will be required for vindication on either account, and similarly for a proposition that is determined to be false. The difference arises when the proposition is neither determinately true nor determinately false. In this situation, we ask: What degree of credence must a credence function assign to such a proposition if it is to be vindicated? Must it assign the truth value of the proposition (represented as 0 or 1) or its chance of being true (represented, as usual, on a scale from 0 to 1)? Joyce gives the former answer; I give the latter. That is, for me:

Truth: Set of full beliefs:: Objective chance: Credence functions

In unpublished work, Alan Hájek (n.d.) has argued for a similar claim. He imagines a situation in which two agents—whom we might call Jane and Edward—are considering a particular piece of radioactive material. Jane has degree of credence 0.6 in the proposition that it will decay within the next ten days; Edward has degree of credence 1. The radioactive material does decay in that period of time; and its objective chance of doing so was 0.6. Who is vindicated? As Hájek notes, there seem to be reasons for saying that Jane is vindicated and reasons for saying that Edward is: after all, both got something right; both were correct in a certain way. But surely it can’t be that both are vindicated. I think that’s right. What’s going on here is this: When we say that an agent, such as

Edward, has degree of credence 1 or 0, we think of him as having a full belief or full disbelief, and we evaluate him according to the standards for those sorts of epistemic state. That is why we think of him as being vindicated in this case; it is because the correct notion of vindication for full beliefs is truth. But really it is only Jane's credence function that is vindicated here; Edward's set of full beliefs might be vindicated, but not his credence function. I think the same thing is going on in the following case: Suppose Jane has credence 0.7, and Edward has credence 0.9. Then we might be tempted to say that each is closer to being right about something; each is closer to the correct credence according to some salient standard of correctness. But again I think we are wrong to say that Edward's credence is closer to any salient standard of correctness *for a credence*. Rather, Edward's credence is closer than Jane's to the value—namely, 1—at which we can say with confidence that Edward has a full belief: anything short of this is potentially vulnerable to lottery paradoxes. And, once he obtains the full belief, *that full belief* is vindicated (though the *credence* is not).

Let us then seek the consequences of this account of vindication. We require separate treatments for the case in which the chances are self-undermining and the case in which they are not.

5.2. *Non-Self-Undermining Chances and the Principal Principle*

I begin with the case in which none of the possible initial chance functions is self-undermining. As we saw above, this is a consequence of certain nonreductionist accounts as well as certain reductionist accounts. In this case, each ch is certain of the proposition C_{ch} that says that the objective chances are given by ch ; in symbols, $ch(C_{ch}) = 1$. It follows that $ch(A) = ch(A | C_{ch})$, for all propositions A . So, if vindication is agreement with the objective chances, we mean this: $v_w(A) = ch_w(A) = ch_w(A | C_{ch_w})$ for all A .

With this definition of \mathcal{V} , we are ready to call upon the Fundamental Theorem of Joycean Inaccuracy Measures to establish (Prob) and (LPP). However, we must make one assumption before we proceed: we must assume that all initial chance functions are probability functions. So before we begin, we must ask if this is reasonable. Again, I appeal to an argument of Hájek's (Hájek n.d.). Philosophers of a positivist persuasion were once wont to identify objective chances with actual relative frequencies, or hypothetical limiting relative frequencies. Thus, they would say that the objective chance of an event is the relative frequency with which

events like that occur in the actual world, or the limiting relative frequency with which events like that would occur in a world in which there are infinitely many such events. Hájek (1997) has provided fifteen arguments against the actual relative frequency account of objective chance and fifteen more against the hypothetical limiting relative frequency account (Hájek 2009). However, while he denies the identification of (limiting) relative frequencies and chances, he nonetheless admits that there is some connection. The connection is this: it must be at least possible for the objective chances to match the (limiting) relative frequencies; that is, for any possible chance function ch , there must either be a possible world that is finite and such that the relative frequencies at that world match the chances given by ch , or a possible world that is infinite and such that the limiting relative frequencies at that world match the chances given by ch . And from this he infers that the objective chance function must be a (finitely additive) probability function. After all, under certain structural assumptions, we can show: (i) the (limiting) relative frequency function of a world is always a (finitely additive) probability function; (ii) any (finitely additive) probability function is the (limiting) relative frequency function of some world (van Fraassen 1983; Shimony 1988).

We are now ready to state our analogue of de Finetti's theorem:

Theorem 5.1: *Suppose C is the set of possible initial chance functions. And suppose each is a probability function and none is self-undermining. Then C^+ is the set of all (finitely additive) probability functions on \mathcal{F} that satisfy (LPP).*

Combining this theorem with Theorem 3.11, we have

- (I'') If $b \in \mathcal{B}$ does not satisfy both (Prob) + (LPP), then there is $p \in \mathcal{B}$ that satisfies (Prob) + (LPP) that is closer to vindication than b at all possible worlds (where $v_w = ch_w$, for all w).
- (II'') If $p \in \mathcal{B}$ does satisfy both (Prob) + (LPP), then there is no different $b \in \mathcal{B}$ that is at least as close to vindication as p at all possible worlds (where $v_w = ch_w$, for all w).

This completes our argument for (Prob) + (LPP).

5.3. Self-Undermining Chances, the New Principle, and the Generalized Principal Principle

The case in which at least some possible initial chance functions are self-undermining is more complicated. The problem is that our conclusion that vindication is agreement with objective chances fails to distinguish

two possible definitions of v_w . We might say: $v_w(A) = ch_w(A)$, for A in \mathcal{F} . That is, it is the chance function itself that is vindicated. Or we might say: $v_w(A) = ch_w(A | C_{ch_w})$, for A in \mathcal{F} . That is, it is the chance function *after it has been informed that it is the chance function* that is vindicated. If ch_w is self-undermining, these give two substantially different accounts of vindication between which we must choose.¹⁴ I do not have strong arguments in favor of one over the other, though it seems clear that the former is the most intuitively plausible. However, here, I will simply spell out the consequences of each. They are as follows: adopting the former notion of vindication gives us (Prob) + (GPP); adopting the latter gives (Prob) + (NP). We establish this by proving two further analogues of de Finetti's theorem.

5.3.1. The Generalized Principal Principle

Theorem 5.2: *Suppose C is the set of possible initial chance functions. And suppose each is a probability function and that each chance function expects the chance of a proposition to be the chance that it assigns to that proposition (that is, $ch_{w'}(A) = \sum_{w \in \mathcal{W}} ch_{w'}(C_{ch_w})ch_w(A)$, for all w' in \mathcal{W} and A in \mathcal{F}). Then C^+ is the set of all (finitely additive) probability functions on \mathcal{F} that satisfy (GPP).*

Note the extra assumption we require to establish our theorem: We require that the expected chance of A calculated relative to the chance function ch must be $ch(A)$. Combining this theorem with Theorem 3.11, we have

- (I''') If $b \in \mathcal{B}$ does not satisfy (Prob) + (GPP), then there is $p \in \mathcal{B}$ that satisfies (Prob) + (LPP) that is closer to vindication than b at all possible worlds (where $v_w = ch_w$, for all w).
- (II''') If $p \in \mathcal{B}$ does satisfy (Prob) + (GPP), then there is no different $b \in \mathcal{B}$ that is at least as close to vindication as p at all possible worlds (where $v_w = ch_w$, for all w).

This completes our argument for (Prob) + (GPP).

14. See Lewis 1994, 488, for an illuminating discussion of exactly how different these quantities are on a reductionist account of chance. Essentially, if A is a proposition about m outcomes of a chance process that is repeated n times throughout history, and the reductionist has extracted ch from (something close to) the relative frequencies with which the various possible outcomes of this chance process occur throughout the n trials, $ch(A)$ and $ch(A | C_{ch})$ will differ more when the ratio m/n is greater.

5.3.2. The New Principle

To state our other analogue of de Finetti's theorem, I introduce some new notation. Suppose ch is a possible initial chance function. Then let $ch^*(\cdot) = ch(\cdot \mid C_{ch})$. Thus, our second account of vindication in the presence of self-undermining chances says: $v_w(A) = ch_w^*(A)$, for all A in \mathcal{F} .

Theorem 5.3: *Suppose C^* is the set of all credence functions ch^* , where ch is a possible initial chance function. And suppose each ch (and thus each ch^*) is a probability function. Then $(C^*)^+$ is the set of all (finitely additive) probability functions on \mathcal{F} that satisfy (NP).*

Combining this theorem with Theorem 3.11, we have

- (I''') If $b \in \mathcal{B}$ does not satisfy (Prob) + (NP), then there is $p \in \mathcal{B}$ that satisfies (Prob) + (NP) that is closer to vindication than b at all possible worlds (where $v_w = ch_w^*$, for all w).
- (II''') If $p \in \mathcal{B}$ does satisfy (Prob) + (NP), then there is no different $b \in \mathcal{B}$ that is at least as close to vindication as p at all possible worlds (where $v_w = ch_w^*$, for all w).

This completes our argument for (Prob) + (NP).

6. Objections

Before turning to ways in which this argument might be extended, I wish to consider two possible objections to it. Both stem from objections that have been raised against Joyce's original argument.

6.1. Bronfman's Objection to Joyce

Aaron Bronfman (n.d.) has raised the following objection against Joyce's argument for probabilism. Joyce's argument establishes that, for any non-probabilistic credence function b and any legitimate inaccuracy measure I , there is at least one probabilistic credence function p that is more accurate than b at all possible worlds according to I . In symbols, this is:

$$(\forall b)(\forall I)(\exists p)(\forall w)[I(p, w) < I(b, w)] \quad (\text{JA1})$$

The problem is that, because of the order of the quantifiers, p depends not only on b but also on I . That is, Joyce does not establish that for any nonprobabilistic b , there is a probabilistic p that is more accurate than b at all possible worlds and according to any legitimate inaccuracy measure.

In symbols, this would be:

$$(\forall b)(\exists p)(\forall I)(\forall w)[I(p, w) < I(b, w)] \tag{JA2}$$

Thus, given nonprobabilistic b and two different legitimate inaccuracy measures I and I' , Joyce's result does not rule out the following situation: all p that are more accurate than b at all worlds *according to* I are not more accurate than b at all worlds *according to* I' , and vice versa. In this situation, it is not clear that the agent is rationally required—or even rationally permitted—to give up his or her nonprobabilistic credence function b . For whichever probabilistic credence function p he or she moves to, there is a legitimate inaccuracy measure—either I or I' —according to which p is not always more accurate than b at all worlds. Indeed, it might be that the credence functions that are always more accurate according to I' are very inaccurate indeed at some worlds according to I , and vice versa. If the agent wishes to avoid any possibility of such great inaccuracy, he or she would do well to stick with b . This is Bronfman's objection. And clearly it is just as powerful as an objection against my arguments for the various chance-credence norms.

But is it powerful against either argument? This depends on what we are doing when we lay down conditions on legitimate inaccuracy measures. One possibility is that we are providing a characterization of the class of all legitimate inaccuracy measures. That is, we are giving not only necessary but also sufficient conditions on being a legitimate inaccuracy measure. If this is the case, then Bronfman's objection is decisive. If any Joycean function is a legitimate inaccuracy measure, then we cannot infer from Joyce's result that an agent ought to obey probabilism. After all, moving from a nonprobabilistic credence function to a probabilistic one might always risk great inaccuracy by the lights of some legitimate inaccuracy measure.

But there are two alternative views of our conditions on inaccuracy measures. On the first, there is a unique correct measure of inaccuracy. We do not know enough about it to characterize it uniquely, but we can know some features of it, namely, the postulates we laid down in section 3. If that is the case, then Bronfman's objection fails. If there is just one legitimate inaccuracy measure, then the variable I in (JA1) and (JA2) ranges over only that measure. Thus, (JA1) and (JA2) are equivalent. And Bronfman accepts that if Joyce has (JA2), then his argument goes through.

On the second alternative view, there is no single measure of inaccuracy that is correct for all agents. However, for any given agent, there is a

single measure of inaccuracy that reflects that agent's epistemic values. On this view, when we lay down conditions on legitimate inaccuracy measures, we lay down basic objective constraints on the sort of epistemic values that an agent should have. But they are not sufficient to constrain those epistemic values completely. They leave room for significant variation between agents, as is witnessed by Lemma 3.10, which says that an agent may take any quadratic scoring rule to be his or her inaccuracy measure. But nonetheless, Bronfman's objection is blocked. Focus attention on any agent. Then he or she considers only one inaccuracy measure correct. Then again the variable I in (JA1) and (JA2) ranges over only that measure. Thus, (JA1) and (JA2) are equivalent.

I favor the latter response to the version of Bronfman's objection that is directed against my argument. After all, it would require a great deal of further argument to make it plausible that there is a single inaccuracy measure that is correct for any agent, even though we are unable to characterize it fully. So my view is this: The correct notion of vindication is agreement with objective chance. And, for any agent, there is a unique function that measures the distance of a credence function from that goal of vindication for that agent. That unique function must satisfy certain basic rational constraints, which are listed as our postulates in section 3. This is enough to establish that any agent ought to obey probabilism and one of the chance-credence norms—which chance-credence norm will depend on one's account of chance. After all, these norms will follow whichever inaccuracy measure an agent has, providing it satisfies the postulates laid down. That is the lesson of Theorem 3.11.

6.2. Fitelson and Easwaran's Objection to Joyce

In recent joint work, Branden Fitelson and Kenny Easwaran (forthcoming) have raised the following objection against Joyce's argument for probabilism. Recall that, for Joyce, an agent with a nonprobabilistic credence function b is irrational because there is at least one credence function p that is better than b no matter how the world turns out; that is, there is p that accuracy-dominates b . But this will fail to render our agent irrational if she has good reason not to adopt the credence function p , or any other credence function that also accuracy-dominates b . And, as Fitelson and Easwaran point out, most agents will have good reasons to avoid many different credence functions. The reason is that most agents accumulate evidence during their epistemic life. And, having accumulated this evidence, they have good reason to avoid any credence function

that does not respect it. For instance, an agent who has learned proposition E with certainty has good reason to avoid credence functions that assign E a degree of credence less than 1. Or consider an agent who has learned evidence in a more realistic way. For instance, he or she might have had an experience that led her to assign degree of credence 0.4 to the proposition E . Then he or she has good reason to avoid credence functions that assign E a degree of credence other than 0.4 (at least in the absence of new evidence). If this is granted, then Joyce's argument fails for many such agents. To state just one example: Suppose we are given the four-element algebra based on atoms A and $\neg A$. Then there is a nonprobabilistic credence function b on this algebra and a Joycean inaccuracy measure I such that (i) $b(A) = 0.4$ and (ii) there is no probabilistic credence function p on this algebra for which $p(A) = 0.4$ and p accuracy-dominates b relative to I . Thus, if an agent has credence function b and if he or she has good reason to retain the degree of credence 0.4 for the proposition A , Joyce's argument has no force against him or her.

How is my argument to avoid this same objection? I avoid it because my argument concerns only an agent's *initial* or *ur-* credence function, the credence function with which he or she begins his or her epistemic life. At that stage, the agent can have no good reason for placing any constraints on the degrees of credence that he or she assigns, for he or she has accumulated no evidence. Thus, there is no credence function that he or she has good reason to avoid. Thus, my version of Joyce's argument goes through.

7. Extensions

So far, I have argued that Joyce's nonpragmatic vindication of (Prob) can be adapted to give a nonpragmatic vindication of (Prob) and either (LPP) or (NP) or (GPP) for agents whose credence functions are defined on a finite algebra \mathcal{F} . As we have seen, the trick is to change the notion of vindication for credence functions. I conclude by considering one way in which the results of this article might be extended.

The obvious extension is to the case in which \mathcal{F} is infinite. Here, I focus only on countably infinite σ -algebras. As is clear from the proof of Theorem 3.11, it holds for credence functions defined on any algebra \mathcal{F} , finite or infinite; though when \mathcal{F} is infinite, we must make sure to define \mathcal{V}^+ to be the *closed* convex hull of \mathcal{V} .¹⁵ And it is easy to see that

15. Given a subset $X \subseteq \mathcal{B}$ and $b \in \mathcal{B}$, we say that b is a *limit point* of X if, for any $\varepsilon > 0$,

our proofs of Theorems 5.1 and 5.2 can be adapted to the countably infinite case, providing each objective chance function obeys Countable Additivity:

Countable Additivity (CA): An agent's credence function b ought to be *countably additive*. That is, for any countably infinite set of mutually inconsistent propositions $\{A_i : i = 1, 2, \dots\}$, the right-hand side of the following identity exists, and the identity holds:

$$b\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} b(A_i)$$

In this case, the theorem says that the closed convex hull of the set of objective chance functions is the set of all *countably additive* probability functions on countably infinite σ -algebra \mathcal{F} that obey (LPP) or (GPP) or (NP). Thus, in the countably infinite case, we can justify (Prob), (CA), and our favored chance-credence norm, providing the objective chance functions obey (CA).¹⁶

The problem, of course, is that it is not clear that the objective chance functions are countably additive. What is certain is that we cannot appeal to Hájek's argument from possible agreement with relative frequencies to justify the claim that they do. In the infinite case, we have to move from relative frequencies to limiting relative frequencies, and it is well known that limiting relative frequencies do not necessarily obey (CA). Furthermore, it seems at least possible that there should be a fair infinite lottery, and it is well known that no countably additive probability function can deliver this.¹⁷ I leave the following as an open question: is it the case that every credence function on a countably infinite σ -algebra \mathcal{F}

there is $b' \in X$ such that, for every $A \in \mathcal{F}$, $|b(A) - b'(A)| < \varepsilon$. That is, any neighborhood around b —however small—contains an element of X . We say that X is *closed* if it contains all its limit points. Thus, the *closed* convex hull of \mathcal{V} is the smallest closed convex subset of \mathcal{B} that contains \mathcal{V} . When \mathcal{F} is finite, the convex hull of \mathcal{V} is closed; that is, the closed convex hull is the convex hull.

16. We also require the following lemma: If all credence functions in a given set obey (Prob), (CA), and our favored chance-credence norm, then so do the credence functions that are the limit points of that set.

17. If an infinite lottery is fair, the chance of each of the infinitely many, exhaustive, and mutually exclusive outcomes is given by the same real number ε . If $\varepsilon > 0$, then the infinite sum of the chances of the outcomes does not exist. If $\varepsilon = 0$, then the infinite sum of the chances of the outcomes does exist, but it is 0, while the probability of the disjunction of all the outcomes is 1. In either case, the objective chance function is not countably additive.

that lies in the closed convex hull of the set of objective chance functions obeys one of the chance-credence norms, even if the objective chance functions do not all obey (CA)?

8. Appendix

8.1. Proof of Theorem 3.11

Theorem 3.11: *Suppose $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$ is Joycean. Then*

- (I) *If $b \notin \mathcal{V}^+$, then there is $p \in \mathcal{V}^+$ such that, for all $w \in \mathcal{W}$, $I(p, w) < I(b, w)$.*
- (II) *If $p \in \mathcal{V}^+$, then there is no $b \in \mathcal{B}$ such that, for all $w \in \mathcal{W}$, $I(b, w) \leq I(p, w)$.*

The proof of (I) follows Joyce’s original proof very closely; the proof of (II) is new.

Proof. Suppose $I : \mathcal{B} \times \mathcal{W} \rightarrow \mathbb{R}$ is Joycean. Then our first task is to use I to define a measure of distance between any two credence functions $b, c \in \mathcal{B}$. Define $D : \mathcal{B} \times \mathcal{B} \rightarrow \mathbb{R}$ as follows:

$$D(b, c) = I(v_w + (b - c), w)$$

for $w \in \mathcal{W}$. Since I is weakly difference dominating, we have:

- The definition of D does not depend on the element $w \in \mathcal{W}$ used in its definition.
- For $b \in \mathcal{B}$ and $w \in \mathcal{W}$, $I(b, w) = D(b, v_w)$.
- For all $b, c \in \mathcal{B}$, $D(b, c) = D(c, b)$.

Now we embark on our proofs of (I) and (II).

(I) Suppose $b \notin \mathcal{V}^+$. Now consider the function $D(b, \cdot) : \mathcal{V}^+ \rightarrow \mathbb{R}$. Since I is proposition-wise continuous, $D(b, \cdot)$ is continuous with respect to the Euclidean metric. Thus, $D(b, \cdot)$ is a real-valued, continuous function on a closed, bounded subset of \mathbb{R}^n . Thus, by the Extreme Value Theorem, the function $D(b, \cdot)$ obtains a minimum on \mathcal{V}^+ . Moreover, since I is disagreement resolving and \mathcal{V}^+ is convex, this minimum must be unique. Let us denote the unique minimum of $D(b, \cdot)$ on \mathcal{V}^+ as p .

Our job now is to show that, for all $w \in \mathcal{W}$, $D(p, v_w) < D(b, v_w)$, and thus $I(p, w) < I(b, w)$. Suppose $w \in \mathcal{W}$. If $p = v_w$, then since I is nontrivial, $D(p, v_w) < D(b, v_w)$, and we’re done. Thus, we suppose that

$p \neq v_w$. Let

$$R = \{\lambda p + (1 - \lambda)v_w : -\infty < \lambda < \infty\}$$

We now prove that there is $m \in R$ such that

- (a) $D(m, v_w) \geq D(p, v_w)$
- (b) $D(b, v_w) > D(m, v_w)$

Consider the function $D(b, \cdot)$ on R . Again, this is a continuous, real-valued function on R . And, since I is unbounded, D tends to infinity as we move away from v_w toward p and beyond. But since $D(b, p) < D(b, v_w)$, the Intermediate Value Theorem gives us that there is $k \in R$ that lies beyond p on the line from v_w through p to infinity and is such that $D(b, k) = D(b, v_w)$. Then, since I is symmetric and disagreement resolving, $\frac{1}{2}k + \frac{1}{2}v_w$ is the unique minimum of $D(b, \cdot)$ on the segment kv_w of this line. Let $m := \frac{1}{2}k + \frac{1}{2}v_w$. We now prove that m satisfies (a) and (b):

- If $m = p$, then certainly $D(m, v_w) \geq D(p, v_w)$. Thus, suppose $m \neq p$. First, we show that p lies on the line segment mv_w . Suppose not. Then m must lie on the line segment pv_w . But then $m \in \mathcal{V}^+$ and $D(b, m) \leq D(b, p)$, which contradicts the minimality of $D(b, \cdot)$ on \mathcal{V}^+ at p . Thus, p lies on the line segment mv_w . Thus, since I is weakly difference dominating, $D(m, v_w) \geq D(p, v_w)$. Thus, (a).
- By the definition of k , we have $D(b, k) = D(b, v_w)$. Since I is weakly difference dominating, we have $D(b, k) = D(b, 2b - k)$. Thus, we have

$$D(b, v_w) = D(b, 2b - k)$$

And thus $D(b, \cdot)$ obtains a unique minimum on the line segment $v_w(2b - k)$ at $\frac{1}{2}(v_w - k) + b$. Thus, in particular,

$$D(b, v_w) > D\left(b, \frac{1}{2}(v_w - k) + b\right)$$

But, again since I is weakly difference dominating, from the definition of m we have

$$D\left(b, \frac{1}{2}(v_w - k) + b\right) = D(m, v_w)$$

Thus, (b).

Putting (a) and (b) together gives us (I).

(II) We define D as before. Suppose, for a reductio, that $p \in \mathcal{V}^+$ and $b \in \mathcal{B}$ and, for all $w \in \mathcal{W}$, $I(b, w) \leq I(p, w)$. Thus, $D(b, v_w) \leq D(p, v_w)$. Then define the following subset $\mathcal{A} \subseteq \mathcal{V}^+$:

$$\mathcal{A} := \{c \in \mathcal{V}^+ : D(b, c) \leq D(p, c)\}$$

Then:

- (i) $p \notin \mathcal{A}$, since I is nontrivial, and thus $D(b, p) > D(p, p)$.
- (ii) $\mathcal{V} \subseteq \mathcal{A}$, by assumption.
- (iii) \mathcal{A} is convex, since I is dominated compromise-respecting, and thus we have: for all $c, c' \in \mathcal{B}$, if $D(b, c) \leq D(p, c)$ and $D(b, c') \leq D(p, c')$, then for all $0 \leq \lambda \leq 1$,

$$D(b, \lambda c + (1 - \lambda)c') \leq D(p, \lambda c + (1 - \lambda)c')$$

Thus, $\mathcal{V} \subset \mathcal{A} \subsetneq \mathcal{V}^+$ and \mathcal{A} is convex. This is a contradiction, since \mathcal{V}^+ is the convex hull of \mathcal{V} . Thus, (II).

8.2. Proof of Theorems 5.1, 5.2, and 5.3

Theorem 5.1 follows from Theorem 5.3 when we note that if ch is not self-undermining, then (i) $ch_w^* = ch_w$, for all w , and (ii) (NP) and (LPP) are equivalent.

Theorem 5.3: *Suppose \mathcal{C}^* is the set of all credence functions ch^* , where ch is a possible initial chance function. And suppose each ch (and thus each ch^*) is a probability function. Then $(\mathcal{C}^*)^+$ is the set of all (finitely additive) probability functions on \mathcal{F} that satisfy (NP).*

Proof. The proof has two parts: first, we show that, if $b \in (\mathcal{C}^*)^+$, then b satisfies (Prob) and (NP); second, we show the converse.

First, we show that if $b \in (\mathcal{C}^*)^+$, then b satisfies (Prob) and (NP).

We show this in two stages:

- We begin by showing that, for all $w \in \mathcal{W}$, ch_w^* satisfies (Prob) and (NP). For each $w \in \mathcal{W}$, ch_w satisfies (Prob) by assumption, and thus so does ch_w^* . Suppose $w, w' \in \mathcal{W}$. Then, if $ch_{w'}^*(C_{ch_w}) > 0$, then $w = w'$, and thus

$$ch_{w'}^*(A | C_{ch_w}) = ch_w(A | C_{ch_w})$$

as required.

- Then we show that if b and b' satisfy (Prob) and (NP), then so does any mixture of b and b' . Suppose b and b' satisfy (Prob) and (NP). Then

$$\begin{aligned}
 (\lambda b + (1 - \lambda)b')(A | C_{ch}) &= \frac{\lambda b(A \wedge C_{ch}) + (1 - \lambda)b'(A \wedge C_{ch})}{\lambda b(C_{ch}) + (1 - \lambda)b'(C_{ch})} \\
 &= \frac{\lambda b(A | C_{ch})b(C_{ch}) + (1 - \lambda)b'(A | C_{ch})b'(C_{ch})}{\lambda b(C_{ch}) + (1 - \lambda)b'(C_{ch})} \\
 &= \frac{\lambda ch(A | C_{ch})b(C_{ch}) + (1 - \lambda)ch(A | C_{ch})b'(C_{ch})}{\lambda b(C_{ch}) + (1 - \lambda)b'(C_{ch})} \\
 &= ch(A | C_{ch})
 \end{aligned}$$

as required.

This suffices to establish that if $b \in (C^*)^+$, then b satisfies (Prob) and (NP). After all, it shows that the set of credence functions that satisfies both norms is convex and contains C^* . Since $(C^*)^+$ is the smallest such set, $(C^*)^+$ is a subset of it, as required.

Second, we show that if b satisfies (Prob) and (NP), then $b \in (C^*)^+$. Suppose b satisfies (Prob) and (NP). Then let $A \in \mathcal{F}$. Then

$$\begin{aligned}
 b(A) &= \sum_w b(A | C_{ch_w})b(C_{ch_w}) \quad (\text{by (Prob)}) \\
 &= \sum_w ch(A | C_{ch_w})b(C_{ch_w}) \quad (\text{by (NP)}) \\
 &= \sum_w ch_w^*(A)b(C_{ch_w})
 \end{aligned}$$

Thus, $b \in (C^*)^+$, as required.

Theorem 5.2: *Suppose C is the set of possible initial chance functions. And suppose each is a probability function and that each chance function expects the chance of a proposition to be the chance that it assigns to that proposition (that is, $ch_{w'}(A) = \sum_{w \in \mathcal{W}} ch_{w'}(C_{ch_w})ch_w(A)$, for all w' in \mathcal{F} and A in \mathcal{F}). Then C^+ is the set of all (finitely additive) probability functions on \mathcal{F} that satisfy (GPP).*

Proof. As above, the proof has two parts: first, we show that, if $b \in C^+$, then b satisfies (Prob) and (GPP); second, we show the converse.

First, we show that if $b \in C^+$, then b satisfies (Prob) and (GPP). We show this in two stages:

- We begin by noting that for all $w \in \mathcal{W}$, ch_w satisfies (Prob) and (GPP) by assumption.

- We then observe that, if b and b' satisfy (Prob) and (GPP), then so does any mixture of b and b' . This is straightforward.

Second, we show that, if b satisfies (Prob) and (GPP), then $b \in C^+$. But this follows immediately from the fact that b satisfies (GPP). For then b is a convex combination of the elements of C : that is, there are λ_w such that $\sum_{w \in \mathcal{W}} \lambda_w = 1$ and

$$b = \sum_{w \in \mathcal{W}} \lambda_w c_{h_w}$$

For $\lambda_w = b(C_{c_{h_w}})$. This completes our proof.

References

- Bogdan, R., ed. 1984. *Henry E. Kyburg, Jr. and Isaac Levi*. Dordrecht: D. Reidel.
- Bronfman, A. n.d. "A Gap in Joyce's Argument for Probabilism." Unpublished manuscript.
- Christensen, D. 2009. "Disagreement as Evidence: The Epistemology of Controversy." *Philosophy Compass* 4, no. 5: 756–67.
- Fitelson, Branden, and Kenny Easwaran. Forthcoming. "An 'Evidentialist' Worry about Joyce's Argument for Probabilism." *Dialectica*.
- Hájek, A. 1997. "Mises Redux—Redux: Fifteen Arguments against Finite Frequentism." *Erkenntnis* 45: 209–27.
- . 2009. "Fifteen Arguments against Hypothetical Frequentism." *Erkenntnis* 70: 211–35.
- . n.d. "A Puzzle about Partial Belief." Unpublished manuscript.
- Hall, N. 1994. "Correcting the Guide to Objective Chance." *Mind* 103: 505–18.
- . 2004. "Two Mistakes about Credence and Chance." *Australasian Journal of Philosophy* 82, no. 1: 93–111.
- Ismael, J. 2008. "Raid! Dissolving the Big, Bad Bug." *Noûs* 42, no. 2: 292–307.
- Joyce, J. M. 1998. "A Nonpragmatic Vindication of Probabilism." *Philosophy of Science* 65, no. 4: 575–603.
- . 2007. "Epistemic Deference: The Case of Chance." *Proceedings of the Aristotelian Society* 107, no. 2: 187–206.
- . 2009. "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief." In *Degrees of Belief*, ed. F. Huber and C. Schmidt-Petri, 263–97. Dordrecht: Springer.
- Lewis, D. 1980. "A Subjectivist's Guide to Objective Chance." In *Studies in Inductive Logic and Probability*. Vol. 2, ed. R. Carnap and R. C. Jeffrey, 263–93. Berkeley: University of California Press. Reprinted with added postscripts in Lewis, D., 1987. *Philosophical Papers*, vol. 2, 83–132. Oxford, UK: Oxford University Press. All references are to the original version.
- . 1994. "Humean Supervenience Debugged." *Mind* 103: 473–90.

Accuracy, Chance, and the Principal Principle

- Mackie, J. L. 1977. *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin.
- Maher, P. 2002. "Joyce's Argument for Probabilism." *Philosophy of Science* 69, no. 1: 73–81.
- Meacham, C. J. G. 2010. "Two Mistakes regarding the Principal Principle." *British Journal for the Philosophy of Science* 61, no. 2: 407–31.
- Schaffer, J. 2003. "Principled Chances." *British Journal for the Philosophy of Science* 54: 27–41.
- Shimony, A. 1988. "An Adamite Derivation of the Calculus of Probability." In *Probability and Causality*, ed. J. Fetzer, 79–89. Dordrecht: D. Reidel.
- Thau, M. 1994. "Undermining and Admissibility." *Mind* 103: 491–504.
- van Fraassen, B. C. 1983. "Calibration: Frequency Justification for Personal Probability." In *Physics, Philosophy, and Psychoanalysis*, ed. R. S. Cohen and L. Laudan, 295–314. Dordrecht: Springer.

