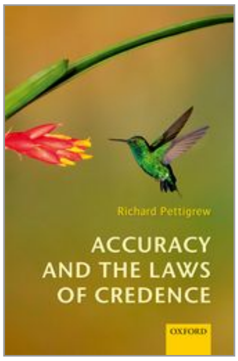
University Press Scholarship Online
Oxford Scholarship Online



Accuracy and the Laws of Credence
Richard Pettigrew

Print publication date: 2016
Print ISBN-13: 9780198732716
Published to Oxford Scholarship Online: May 2016
DOI: 10.1093/acprof:oso/9780198732716.001.0001

Vindication and chance

Richard Pettigrew

DOI:10.1093/acprof:oso/9780198732716.003.0010

Abstract and Keywords

This chapter introduces a first attempt at an accuracy-based argument for chance-credence principles. It adapts the accuracy-based argument for Probabilism by changing the notion of vindication from that proposed by the veritist to an alternative that has been endorsed by Alan Hájek. The chapter argues that, ultimately, it fails.

Keywords: Vindication, accuracy, Principal Principle, chance, Alan Hájek

So much for the formulation of the Principal Principle as PP_0. Let us turn now to its justification. In two recent papers, I offered two accuracy-based arguments in its favour (Pettigrew, 2012, 2013). In this chapter, we consider the first; in the next, we consider the second. In the end, I will reject the first argument and accept an improved version of the second. Readers more interested in the more promising second argument can skip this chapter without loss.

To introduce the first argument, it will help to recall our accuracy argument for Probabilism from above:

(I_P^*) **Veritism** The ultimate source of epistemic value is accuracy.

(II^*_P) **Brier Alethic Accuracy** The inaccuracy of a credence function at a world is measured by its Brier score at that world.
This is a consequence of the following axioms:

(i) **Perfectionism** The accuracy of a credence function at a world is its proximity to the vindicated credence function at that world.

(ii) **Squared Euclidean Distance** Distance between credence functions is measured by squared Euclidean distance.
This is a consequence of the following axioms:

- **Perfectionism** (cf. Chapter 4)
- **Divergence Additivity** (cf. Section 4.1)
- **Divergence Continuity** (cf. Section 4.2)
- **Decomposition** (cf. Section 4.3)
- **Symmetry** (cf. Section 4.4)
- **Theorem** 4.4.1

(iii) **Alethic Vindication** The vindicated credence function at a world is the omniscient credence function at that world.

(III^*_P) **Immodest Dominance**

(IV^*_P) **Theorems I.D.5** and **I.D.7** and **I.A.2** (which together entail **Theorem 4.3.4**) and **Proposition I.B.2.**
Therefore,
(V^*_P) **Probabilism**

The first accuracy-based argument for PP_0 adapts this argument for Probabilism by offering an alternative account of vindication: it replaces Alethic Vindication with a (p.107) claim that I will call Ur-Chance Initial Vindication; this results in replacing the second premise —namely, Brier Alethic Accuracy—with a new axiom called Brier Ur-Chance Initial Accuracy.

Ur-Chance Initial Vindication The vindicated initial credence function at a world is the ur-chance function at that world.

This is one way of making precise a proposal of Alan Hájek's. Hájek asks us to fill in the blank in the following sentences (Hájek, ms):

Truth is to belief as____is to degrees of belief

and

> A belief is vindicated by truth; a degree of belief is vindicated by____

He offers the following answer: vindication for credences is 'agreement with the objective chances'. But which objective chances? Later in this chapter, we'll consider one natural answer: vindication for credences is agreement with the objective chances *at the time the credence is held*. But, to begin with, we consider the answer given by Ur-Chance Initial Vindication: vindication for *initial* credences is agreement with the *ur*-chances.

Together with Perfectionism, this entails that the accuracy of an initial credence function at a world is its proximity to the ur-chance function of that world; so its inaccuracy is its distance from that ur-chance function. Now, in Part I, we argued that the distance from one credence function *c* to another *c′* defined on the same set of propositions is given by the squared Euclidean distance between *c* and *c′*. Thus, granted Ur-Chance Initial Vindication, the second premise of our accuracy argument becomes:

> (II_{PP_0}) **Brier Ur-Chance Initial Accuracy** The inaccuracy of an initial credence function at a world is the squared Euclidean distance from the ur-chance function at that world to the credence function.
> That is:
> $$\mathfrak{J}(c_0, w) = \sum_{X \in \mathcal{F}} |ch_w(X) - c_0(X)|^2$$

Of course, our arguments in favour of squared Euclidean distance above—in particular, our argument for Decomposition—relied on Alethic Vindication. Thus, having rejected that account of vindication, we have no right to assume that distances are given by that particular divergence. But let us assume that for the sake of argument.[1] I will be rejecting the argument on other grounds.

(p.108) To see how this allows us to argue for PP_0, recall again the mathematics behind our accuracy argument for Probabilism. If the inaccuracy of a credence function is given by its distance from the vindicated credence function, and if the distance between two credence functions is measured by squared Euclidean distance—or, indeed, by any additive Bregman divergence—then we have the following:

(I) If a credence function lies outside the closure of the convex hull of the set of vindicated credence functions, then it is accuracy-dominated by a credence function that lies inside the convex hull (where the convex hull of a set of credence functions is the smallest convex set that contains all of the credence functions; that is, it is the set of all mixtures of those credence functions).

(II) If a credence function lies inside the convex hull of the set of vindicated credence functions and is a probability function (which is guaranteed by its being inside the convex hull if all vindicated credence functions are probability functions), then it expects itself to be most accurate.

This is the content of Theorems I.D.5 and I.D.8. Thus, in order to apply Immodest Dominance and derive a principle of credal rationality from Ur-Chance Initial Vindication together with Perfectionism and the claim that the distance between credence functions is the squared Euclidean distance between them, we need only find properties that are shared by all credence functions in the closure of the convex hull of the ur-chance functions. These properties will then be rationally required of a credence function: after all, if a credence function fails to have any one of them, then the credence function must lie outside the closure of the convex hull; and thus it must be dominated by an immodest credence function. As it turns out, all credence functions in the closure of the convex hull satisfy PP_0.[2]

Theorem 9.0.9 *Let $\mathscr{C}_0$ be the set of possible ur-chance functions. Then, if c is in* $\mathrm{cl}(C_0^+)$, *then c satisfies PP_0.*[3]

(This is a corollary of Theorem III.D.2. It follows if we let $\mathscr{S} = \mathscr{C}_0$ in that theorem.) Note that this is analogous to De Finetti's Characterization Theorem (Theorem I.A.2), which (p.109) characterizes the probability functions as the elements of the closure of the convexhull of the omniscient credence functions. However, unlike De Finetti's Characterization Theorem, this theorem only shows a necessary condition on being in the closure of the convex hull of the vindicated credence functions; de Finetti's result also shows that that condition is suffcient.

Thus, we have the following argument for PP_0:

(I_{PP_0}) **Veritism** The sole fundamental source of epistemic value is accuracy.

(II_{PP_0}) **Brier Ur-Chance Initial Accuracy** The inaccuracy of an initial credence function at a world is the squared Euclidean distance from the ur-chance function at that world to the credence function.

This is a consequence of the following axioms:

(i) **Perfectionism** The accuracy of a credence function at a world is its proximity to the vindicated credence function at that world.

(ii) **Squared Euclidean Distance** Distance between credence functions is measured by squared Euclidean distance.

(iii) **Ur-Chance Initial Vindication** The vindicated initial credence function at a world is the ur-chance function at that world.

(III_{PP_0}) **Immodest Dominance**

(IV_{PP_0}) **Theorems I.D.5** and **I.D.8** and **I.B.2** and **9.0.9** Therefore,

(V_{PP_0}) **Probabilism + PP**$_0$.

After stating my final version of the accuracy argument for Probabilism, I noted that, while it is necessary for the success of that argument that we assume that there is a unique legitimate inaccuracy measure, and while it's necessary that that unique measure must be one that is generated by an additive Bregman divergence, it is not necessary for the success of the argument that the additive Bregman divergence we use is squared Euclidean distance. Any additive Bregman divergence will do just as well. The same is true of this argument. For any additive Bregman divergence $\mathfrak{D}$, if we replace Squared Euclidean Distance with the claim that $\mathfrak{D}(c, c')$ gives the distance between credence functions c and c', then the argument will still go through. This is due to the power of Theorems I.D.5 and I.D.8.

Thus, what is wrong with Cleos credences is this: There are immodest credence functions that accuracy dominate hers. In Part I, where we assumed Alethic Vindication, this was only true of credence functions that violate Probabilism. But, in the presence of Ur-Chance Initial Vindication, even if we assume that Cleo's credence function satisfies Probabilism, if the only possible ur-chance functions are ones on which the ur-chance of heads is 60% or 70% and if both exceed Cleo's credence in heads, then there is a credence function that is closer to all possible ur-chance functions than Cleo's is, and which expects its distance from the ur-chance function to be lowest. Figure 9.1 illustrates

Cleo's situation; Figure 9.2 further illustrates the argument for PP_0 given above. (p.110)

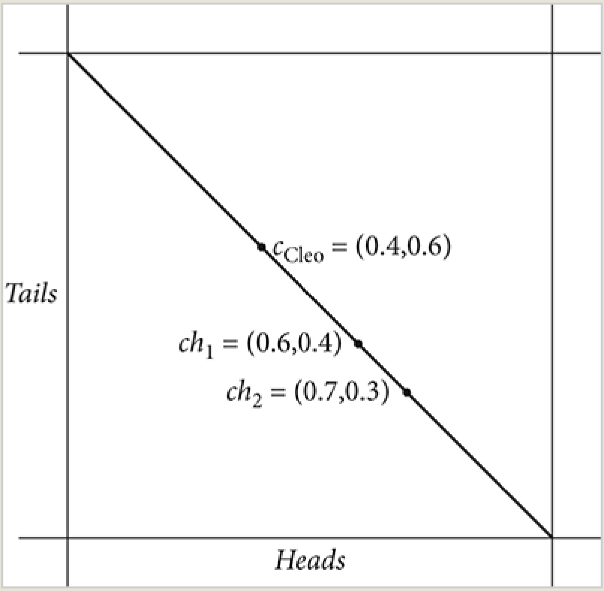


Figure 9.1 Recall: Cleo knows that the ur-chance of the coin landing heads is either 60% or 70%. Thus, the possible ur-chance functions are ch_1 and ch_2, as marked in the figure. By Theorem 9.0.9, all of the credence functions in the convex hull of these two credence functions—that is, those on the straight line between ch_1 and ch_2—satisfy PP_0. Indeed, in this situation, since there are only finitely many possible ur-chance functions, the converse holds as well: that is, every credence function that satisfies PP_0 lies in that convex hull. Cleo's credence in *Heads* is less than 0.5. Let's assume, for instance, that it is 0.4; and her credence in *Tails* is 0.6. Then she violates PP_0. And we can see that there is an alternative credence function—for instance ch_1 itself—that is closer to ch_1 and to ch_2.

(p.111)

9.1 Objections to Ur-Chance Initial Vindication

Michael Caie rejects Ur-Chance Initial Vindication (Caie, 2015). His main argument turns on a comparison with Alethic Vindication.

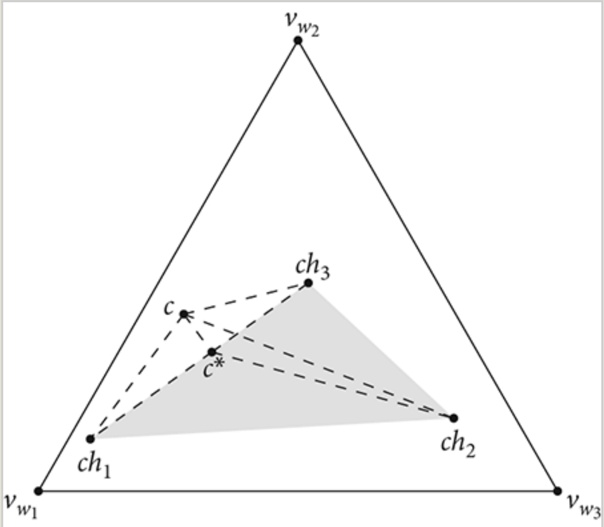


Figure 9.2 In this example, our agent has credences in the three elements of a partition $\{X_1, X_2, X_3\}$. Let w_1, w_2, and w_3 be the possible worlds relative to this set of propositions, where X_i is true at world w_i. Then the triangle in the figure represents the set of probabilistic credence functions—it is the convex hull of the set of omniscient credence functions, v_{w1}, v_{w2}, and v_{w3}. Suppose that there are three possible ur-chance functions, ch_1, ch_2, and ch_3, as pictured. And suppose that c violates PP_0. Therefore, by Theorem 9.0.9, c lies outside the convex hull of these possible ur-chance functions, which is the shaded grey triangle. And thus we can see that there is a credence function inside that convex hull—for instance, c^*—that is closer to each of the possible ur-chance functions than c is. Thus, if vindication is agreement with the ur-chance function, then c is dominated.

So far in this book, we have been focussing on agents who have credences only in what we might call *timeless* or *eternal* propositions. These are propositions whose truth values are invariant across different times in a given possible world: thus, an eternal proposition, if true at some time in a possible world, is true at all times in that world; if false at some time, it is false at all times. For instance, the proposition *Salt always dissolves in water* is eternal; similarly, *A sea battle takes place on 21 October 1805*. Since the truth value of such a proposition is determined only relative to a possible world—and not relative to a possible world together with a time within it— omniscient credence functions are given only relative to a possible world. As a result of that, in order to know whether or not an agent's credences are vindicated, one need only know the world that she inhabits. Thus, we say that an agent has a vindicated credence function iff it matches the omniscient credence function of the world that she inhabits—it does not matter the time at which she has the credence function.

However, suppose we were to consider agents with opinions about propositions that are not eternal: that is, propositions whose truth value may change within a given possible world. For instance, the proposition *Salt currently dissolves in water* is such a proposition; similarly, *A sea battle takes place tomorrow*. Both propositions may be true today and not tomorrow, or true 200 years ago but not 300 years ago. How then are we going to define omniscient credence functions? And how are we going to state Alethic Vindication? The natural move is this: An omniscient credence function is given relative to a world *and a time within that world*. An agent has a vindicated credence function iff it matches the omniscient credence function of the world *and time* that she inhabits when she has that credence function. Thus, if the propositions in $\mathscr{F}$ can change truth value over time, vindication is agreement with the omniscient credences *at the world and the time at which the credences are held*.

Contrast this with Ur-Chance Initial Vindication. Just as the truth values of non- eternal propositions can vary over time, so can the chances of even eternal propositions. The chance of *A sea battle takes place on 21 October 1805* may well change between 20 October 1805—when it may be less than 1—and 22 October 1805—by which time it may have become 1. Thus, the same question arises: When is an agent's initial credence function vindicated? According to Ur-Chance Initial Vindication, it is vindicated iff it matches the ur-chance function of the possible world that the agent inhabits. But the natural analogy with the alethic case says instead that it is vindicated iff it matches the chance function of the possible world *and time* at which the agent has that credence function—that is, the chance function of the possible world she inhabits at the moment at which she begins her epistemic life. Ur-Chance Initial Vindication, by contrast, is analogous to the following alethic account of vindication in the presence of non-eternal propositions: an agent's credence function is vindicated iff (p.112) it matches the omniscient credence function at the world she inhabits *and the earliest time in that world*. Thus, the vindicated initial credence for an agent to have in *A sea battle takes place tomorrow* is 1 if a sea battle took place the day after the earliest moment in the history of the world she inhabits; and 0 if one did not. But this seems unmotivated. What is so special about the earliest moment of a world that would mean that vindication is agreement with the omniscient credence function at that moment? But if it is unmotivated in the alethic case, it seems similarly unmotivated for one who agrees with Hájek that vindication is agreement with some chance function. Thus, analogously, we can ask: What is so special about the earliest moment in the history of the world our agent inhabits that would mean that vindication is agreement with the chance function at that moment, i.e., the ur-chance function? Why not say instead that vindication is agreement with the current chances? That is, why not adopt the following account of vindication?

> **Current Chance Initial Vindication** The vindicated initial credence function at a world and a time is the chance function at that world at that time.

In fact, there is a possible answer to this question that might be offered by the defender of Ur-Chance Initial Vindication: One might think that chance is valuable as what Ned Hall calls an *analyst expert* (Hall, 1994). Hall distinguishes two ways in which an expert function might be valuable. On the one hand, it might encode a great deal of evidence about the world: in this case, we say that it is a *database expert*. An encyclopaedia or an expert scientist or a University Challenge contestant might be a database expert in this sense. On the other hand, it might be very good at analysing evidence and assigning probabilities on the basis of that evidence: in this case, we say that it is an *analyst expert*. A clear-headed, careful thinker might be an analyst expert in this sense. Now suppose you are at the beginning of your epistemic life. You have not acquired any evidence. However, you would like to have a credence function that is very good at assessing evidence and reacting to it appropriately. Thus, you want your initial credence function to be a good analyst. In that case, it might seem that your initial credence function is vindicated just in case it matches the ur-chance function of the world, since that function is considered the ultimate analyst expert: like you, it lacks any evidence about the world; but it is very good at analysing any evidence that comes in. Thus, Ur-Chance Initial Vindication.

It may be correct to say that the chances are valuable as analyst experts. But whether or not that is so, it is a mistake to think that they are not also valuable as database experts. After all, the chances at any given moment encode complete information about the history of the world up to that moment. Thus, if we take vindication for an initial credence function to be agreement with the ur-chance function, then we deny the value of all the extra information that the current chance function holds. Thus, I agree with Caie that Current Chance Initial Vindication is superior to Ur-Chance Initial Vindication. This gives the following alternative to Brier Ur-Chance Initial Accuracy:

(p.113) **Brier Current Chance Initial Accuracy** The inaccuracy of an initial credence function held at a world and a time is the squared Euclidean distance from the chance function at that world and time to the credence function.

9.2 Introducing the Temporal Principal Principle

In this section, we explore the consequences of substituting Current Chance Initial Vindication in place of Ur-Chance Initial Vindication in the argument for the Principal Principle given above. To state the chance-credence principle that we thereby justify, we must introduce new terminology.

Suppose ch is a probability function that is amongst the possible chance functions at the time of the agent's initial credence function. Then we let T_{ch} be the proposition that says that the current chances are given by ch. Thus, T_{ch} is true at a world w and a time t if ch is the chance function at w and t; that is, $ch = ch_{tw} = ch_w(-|H_{tw})$. Note that this proposition is non-eternal. That is, its truth value may change from time to time within a given world: thus, it might be true at one time—when the chances are indeed given by ch—and false at another—when they are not. We call a proposition of this form a *current chance hypothesis*. As before, of course, there will typically be uncountably many possible current chance functions, and therefore uncountably many possible current chance hypotheses. But, also as before, we restrict attention to agents who have opinions only about finitely many of them.

With this in hand, we are ready to state the Temporal Principal Principle (Caie, 2015).

TPP$_0$ If an agent has an initial credence function c_0 defined on $\mathscr{F}$, then rationality requires that

$$c_0(X \mid T_{ch}) = ch(X)$$

for all propositions X in $\mathscr{F}$ and all possible ur-chance functions ch such that T_{ch} is in $\mathscr{F}$ and $c_0(T_{ch}) > 0$.

Thus, an agent is irrational at the beginning of her epistemic life unless her credence in each proposition conditional on ch giving the current chances is equal to the probability that ch assigns to that proposition.

Now, notice that we have introduced non-eternal propositions into the set $\mathscr{F}$—we had to do this in order to state TPP$_0$. And note that we have assumed in the statement of TPP$_0$ that each possible chance function is defined for each proposition in $\mathscr{F}$—this assumption is required to place the restriction on the conditional initial credences of our agent. Thus, in particular, each is defined on each non-eternal proposition in $\mathscr{F}$. But that raises the question: what is the chance of a non-eternal proposition? Typically, we take chances to be defined on eternal propositions, such as *A sea battle takes place on 21 October 1805*, and not on non-eternal propositions, such as *A sea battle takes place tomorrow*. The reason is that the truth of the latter sort of proposition depends (p.114) in part on the current time and it seems odd to say that there is a well-defined chance that it is currently one time rather than another. Caie offers a solution to this problem. According to him, the chance at t of a non-eternal proposition X is the chance at t of the eternal proposition X^t that is true at a world w iff X is true at w at t. Thus, for instance:

$$(A \text{ sea battle takes place now})^t \equiv A \text{ sea battle takes place at } t.$$

So:

$$(T_{ch})^t \equiv \text{The chance function at } t \text{ is } ch.$$

And therefore, the chance on 20 October 1805 of *A sea battle takes place today* is the chance on that day of *A sea battle takes place on 20 October 1805*. And the chance at t of T_{ch} is the chance at t of *The chance function at t is ch*.

This seems right to me. Essentially, the chance at t of some proposition that includes a temporal indexical, such as 'now', 'today', 'tomorrow', is just the chance of the proposition that results from rigidifying that indexical so that it refers to precisely the time that it refers to when the proposition is evaluated at t.

It is easy to show that, if p is a probability function over the algebra $\mathcal{F}_{\text{et}} \subseteq \mathcal{F}$ of eternal propositions in $\mathcal{F}$, and if we extend p to $\mathcal{F}$ using Caie's rigidification trick, then p is a probability function on $\mathcal{F}$.

Now we have the following theorem, due to Caie:

> **Theorem 9.2.1 (Caie)** *Let C be the set of possible current chance functions at the time of the agent's initial credence function. Then, if c is in* $cl(\mathcal{C}^+)$*, then c satisfies* TPP_0.

(This is a corollary of Theorem III.D.2. It follows if we let $\mathcal{S} = \mathcal{C}$ in that theorem.) Note: when we talk of a possible current chance function, we mean a probability function *ch* such that it is compatible with the agent's evidence that *ch* is currently the chance function. Now, the agent's evidence may not be strong enough to pin down a particular time as the present time. Thus, the set $\mathcal{C}$ of the possible chance functions may contain two chance functions *ch, ch′* that the agent knows could only be the chance functions at t and t' respectively. But the agent may not know whether the time is t or t' (or neither), so both are included in $\mathscr{C}$. Thus, if you know a fair coin will be tossed at 7:30 a.m., and you awake not knowing whether it is before or after 7:30 a.m., then there is a possible current chance function for you on which it is currently before 7:30 a.m. and the chance of heads is 50%, and there is a possible current chance function for you on which it is currently after 7:30 a.m. and the chance of heads is 1, and one on which it is 0.

Together with Bregman Current Chance Initial Accuracy and Theorems I.D.5 and I.D.8, the theorem just stated (Theorem 9.2.1) gives us an argument for TPP_0 Thus, consider an agent at the beginning of her epistemic life. Her credence function is her initial credence function. Suppose she violates TPP_0. That is, her credence in some proposition conditional on the current chances being given by *ch* is different from (p.115) the probability that *ch* assigns to that proposition. Then, according to the accuracy argument at which we have arrived, she is irrational because there is an alternative credence function that is guaranteed to be closer to the true current chance function than hers is; moreover, that alternative credence function expects itself to be closer to the true current chance function than it expects any other credence function to be.

9.3 Beyond the initial credence function

Both PP_0 and TPP_0 apply to an agent only at the beginning of her epistemic life: that is, both principles govern only her initial credence function. Yet chance-credence principles seem to govern us at later points in our epistemic lives as well. If tomorrow I encounter a coin that I know to be fair, it is surely just as irrational for me to assign a credence of 0.6 to it landing heads now as it would have been for me to do this at the beginning of my epistemic life. Thus, we should seek chance-credence principles that make this so. In (Pettigrew, 2012), I extended PP_0 as follows:

Extended Principal Principle (PP) If an agent has a credence function *c* and total evidence *E*, then rationality requires that

$$c(X \mid c_{ch}) = ch(X \mid E)$$

for all propositions *X* in $\mathscr{F}$ and all possible ur-chance functions *ch* such that C_{ch} is in $\mathscr{F}$ and $c(C_{ch}) > 0$.

That is, at any point in her epistemic life, an agent's credence in a proposition conditional on the ur-chances being given by *ch* must match the chance assigned to that proposition by *ch* once it has been brought up to speed with the agent's total evidence.

Having introduced this extended version of the Principal Principle, I defended it by extending Ur-Chance Initial Vindication as follows:

> **Ur-Chance Evidential Vindication** The vindicated credence function at a world and a time is the ur-chance function at that world conditional on the agent's total evidence at the time.

Substituting this for Ur-Chance Initial Vindication entails PP.

Caie rejects Ur-Chance Evidential Vindication. He worries that it requires a sense of 'ought' that is caught uncomfortably between the subjective sense and the objective sense. Take Alethic Vindication or Ur-Chance Initial Vindication. Both of these entail propositions that are stated using the objective sense of 'ought': an agent's credence function ought to match the omniscient credence function at the world she inhabits; an agent's initial credence function ought to match the ur-chance function at the world she inhabits. They are objective because they make the same demand of any two agents who inhabit the same world, even if those agents have different total evidence. On the other hand, take an evidentialist notion of vindication: An agent's credence function (p.116) is vindicated if it respects her total evidence, where we suppose that an agent has unimpeded access to whether or not she respects her total evidence. This entails a proposition that is stated using the subjective sense of 'ought': An agent's credence function ought to respect her total evidence. It is subjective because it makes the same demand of any two agents who share the same evidence, even if they inhabit different worlds. Ur-Chance Evidential Vindication entails a proposition that is stated using a third sense of 'ought', which lies somewhere between objective and subjective: An agent's credence function ought to match the current chances conditional on her total evidence. This is not objective, since it could make different demands on two agents who inhabit the same world at the same time but have different evidence. It is not subjective, since it could make different demands on two agents who share the same evidence but inhabit different worlds. Caie claims that there is no such sense of 'ought' and concludes that Ur-Chance Evidential Vindication is false.

I'm not sure I share Caie's discomfort with this sense of 'ought'. Consider an analogous sense that we might encounter when we are talking about a practical decision. We might wish to say that an agent ought to perform the act that has greatest utility by that agent's lights. For instance, I might say that Philip ought to go to Restaurant A rather than Restaurant B because Restaurant A serves pizza while Restaurant B doesn't, and pizza is Philip's favourite food. In this sentence, the 'ought' is not objective, since the act demanded is determined in part by the agent's utility function—which encodes Philip's preference for pizza—and that is subjective. But it is not subjective either, since the act is determined in part by how the world is—it is determined in part by which restaurant serves which food—and that is an objective fact. Thus, it is again the hybrid sense of 'ought' to which Caie objects. But it seems entirely unproblematic in this case. There seems to be no problem that arises from what you ought to do depending in part on your situation and in part on the world.

Nonetheless, I think Caie is right to reject Ur-Chance Evidential Vindication, even if he does so for the wrong reason. The concern is simply that to which we gave voice above. There is no satisfactory answer to the following question: What is so special about the initial moment in a world's history such that it is to the chance function at that moment, brought up to speed with my total evidence, that I should defer at any point in my epistemic life?

In response, Caie formulates an alternative account of vindication, which is purely objective, and which supports an argument in favour of a generalized version of TPP_0.

> **Current Chance Vindication** The vindicated credence function at a world and a time is the chance function at that world at that time.

Substituting this in place of Current Chance Initial Vindication gives an accuracy argument for the following chance-credence principle:

> **Extended Temporal Principle (TPP)** If an agent has a credence function c, then rationality requires that (p.117)
>
> $$c(X \mid T_{ch}) = ch(X)$$
>
> for all propositions X in $\mathcal{F}$ and all possible chance functions ch such that T_{ch} is in $\mathcal{F}$ and $c(T_{ch}) > 0$.

Thus, TPP is obtained from PP_0 in two steps: first, like PP, it applies to any point in an agent's life; second, the ur-chance hypotheses are replaced by current chance hypotheses.

The problem with TPP is that, while it is correct for nearly all situations we are ever likely to encounter, it is not generally true. Suppose you have been told the outcome of a future coin toss by a wizard with magical powers of prediction—he has told you that the coin will land heads (cf. (Lewis, 1980, 272–6)). Then you should be certain that it will land heads, it seems—that is, you should assign credence 1 to the future coin toss landing heads. So your credence in heads conditional on the current chance of heads being 50% will, likewise, be 1. However, TPP insists that it should be 0.5. The problem is that your evidence is inadmissible in Lewis' sense. Thus, to make TPP plausible, we must do one of two things: we might restrict TPP so that it applies only to agents with no inadmissible evidence; or we might amend TPP so that it makes the correct demands on all agents, even if they have inadmissible evidence.

On the first, we have:

> **Admissible Temporal Principle (ATP)** If an agent has a credence function c, then rationality requires that
>
> $$c(X \mid T_{ch}) = ch(X)$$
>
> for all propositions X in $\mathcal{F}$ *about which the agent has no inadmissible evidence*, and all possible chance functions ch such that T_{ch} is in $\mathcal{F}$ and $c(T_{ch}) > 0$.

The problems with this proposal are obvious. It is not fully general: it makes no demands on an agent's credences in propositions about which she has inadmissible evidence. It relies on a notion of inadmissibility for which it offers no analysis. And there is no obvious accuracy-based explanation for the differential treatment of agents with and agents without inadmissible evidence, at least if one follows Hájek and takes vindication to be agreement with the chances.

Here is the second proposal:

> **Evidential Temporal Principle (ETP)** If an agent has a credence function c *and total evidence E*, then rationality requires that
>
> $$c(X \mid T_{ch}) = ch(X|\mathrm{E})$$
>
> for all propositions X in $\mathcal{F}$, and all possible chance functions ch such that T_{ch} is in $\mathcal{F}$ and $c(T_{ch}) > 0$.

This has the advantage that it is fully general and does not appeal to an unanalysed notion of inadmissible evidence. Indeed, it provides an account of inadmissibility. A (p.118) body of total evidence E is admissible relative to the proposition X and a possible chance function ch iff $ch(X|E) = ch(X)$; that is, iff ch renders X and E stochastically independent. It also makes the intuitively correct demands on agents. Moreover, there is an accuracy-based argument in its favour. It is based on the following account of vindication:

Current Chance Evidential Vindication The vindicated credence function at a world and a time is the chance function at that world at that time conditional on the agent's total evidence at that time.

And it requires the following theorem, which is analogous to De Finetti's Characterization Theorem:

Theorem 9.3.1 *Let $\mathscr{C}$ be the set of possible current chance functions; let E be the agent's total current evidence; and let* $\mathscr{C}_E := \{ch(-|E) : ch \in \mathscr{C}\}$, *as above. Then, if c is in* $\mathrm{cl}(\mathrm{C}^+_E)$, *then c satisfies ETP.*

This says the following: take all of the possible current chance functions; conditionalize each on evidence E; take their convex hull and then take its closure; every credence function in that set satisfies ETP. (It is a corollary of Theorem III.A.2. It follows if we let $\mathscr{S} = \mathscr{C}_E$ in that theorem.)

Thus, we have:

(I_{ETP}) **Veritism** The sole fundamental source of epistemic value is accuracy.
(II_{ETP}) **Brier Current Chance Evidential Accuracy** The inaccuracy of an agent's credence function at a world is the squared Euclidean distance from the current chance function at that world conditional on the agent's total evidence to the agent's credence function.
This is a consequence of the following axioms:
(i) **Perfectionism** The accuracy of a credence function at a world is its proximity to the vindicated credence function at that world.
(ii) **Squared Euclidean Distance** Distance between credence functions is measured by squared Euclidean distance.
(iii) **Current Chance Evidential Vindication** The vindicated credence function at a world and a time is the chance function at that world at that time conditional on the agent's total evidence at that time.

(III_{ETP}) **Immodest Dominance**
(IV_{ETP}) **Theorems I.D.5** and **I.D.8** and **I.B.2** and **9.3.1**
Therefore,
(V_{ETP}) **Probabilism + ETP**.

Of course, Current Chance Evidential Vindication falls foul of Caie's objection to Ur-Chance Evidential Vindication: it involves a sense of 'ought' that lies uncomfortably between the subjective and the objective notion. But as we saw above that is not a (p.119) serious problem. In the next section, we will formulate a more powerful objection against Current Chance Evidential Vindication.

9.4 An objection to Current Chance Evidential Vindication

In this section, I'd like to raise an objection against Current Chance Evidential Vindication. So far in this chapter, we have raised detailed objections against accounts of vindication on the grounds that they are the wrong way to make precise Alan Hájek's claim that vindication for credences is agreement with the objective chances. In this section, I'd like to raise an objection against Hájek's claim itself, however it is made precise. This will motivate our second accuracy-based argument for chance-credence principles, which will occupy us in the next chapter.

Why think with Hájek that, in just the way that a belief in a proposition is vindicated if the proposition is true and a disbelief in a proposition is vindicated if the proposition is false, a credence in a proposition is vindicated if it matches the actual chance of that proposition? More precisely: Why think that there is *any* credence function and *any* time such that vindication for that credence function is agreement with the actual chance function at that time? More carefully still: Why think there is any credence function and any time *at which there are still chancy events left open* such that vindication for that credence function is agreement with the chance function at that time? After all, at a time at which no chancy events are left open, the actual chance function at that time simply coincides with the actual omniscient credence function, since it assigns a chance of 1 to all truths and 0 to all falsehoods. So even the proponent of Alethic Vindication will accept that *some* chance functions provide the correct standard for vindication, namely, the ones that agree with the relevant omniscient credence functions. So let us restrict our attention to chance functions that leave some events still chancy. Why think that such a chance function provides the function that an agent's credence function must match in order to be vindicated?

To answer this question, let us divide up accounts of the metaphysics of chance rather crudely into two camps: reductionist accounts and non-reductionist accounts. On a reductionist account, chance facts reduce to non-modal facts; on a non- reductionist account, they don't.

9.4.1 Reductionist accounts of chance

Let's consider reductionist accounts first: the two standard accounts in this camp are actual frequentism (Venn, 1876) and the best-system analysis (Lewis, 1994).[4] (p.120) According to actual frequentism, the chance of a proposition that reports an event occurring is simply the frequency with which events of the same kind actually occur. Thus, the chance that a particular coin lands heads on a given toss is the frequency with which that very coin actually lands heads when tossed—that is, the proportion of actual tosses on which it lands heads. The best-system analysis, on the other hand, is a little more sophisticated. According to that account of the metaphysics of chance, the chance of a proposition is whatever the best theory of the actual world says it is. Theories, on this account, are evaluated for strength, simplicity, and fit, where the fit of a theory is simply the chance that it assigns to the world unfolding as it actually does. Including chances in your theory of the world can greatly increase its simplicity. They serve as simple summaries of complicated and varied patterns of particular matters of fact. They avoid the need to posit laws that entail each of those particular matters of fact individually, which would give rise to a very complicated theory. Of course, positing chances also reduces the fit of the theory. A theory whose laws simply entail each of the particular matters of fact has maximal fit, since it assigns a chance of 1 to every matter of fact in the actual world and thus a chance of 1 to the world unfolding as it actually does. Positing non-extremal chances will not have maximal fit. But often, the best-system analysis claims, that loss of fit is outweighed by the gain in simplicity.

I will not consider the standard objections to these accounts (Hájek, 1997). Rather, I wish to ask: Do any of these accounts support a chance-based account of vindication? The answer, I think, is no. What is distinctive about these accounts is that, on both, the chance facts are *summaries* of the non-modal facts. What's more, these summaries do not preserve information—that is, it is not possible to recover the underlying non- modal facts from the summary of them that chance provides; there are different ways that the non-modal facts might be that will be summarized in the same way by chance facts. For instance, on the frequency account, if the chances assign a probability of $\frac{1}{2}$ to a coin landing heads, that summarizes the actual sequence of heads and tails that the coin produces when tossed: it summarizes that sequence by telling us that half of the tosses in it come up heads. But of course, there are many sequences that have this feature. So the summary contains less information than would be contained in a comprehensive description of the actual sequence of tosses. And something similar holds for the best-system analysis. This raises the following question: suppose the coin is to be tossed ten times. For each coin toss, I have a credence in the proposition that it will land heads on that toss. That is, the following propositions are amongst those to which I assign a credence: *Coin will land heads on toss 1*, …, *Coin will land heads on toss 10*. Why would we value having credence $\frac{1}{2}$ in each of these propositions more (p.121) highly than having credence 1 in each proposition that is true and 0 in each proposition that is false? Why would we value a state that matches a summary that contains less than maximal information more than we value the omniscient state, which contains maximal information?

One answer might be that we value simpler doxastic states more than we value complex ones. While the omniscient credence function contains more information than a credence function that matches the chances, it is in some sense less simple— the credence function that matches the chances takes the same attitude to the outcome of each coin toss; the omniscient credence function takes different attitudes and those attitudes lack any discernible pattern. Extending this claim, we might say that we value doxastic states in much the same way that we value scientific theories; and we might adopt the account of value for scientific theories offered by Lewis' best-system analysis. Thus, the value of a credal state at a world is given by a combination of its strength, its simplicity, and its fit to that world (which is an increasing function of the credence that it assigns to that world being actual). If we assume this, something like Hájek's thesis seems to follow from Lewis' best-system analysis: if the best system or theory is determined by the same factors that determine the most valuable doxastic state, and if the chances at a world and a time are the chances postulated by the best-system analysis, then the most valuable doxastic state is the one that embodies the best system and thus includes a credal state that embodies the chance function postulated by that system, which is, by hypothesis, the true chance function at the world and time in question.

Now, it may initially seem plausible that the virtues of a scientific theory are equally virtues of a doxastic state and vice versa. After all, we are talking about those virtues of a scientific theory that might lead us to accept it; and acceptance is a doxastic pro- attitude. However, while I might be happy to say that a scientific theory with those virtues is one to which I should assign a high credence, I am not happy to say that it is one that my doxastic state should *reflect* or *embody* in the sense that I should adopt the chance function at a time that it postulates as my credence function at that time. The strength, simplicity, and fit of a scientific theory may give you reason to assign a high credence to the whole package, including the chance hypothesis that the theory entails. But that is different from saying that these virtues are the ones we wish our credence functions themselves to have.

Perhaps there is another reason to think that simplicity is a virtue of credal states. We are bounded creatures with a limited capacity to adopt highly complex doxastic states. The simpler a credal state, the more easily we will be able to update it quickly and reason from it effectively While this may be true, it is not relevant to the discussion here. In this book, we seek purely epistemic arguments for principles of rationality. There may be a range of practical reasons for violating these principles because our overall utility function incorporates many other factors besides the purely epistemic ones—indeed, the vast literature on heuristics and biases suggests that our evolutionary history did indeed include strong practical reasons for deviating from (p.122) the principles of rationality that we attempt to justify here. But that does not derail our arguments—we are interested here in what is required of an agent without such limitations.

So: on a reductive account of objective chance, there seems to be no good reason to hold that there is ever a situation in which a credence function is vindicated iff it matches a non-trivial objective chance function—that is, a chance function on which some events are still chancy. It is always better to have maximal credence in truths—including true chance hypotheses—and minimal credence in falsehoods—including false chance hypotheses—than to have credences that match the information-losing summaries of the patterns of truths and falsehoods that the chances provide.

9.4.2 Non-reductionist accounts of chance

Let us turn now to non-reductive accounts of chance. On these accounts, ur-chance hypotheses describe irreducibly modal features of the world. For instance, on a propensity account of chances, the chance of an event measures the strength of the disposition for that event to occur—we might think of this as the disposition that the world has to contain that event. If this is correct, then, according to Hájek’s thesis, your initial credence function is vindicated iff your credence in each proposition matches the strength of the world’s disposition to make that proposition true. However, the same question arises here as in the previous section: Why would you value a credence function that matches the chances more than you value one that matches the omniscient credence function? The omniscient credence function assigns 1 to the true chance hypotheses: so it seems that it contains all the information that the chance function contains, as well as much more detailed information about whether or not particular chancy events in fact occur, irrespective of what was the strength of their disposition to occur. Put another way: it seems valuable to represent both the strength of the disposition of an event to occur and whether or not it did really occur—the omniscient credence function represents both; the chance function only represents the former.

Thus, as in the case of reductive accounts, we have found no reason to say that there is any credence function and any time at which there remain undetermined chancy events such that that credence function is vindicated iff it matches the chances at that time. In particular, we must abandon Hájek’s thesis and, with it, Current Chance Evidential Vindication and our first argument for ETP. In the next chapter, we will give a second, more promising argument for ETP.

Notes:

([1]) In fact, in (Pettigrew, 2012), I appealed to conditions on distance measures that are much closer to those given by Joyce (1998). I am less confident of those conditions now, partly for the reasons given above. Moreover, they support an argument for PP_0 only in the presence of Undominated Dominance, which, we have seen, is too strong (Section 2.2). In any case, my reasons for abandoning this accuracy-based argument for PP_0 are orthogonal to the choice of distance measure. As we will see, my concern is with Ur-Chance Initial Vindication.

[2] In (Pettigrew, 2012) and (Pettigrew, 2013), I assumed that there are only finitely many possible ur- chance functions and that our agent has opinions about all of them. In the presence of that assumption, it is possible to show that the closure of the convex hull of $\mathscr{C}$ is *precisely* the set of probability functions that satisfy PP_0. But this assumption is clearly too restrictive, as Caie (2015) points out. Fortunately, the argument can easily be adapted to allow infinitely many possible ur-chance functions, providing the agent only has opinions about finitely many of them. I will present the adapted argument here. Note one difference between the adapted argument I present and Caie's adaptation. Caie assumes that the set of possible ur- chance functions is closed. This seems implausible to me. It seems quite possible that there is some event such that, between them, the possible ur-chance functions assign to this event all and only the possible probability values greater than $\frac{1}{2}$. If that's so, the set of possible ur-chance functions is not closed, since there are infinite sequences of possible ur-chance functions whose limit assigns a probability of $\frac{1}{2}$ to that event; but, by hypothesis, the probability function that is that limit isn't amongst the possible ur-chance functions. As we see in Theorem 9.0.9, the assumption of closure is not necessary. So we do without it here.

[3] Recall: if $\mathscr{X}$ is a set of credence functions, $\mathscr{X}^+$ is the convex hull of $\mathscr{X}$ and $\mathrm{cl}(\mathscr{X}^+)$ is the closure of that convex hull.

[4] I am considering only non-reductionist accounts of chance on which the chances are still mind- independent features of the world. I do not consider expressivist accounts of chance on which the sentence 'The chance of X is 0.9' merely expresses the agent's credence of 0.9 in X. It is pretty clear that, if this is the correct semantics for chance talk, the chances cannot provide a probative notion of vindication—if the true chances for a particular agent simply are her credences, then vindication comes too easily, and it cannot ground any interesting principles of rationality. This possibility has been explored in unpublished work separately by Wo Schwarz and Cian Dorr.