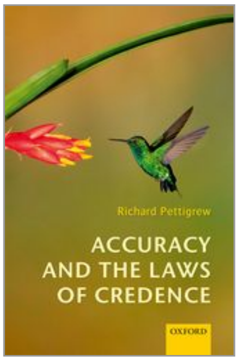
University Press Scholarship Online
Oxford Scholarship Online



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

Self-undermining chances

Richard Pettigrew

DOI:10.1093/acprof:oso/9780198732716.003.0012

Abstract and Keywords

In the previous chapters in this part of the book, it has been assumed that the objective chances are not self-undermining—that is, they assign maximal probability to the hypothesis that says that they themselves are the chances. This chapter explores what happens to the arguments when this assumption is dropped. The chapter discusses the New Principle as well as Jenann Ismael's General Recipe.

Keywords: Undermining futures, Self-undermining chances, chance, New Principle, Jenann Ismael

So far in this part of the book, we have assumed that while our agent may be uncertain which is the true ur-chance or current chance function, she is nonetheless certain that the true ur-chance or current chance function is what we called *non-self-undermining*. That is, we have assumed that every possible ur-chance function is certain that it is the ur-chance function, and every possible current chance function is certain that it is the current chance function. In symbols, we have assumed that, for all ch in $\mathscr{C}_0$, $ch(C_{ch}) = 1$ and, for all ch in $\mathscr{C}$, $ch(T_{ch}) = 1$. In this chapter, we ask what happens to our chance-credence principles and our arguments in their favour if we drop this assumption. We begin by asking what might lead us to drop this assumption. Then we consider the consequences for the arguments we gave above, and the principles we have been using those arguments to try to justify.

11.1 Self-undermining chance functions

Recall: we say that a possible current chance function *ch* is *self-undermining in the presence of evidence E* if $ch(T_{ch}|E) < 1$, where, as above, T_{ch} says that *ch* is the current chance function, and the chance at a time *t* of a non-eternal proposition *X* is just the chance at *t* of X^t, the eternal proposition that is true at a world just in case *X* is true at that world at *t*.

Now consider a particular version of actual frequentism about chances. On this account, we begin by assigning chances to basic events on the basis of frequencies: in particular, the chance of a particular basic event occurring is the frequency with which events of that type occur amongst all events in some relevant broader type. For instance, the chance of a coin landing heads on its third toss is simply the frequency of heads amongst all tosses of the coin. Having assigned chances to basic events in this way, we go on to assign chances to more complex events by using causal dependence information and the laws of probability. For instance, suppose that the causal relations in the world are such that the tosses of this coin are independent of one another. Then the chance of the complex event that consists of the first toss landing heads and the second toss landing tails, for instance, is simply the product of the chance of the first landing heads with the chance of the second landing tails.

(p.134) Now, suppose we are considering a particular coin. And suppose it is determined that it will be tossed exactly four times, and that the outcomes will be independent; and suppose it is currently prior to the first toss. Thus, there are sixteen possible futures: *HHHH, HHHT, HHTH,…, TTTT*. Moreover, each possible future determines a current chance hypothesis. Thus, *HHHH* determines that the current chance of heads is 1, since 100% of all tosses land heads on this outcome. Similarly, *HHTH* and *HTHH* and *THHH* and *HHHT* all determine that the current chance of heads if $\frac{3}{4}$, since 75% of all tosses land heads on these outcomes. And so on. Thus, we have the following relationship between the possible outcomes of the coin toss and the current chance hypotheses. First, the possible current chance functions are as follows, where H_i says that the coin lands heads on the *i*th toss: for *i* = 1,2,3,4,

$$ch_0(H_i) = 0 \quad ch_1(H_i) = \frac{1}{4} \quad ch_2(H_i) = \frac{1}{2} \quad ch_3(H_i) = \frac{3}{4} \quad ch_4(H_i) = 1$$

And, for each $k = 0, \dots 4$, and each $1 \leq i, j \leq 4$,

$$ch_k(H_i \ \& \ H_j) = ch_k(H_i) \times ch_k(H_j)$$

since the outcome of a particular toss is independent of the outcomes of all the other tosses. Thus, the possible current chance hypotheses are: $T_{ch_0}, \ldots, T_{ch_4}$. And they are related to the possible outcomes as follows:

$$\begin{array}{lcl} T_{ch_0} & \equiv & TTTT \\ T_{ch_1} & \equiv & TTTH \vee TTHT \vee THTT \vee HTTT \\ T_{ch_2} & \equiv & TTHH \vee THTH \vee THHT \vee HHTT \vee HTHT \vee HTTH \\ T_{ch_3} & \equiv & HHHT \vee HHTH \vee HTHH \vee THHH \\ T_{ch_4} & \equiv & HHHH \end{array}$$

Now, we note first that the ur-chance functions ch_0 and ch_4 are not self-undermining:

$$ch_4(C_{ch_4}) = ch_4(HHHH) = ch_4(H_1)ch_4(H_2)ch_4(H_3)ch_4(H_4) = 1 \times 1 \times 1 \times 1 = 1$$

And similarly for ch_0.

$$ch_0(C_{ch_0}) = ch_0(TTTT) = ch_0(T_1)ch_0(T_2)ch_0(T_3)ch_0(T_4) = 1 \times 1 \times 1 \times 1 = 1$$

However, the other three possible ur-chance functions are self-undermining. For instance,

$$\begin{array}{lcl} ch_1(T_{ch_1}) & = & ch_1(TTTH \vee TTHT \vee THTT \vee HTTT) \\ & = & ch_1(TTTH) + ch_1(TTHT) + ch_1(THTT) + ch_1(HTTT) \\ & = & 4 \times ch_1(T_1) \times ch_1(T_2) \times ch_1(T_3) \times ch_1(H_4) \\ & = & 4 \times \frac{3}{4} \times \frac{3}{4} \times \frac{3}{4} \times \frac{1}{4} = \frac{27}{64} < 1 \end{array}$$

(p.135) So this version of actual frequentism gives rise to self-undermining chance functions. The source of the problem is that the current chance hypotheses are equivalent to hypotheses about the pattern of heads and tails that in fact eventuate from the four tosses of the coin. But these latter facts—and therefore also the former facts—are amongst those that are chancy prior to any tosses. Thus, some current chance functions assign a non-zero chance to a sequence of tosses that determines a current chance hypothesis other than the one corresponding to that current chance function. That is, they assign non-zero chance to some other current chance function being the true current chance function. Thus, they assign less than maximal chance to themselves being the true current chance function. And that is precisely what it means to be selfundermining.

Of course, actual frequentism—and, in particular, this version of it—is a hopeless account of chance, as Alan Hájek has documented (Hájek, 1997). Thus, one might think that we can simply avoid the problem of self-undermining chance functions by moving to a more plausible account of chance. But, while the problem of selfundermining chance functions is most easily and dramatically illustrated using the actual frequentist account, it in fact affects many reductionist accounts of chance. Most notably, it affects David Lewis' best-system analysis of chance, as Lewis himself saw clearly. After all, on that account, just as on the actual frequentist account, there are nearly always ways the non-modal facts might be such that the best system for those facts includes a chance function that assigns non-zero probability to some other way the non-modal facts might be such that the best system that accounts for those alternative facts posits a different chance function. Indeed, in the very simple world that includes only the four coin tosses considered above together with physical facts that make them independent, the best-system analysis is likely to coincide with the actual frequentist account, and so both will give rise to the same self-undermining possible chance functions.

Now, it is worth noting that not all reductive accounts of chance give rise to selfundermining chances (Schaffer, 2003). Here is one that doesn't. It proceeds exactly like the actual frequentist account until the last step, where, instead of taking ch_0, ch_1, …, ch_4 to be the possible chance functions in the simple world that just contains the four coin tosses, it takes the possible chance functions to be the following:

$$ch_0^*(-) := ch_0(-|TTTT)$$
$$ch_1^*(-) := ch_1(-|TTTH \vee TTHT \vee THTT \vee HTTT)$$
$$ch_2^*(-) := ch_2(-|TTHH \vee THTH \vee THHT \vee HHTT \vee HTHT \vee HTTH)$$
$$ch_3^*(-) := ch_3(-|HHHT \vee HHTH \vee HTHH \vee THHH)$$
$$ch_4^*(-) := ch_4(-|HHHH)$$

And it declares ch_k^* to be the chance function in precisely those situations in which the actual frequentist declares ch_k to be the chance function. Thus, ch_0^* is the chance function in worlds in which there are no heads, ch_1^* is the chance function in worlds (p.136) in which there is exactly one head, and so on.[1] So, on this account, the chance of an event is whatever the chance of that event would be according to the actual frequentist *conditional on the proposition that declares the actual frequentist's chance function the correct one according to her account of chance*. Now, clearly each ch_i^* is not selfundermining. For instance, $ch_1^*(C_{ch_1^*}) = ch_1^*(TTTH \vee TTHT \vee THTT \vee HTTT) = 1$. Indeed, this trick easily turns any reductive account that gives rise to self-undermining chance functions into one that doesn't. Given an account of chance, if we take the possible chance functions according to that account and apply this trick, we call the resulting probability functions the possible chance* functions. But notice that, on this account, the coin tosses are no longer independent (Arntzenius & Hall, 2003). The chance of a final (fourth) heads conditional on three previous tails is 1 according to ch_1^*, whereas its unconditional probability is $\frac{1}{4}$. And indeed something similar will happen for any reductionist attempt to avoid self-undermining chance functions. Non-selfundermining chance functions must assign zero probability to each future on which it is not the chance function.

But those futures will most often be composed of individual non-modal events to each of which the non-self-undermining chance function assigns non-zero unconditional probability. Thus, to ensure that the total future receives zero probability, the chance function must not render those events independent. Thus, it seems that any plausible reductive account of chance will give rise to self-undermining chance functions.

It is an interesting question, though one we will not be able to pursue here, whether there are sensible non-reductionist accounts of chance on which some possible chance functions are self-undermining. There is certainly no logical obstacle. But most nonreductionists take the chance facts to be determined at each point in the history of the universe. And thus, they would take the true chance of the true chance hypothesis to simply be 1.

As we saw above, self-undermining chance functions give rise to strange consequences when we apply the chance-credence norms that we have been considering so far in this part of the book. Witness, for instance, the following theorem, which we mentioned above:

> **Theorem 11.1.1** *Suppose ch is a self-undermining possible current chance function relative to the agent's evidence—that is,* $ch(T_{ch}|E) < 1$. *Then, if our agent has credence function c,* T_{ch} *is in* $\mathscr{F}$*, and c satisfies ETP, then* $c(T_{ch}) = 0$.

That is, if an agent considers any self-undermining chance function to be a possibility, then she must assign it credence 0. Consider, for instance, the possible chance functions enumerated by the actual frequentist above, namely, $ch_0,\ldots, ch_4$. Any agent who takes those to be the possibilities, and who has acquired no evidence, is required on pain of irrationality to assign credence 0 to T_{ch_1}, T_{ch_2}, and T_{ch_3}. That is, she is required to be certain that either ch_0 or ch_4 is the true chance function. So she must be (p.137) certain from the start that the coin is not chancy at all, but deterministic, since ch_0 and ch_4 both assign extremal probabilities to the outcomes of the coin tosses, and that is just what it means for a process to be deterministic in a finite event space. Equivalently, she must be certain that the coin will either land only heads or land only tails. For those like Lewis who defend Regularity—the principle that only propositions incompatible with one's evidence may rationally receive credence 0—this is a disaster.

How should we respond to this if we wish to retain an account of chance that gives rise to self-undermining chance functions? One possibility is what Michael Titelbaum and I have dubbed *the restricted-*

scope response (Pettigrew & Titelbaum, 2014).[2] The idea is this. According to ETP, an agent should defer to the current chance functions when setting her credence in each proposition about which she has an opinion, *even those propositions that concern the chances themselves*. This fact is crucial in the proof of Theorem 11.1.1. But, the restricted-scope response goes, while self-undermining chance functions are worthy of deference when we are setting credences in the outcome of a particular coin toss, they are not worthy of deference when we are setting credences in chance hypotheses themselves. There may be different reasons for this restriction. Ismael, for instance, suggests that chance functions are not even defined on hypotheses about the chances (Ismael, 2008, 301). If that's the case, it is certainly not possible to defer to those chance functions when you are setting your credences in those hypotheses. Alternatively, you might simply think that the fact that an expert is worthy of deference with respect to some subject matter—coin tosses, for instance— does not entail that they are experts about their own expertise. I may defer to a weather forecaster on the matter of the weather, but not on the matter of her own expertise. Thus, our chance-credence principles should not demand deference on anything other than those propositions that it is the core business of chance functions to evaluate, namely, the particular matters of non-modal fact. Whichever is your motivation for the restricted-scope response, the result is the same: according to you, the correct version of ETP in fact restricts the requirement of deference to some subset of the propositions about which the agent has an opinion—in particular, it restricts to a subset that does not include the chance hypotheses. Thus, given a subset $\mathscr{F}' \subseteq \mathscr{F}$ such that no current chance hypothesis T_{ch} is in $\mathscr{F}'$, we end up with the following version of the Temporal Principle:[3]

(p.138) **Restricted Evidential Temporal Principle (ETP_)** If an agent has a credence function c and total evidence E, then rationality requires that

$$c(X|T_{ch}) = ch(X|E)$$

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

The problem with the restricted-scope response for the reductionist is that the exclusion of chance hypotheses is not well motivated. Let's consider Ismael's motivation first. For Ismael, the problem is that chance functions are simply not defined on chance hypotheses.

> [A] theory of chance that assigns probabilities to finite strings of future events will typically be silent on total histories, not because history is necessarily infinite, but because for any finite string of events, there is always the *possibility* of events to follow. Short of an assumption that the universe has a history of a finite, specified duration, there is no general way of turning a distribution over finite strings of future events into a distribution over total histories. And without an assignment of probabilities to total histories, we don't have an assignment of probabilities to the Humean truthmakers of theories of chance. (Ismael, 2008, 301)

Of course, in the toy example given above to illustrate the problem of self-undermining chance functions for actual frequentists, we did assume that the universe has a history of a finite, specified duration—we assumed that it contained only four tosses of the coin. It is only in the presence of that assumption that *TTTT* is equivalent to T_{ch_0}, and so on. In the absence of that assumption it is the following proposition that is equivalent to T_{ch_0}:

T & The universe contains one coin toss
∨ *TT* & The universe contains two coin tosses
∨ *TTT* & The universe contains three coin tosses
∨ …

And similarly for $T_{ch_1}, T_{ch_2}, \ldots$. But surely this proposition, like *TTTT* or *HTTT* ∨ *THTT* ∨ *TTHT* ∨ *TTTH*, is exactly the sort of proposition to which chance functions assign chances. After all, it is a (countable) disjunction of (finite) conjunctions of particular, non-modal events: the first disjunct is the conjunction of a proposition about the first coin toss and a proposition about the duration of the universe (or at least the number of coin tosses it contains); the second disjunct is the conjunction of a proposition about the first coin toss, a proposition about the second coin toss, and a different proposition about the duration of the universe; and so on. Each of these particular non-modal events is surely assigned a chance; and they might very well be independent of one another; and so the chance of each disjunct above might be defined; and that allows us to determine the chance of the disjunction (at least in the presence of countable additivity). But the disjunction is equivalent to a chance hypothesis, which is exactly the sort of proposition that Ismael hoped to exclude from (p.139) the domain of the chance function. So I don't think Ismael's version of the restricted-scope response will work for the reductionist: for the reductionist, chance hypotheses are disjunctions of total world histories; and total world histories are conjunctions of propositions about particular matters of non-modal fact, either concerning the outcome of a particular event, or concerning the number of events that the universe contains. These same considerations suggest not only that the chance functions are defined on chance hypotheses, but that it is part of the core business of chance functions to give verdicts about the chance hypotheses, given that these hypotheses are equivalent to disjunctions of conjunctions of propositions concerning particular matters of non-modal fact.

For these reasons, it seems to me that the restricted-scope response fails. In the light of this: what is the reductionist to do? I will consider two options. On the first, we simply run our second accuracy-based argument for ETP as it is, but dropping the assumption that each possible current chance function is non-self-undermining, and we see what principle it justifies. On the second, we run the second accuracy-based argument for ETP, but we replace Current Chance Evidential Dominance with an analogous principle that appeals not to the self-undermining chance functions but instead to the non-self-undermining chance* functions, which are defined in the way suggested above when we defined ch_i^* in terms of ch_i, and we see what chance-credence principle the resulting argument justifies. As we will see, the principles that are justified by these two options are close to those that have been suggested in the literature, namely, Ned Hall's and Michael Thau's New Principle, and Jenann Ismael's General Recipe (Hall, 1994; Thau, 1994; Ismael, 2008).

11.2 An accuracy-based argument for Ismael's General Recipe

Recall the structure of our accuracy-based argument for ETP. The accuracy-based argument for Probabilism begins with an account of epistemic value—namely, Veritism combined with Brier Alethic Accuracy—and combines that with a decision- theoretic principle—namely, Immodest Dominance. From that conjunction, we conclude that rationality requires that an agent's credence function should lie in the closed convex hull of the set of omniscient credence functions. And from that, together with de Finetti's proof that every member of the closed convex hull of the omniscient credence functions is a probability function, we obtain that rationality requires an agent to satisfy Probabilism. The accuracy-based argument for ETP, on the other hand, begins with the same account of epistemic value—again, Veritism together with Brier Alethic Accuracy—but combines it with a different decision-theoretic principle— namely, Current Chance Evidential Immodest Dominance. From that conjunction, we conclude that rationality requires that an agent's credence function should, on pain of irrationality, lie in the closed convex hull of the set of possible current chance functions (p.140) after they have been brought up to speed with the agent's total current evidence. And from that, together with our proof that every member of the closed convex hull of the current chance functions conditional on her total evidence satisfies ETP, we obtain that rationality requires adherence to that principle.

It is only in the final step of this argument that we appeal to the assumption that all possible chance functions are non-self-undermining relative to the agent's evidence: if we drop that assumption, it is simply not true that all members of the closed convex hull of the possible chance functions conditional on that evidence satisfy ETP. But the argument leading up to that final step remains unaffected. Thus, the argument establishes at least that rationality requires that an agent's credence function belong to the closed convex hull of the set of possible current chance functions. That is, it establishes ETP$^+$, which we met in Section 10.2 above:

> **Evidential Temporal Principle$^+$ (ETP$^+$)** If an agent has a credence function c and total evidence E, then rationality requires that c is in $\mathrm{cl}(C_E^+)$.

Note that, without further assumptions, ETP$^+$ and ETP are logically independent of one another. As we noted above, if (i) there are only finitely many possible current chance functions, (ii) our agent has credences about all of them, and (iii) all of them are non-self-undermining, then ETP and ETP$^+$ are equivalent. Moreover, if (i) and (ii) hold, but (iii) doesn't, then ETP is strictly stronger—that is, more demanding—than ETP$^+$. If none of (i), (ii), or (iii) hold, then neither is stronger than the other.

Note also that Cleo violates ETP$^+$, since her credence function lies outside the closed convex hull of the current chance functions that are epistemically possible for her: the epistemically possible current chance functions all assign chances between 60% and 70%, but she assigns a credence below 0.5.

Interestingly, Jenann Ismael has proposed a chance-credence principle that comes close to the principle we have just stated. She states it for the case in which there are only finitely many possible current chance functions; and she doesn't consider cases in which you have (inadmissible) evidence about the future. In our notation, and introducing the possibility of inadmissible evidence, the principle is as follows (Ismael, 2008, 298):

General Recipe (GR) If an agent has a credence function c and total evidence E, then rationality requires that

$$c(X) = \sum_{ch \in C} c(T_{ch}) ch(X|E)$$

for all X in $\mathscr{F}$.

In the presence of self-undermining chance functions (relative to E), GR is strictly stronger than ETP$^+$, since it imposes greater restrictions on the weights in the weighted sum. Indeed, it imposes the same restrictions on those weights that ETP imposes when there are finitely many possible chance functions and the agent has (p.141) opinions about all of them. But it is strictly weaker than ETP and does not suffer from same problem. Nonetheless, as we will see below, it does suffer from related problems.

First, however, let's see how it relates to our second accuracy-based argument. For the remainder of the section, we will assume that there are just finitely many possible current chance functions, since this is a presupposition of Ismael's principle GR. In the absence of further information about the nature of the possible current chance functions, the second accuracy-based argument establishes ETP$^+$ at most. However, if we know something more about each of the finitely many possible chance functions, then the second accuracy-based argument establishes GR. What we have to know is this: each of the finitely many possible current chance functions *expects itself to give the correct current chances*. If we know that, then ETP$^+$ entails GR.

Let's see how this works. First, recall that a possible current chance function ch is *non-self-undermining* if $ch(T_{ch}) = 1$. That is, a possible chance function is non-self-undermining if it is certain that it gives the true current chances. However, there are ways for a possible chance function to endorse the chances it provides that don't involve it being certain that it provides the true chances. For instance, a selfundermining possible chance function might expect itself to give the chances (in the presence of some body of evidence) even though it is not certain that it will (even in the presence of that evidence)—just as I might *expect* the weight of my Christmas cake to be 2.2 kg, even though I'm not *certain* that it will have that weight. We call such a possible chance function *expectationally non-self-undermining in the presence of E*. Thus, ch is expectationally non-self-undermining in the presence of E if

$$ch(X) = \sum_{ch \in C} ch(T_{ch}) ch(X|E)$$

Now, we have the following result:

Theorem 11.2.1 *If each possible current chance function is expectationally non-self#x2010;undermining in the presence ofE, then ETP$^+$ entails GR.*

The problem is that there's no good reason to think that the self-undermining possible current chance functions to which a reductive account will typically give rise are nonetheless expectationally non-self-undermining. For instance, the chance functions to which the actual frequentist account sketched above gives rise are demonstrably not.

$$ch_1(HHHH) = \frac{1}{256} \neq \frac{275}{8192} = \sum_{k=0}^{4} ch_1(C_{ch_k})ch_k(HHHH)$$

So, without an assumption about the possible current chance functions, we have no accuracy-based reason in favour of GR. And the assumption required is unmotivated.

What's more, as I noted in (Pettigrew, 2014c), GR faces some serious problems independent of the existence or otherwise of an accuracy-based justification. The most serious stems from the following theorem. Say that a possible current chance (p.142) function *tolerates non-self-undermining chances* (*in the presence of evidence E*) if it assigns positive probability to a current chance hypothesis on which the current chance function is non-self-undermining (in the presence of *E*). Thus, ch_1, ch_2, and ch_3 are all self-undermining, but tolerant of non-self-undermining chances in the absence of any evidence, since each assigns positive probability to T_{ch_0} and T_{ch_4}, both of which are non-self-undermining in the absence of any evidence.

Theorem 11.2.2 *Suppose ch is a possible current chance function; and suppose ch is selfundermining, but tolerant of non-self-undermining chances in the presence of E. Then, if credence function c satisfies GR, then* $c(T_{ch}) = 0$.

That is, just as the Temporal Principle demands that we assign no credence to chance hypotheses that posit self-undermining chance functions, so the General Recipe demands that we assign no credence to chance hypotheses that posit self-undermining chance functions *that tolerate non-self-undermining chances*. As we saw above, in the toy example for actual frequentism described above, this would again demand that we distribute our credence only over T_{ch_0} and T_{ch_4}. That is, we must be certain that the chances are extremal and that the coin will land only heads or only tails. As before, this is an unpalatable consequence, and it is disastrous for anyone who endorses Regularity as a principle of rationality.

So Ismael's General Recipe gives rise to unpalatable consequences in the presence of self-undermining chance functions that tolerate non-self-undermining. Moreover, we cannot give an accuracy argument in its favour without making an unmotivated assumption about the expectational non-self-undermining of the chance functions. I conclude that the reductionist is best advised not to adopt the General Recipe, but at most to endorse ETP^{+}. There is an accuracy argument for that latter principle; it is intuitively plausible; and it does not have the unwelcome consequences of PP or ETP in the presence of the sort of chance function to which a reductionist is typically committed.

11.3 An accuracy-based argument for the New Principle

So much for Ismael's General Recipe. In this section, we turn to another chance- credence principle that has been offered as an alternative to Lewis' problematic Principal Principle and its replacement—the Temporal Principle— for those who give a reductive account of chances. This is Michael Thau's and Ned Hall's so-called New Principle, which Lewis himself came to endorse, albeit tentatively and reluctantly (Lewis, 1994; Hall, 1994; Thau, 1994). I state it here in terms of current chance functions, and I include the possibility of inadmissible evidence.

> **New Temporal Principle (NTP)** If an agent has a credence function c and total evidence E, then rationality requires that
>
> $$c(X|T_{ch}) = ch(X|T_{ch} \,\&\, E)$$
>
> (p.143) for all propositions X in $\mathscr{F}$, and all possible chance functions ch such that T_{ch} is in $\mathscr{F}$ and $c(T_{ch}) > 0$.

Thus, according to this principle, when you set your credence in a proposition X conditional on a current chance hypothesis T_{ch}, you bring the corresponding current chance function ch up to speed with your evidence—as the most plausible version of the Temporal Principle also exhorts you to do—but you also bring it up to speed with the proposition conditional on which you are setting your credence in X, namely, T_{ch}. You then take as your credence the probability that it assigns to X once brought up to speed in this way.

It is easy to see that this avoids the problems that the Temporal Principle faces in the presence of self-undermining current chance functions. After all, the New Temporal Principle is essentially obtained from the original Temporal Principle as follows: instead of demanding that we defer to the possible current chance functions in the way the Temporal Principle makes precise, the New Temporal Principle demands that we defer to the possible current chance* functions in the way the Temporal Principle makes precise, where, if you recall, the chance* functions are the chance functions made non-self-undermining in the way described above. Thus, you can see the New Temporal Principle either as proposing a different way of deferring to the current chance functions, or as endorsing the way of deferring codified in the original Temporal Principle, but proposing a different target for the deference, namely, the chance* functions, rather than the chance functions.

This suggests a way in which we might adapt the accuracy-based argument for the Temporal Principle in the presence of non-self-undermining chance functions to produce an accuracy-based argument for the New Temporal Principle in the presence of self-undermining chance functions. The key is to replace the reference to possible current chance functions in Current Chance Evidential Undominated Dominance and its epistemological cousin Current Chance Evidential Immodest Dominance with reference to possible current chance* functions. Thus, we have:

> **Current Chance* Evidential Undominated Dominance** Suppose $\mathcal{O}$ is the set of options, $\mathcal{W}$ is the set of possible worlds, and $\mathfrak{A}$ is a utility function. Suppose *o, o′* in $\mathcal{O}$. Then, if
>
> (i) *o** strongly current chance* $\mathfrak{A}$-dominates *o* relative to the agent's current total evidence *E*, and
> (ii) there is no *o′* that weakly current chance* $\mathfrak{A}$-dominates *o* relative to *E* then
> (iii) *o* is irrational for any agent with utility function $\mathfrak{A}$.

And we have:

> **Current Chance* Evidential Immodest Dominance** Suppose $\mathfrak{I}$ is a legitimate measure of inaccuracy and *E* is a proposition. Then, if (p.144)
>
> (i) *c* is strongly current chance* $\mathfrak{I}$-dominated by probabilistic *c** relative to *E*,
> (ii) *c** is not weakly current chance* $\mathfrak{I}$-dominated by any credence function relative to *E*, and
> (iii) *c** is not extremely $\mathfrak{I}$-self-undermining then
> (iv) *c* is irrational as a credence function for any agent with inaccuracy measure $\mathfrak{I}$ and evidence *E*.

Thus, according to these principles, it is not the unanimous verdict of the current chance functions that should guide your decision-making; rather it is the unanimous verdict of the current chance* functions. If we accept this, we have the following argument for NTP:

(I_{NTP}) **Veritism**
(II_{NTP}) **Brier Alethic Accuracy**
(III_{NTP}) **Current Chance* Evidential Immodest Dominance**
(IV_{NTP}) **Theorem 10.1.1** and **I.B.2** Therefore,
(V_{NTP}) **Probabilism + NTP**

(p.145) Appendix II: A summary of chance-credence principles

In this part of the book, we have encountered an array of chance-credence principles, some of which differ only slightly. In this appendix, we provide a summary of all the principles together with a short explanation of their shortcomings.

- **PP**$_0$ $c_0(X|C_{ch}) = ch(X)$
  Problems:
  (i) Limited to initial credence functions.
  (ii) Requires initial credence functions to defer to ur-chance functions, which are most likely the chance functions at a very distant time.

- **PP**$_{\text{Lewis2}}$ $c_0(X|C_{ch}\ H_t) = ch_{H_t}(X)$
  Problems:
  (i) This is equivalent to PPo and inherits its problems.

- **PP**$_{\text{Lewis1}}$ $c_0(C_{ch}\ \&\ H_t\ \&\ E) = ch_{H_t}(X)$, providing E is admissible for X relative to the ur-chances conditional on H_t.
  Problems:
  (i) Limited to initial credence functions.
  (ii) Relies on a notion of admissibility, which is not fully specified.

- **TPP**$_0$ $c_0(X|T_{ch}) = ch(X)$
  Problems:
  (i) Limited to initial credence functions.

- **PP** $c(X|C_{ch}) = ch(X|E)$
  Problems:
  (i) Requires all credence functions to defer to ur-chance functions brought up to speed with their evidence. That is, treats chance functions as only analyst experts, not database experts.

- **TPP** $c(X|T_{ch}) = ch(X)$
  Problems:
  (i) Ignores agent's evidence and thus cannot deal with inadmissible evidence.

- **ATP** $c(X|T_{ch}) = ch(X)$, providing the agent has no inadmissible evidence about X.
  (p.146) Problems:

(i) Relies on a notion of admissibility, which is not fully specified.

- **ETP** $c(X|T_{ch}) = ch(X|E)$

Problems:

(i) If all possible current chance functions are non-self-undermining relative to E, then there are no problems.

(ii) If some possible current chance functions are self-undermining relative to E, then ETP demands that they are assigned credence 0. So it is too strong.

- **ETP**$^{+}$ c is in cl($\mathscr{C}^{+}$).

Problems: None.

- **GR** $c(X) = \sum_{ch \in C} c(T_{ch}) ch(X|E)$

Problems:

(i) If all possible current chance functions are non-self-undermining relative to E, then there are no problems.

(ii) If some possible current chance functions are self-undermining relative to E and tolerate non-self-undermining, then GR demands that they are assigned credence 0. So it is too strong.

- **NTP** $c(X|T_{ch}) = ch(X|T_{ch} \mathbin{\&} E)$

Problems:

(i) While we have not discussed this here, the central problem with NTP is that its consequences are sometimes in tension with our intuitions, as Lewis (1994) recognized. Nonetheless, Lewis argued, it approximates those intuitions closely enough. I won't explore this debate further here.

(p.147) Appendix III: The mathematical results

III.A Proof of Theorem III.A.2

Let $\mathscr{P}$ be a set of probability functions defined on $\mathscr{F}$. Suppose there is a finite subset $\mathscr{S} \subseteq \mathscr{P}$ and, for each p' in $\mathscr{S}$, there is a proposition $S_{p'}$ in $\mathscr{F}$ such that, if p' is in $\mathscr{S}$ and p is in $\mathscr{S}$, $p(S_{p'}) = 1$ if $p = p'$ and $p(S_{p'}) = 0$ if $p \neq p'$.

Definition III.A.1 (Deference to $\mathscr{S}$) *Suppose c is a credence function defined on $\mathscr{F}$. Then we say that c defers to $\mathscr{S}$ if, for all p' in $\mathscr{S}$ such that* $c(S_{p'}) > 0$,

$$c(X|S_{p'}) = p'(X)$$

for all X in $\mathscr{F}$.

Equivalently,

$$c(X \,\&\, S_{p'}) = p'(X)c(S_{p'})$$

for all X in $\mathscr{F}$.

Theorem III.A.2 *If c is in* cl($\mathscr{S}^+$), *then c defers to $\mathscr{S}$.*

Proof. To prove the theorem, it is sufficient to show:

(1) If p is in $\mathscr{S}$, then p defers to $\mathscr{S}$.
Proof. Suppose p' in $\mathscr{S}$ and p in $\mathscr{S}$. Then $p(S_{p'})$ if $p = p'$ and $p(S_{p'}) = 0$ if $p \neq p'$. Thus,

- If $p \neq p'$, then
$$p(X \,\&\, S_{p'}) = p(X) = p'(X) = p'(X)p(S_{p'})$$
- If $p \neq p'$, then
$$p(X \,\&\, S_{p'}) = 0 = p'(X)p(S_{p'})$$

(2) Suppose c, c' defer to $\mathscr{S}$. Then $\lambda c + (1-\lambda)c'$ defers to $\mathscr{S}$, for any $0 \leq \lambda \leq 1$. *Proof*. Suppose c, c' both defer to $\mathscr{S}$.
$$\begin{aligned}(\lambda c + (1-\lambda)c')(X \,\&\, S_{p'}) &= \lambda c(X \,\&\, S_{p'}) + (1-\lambda)c'(X \,\&\, S_{p'})\\ &= \lambda p'(X)c(S_{p'}) + (1-\lambda)p'(X)c'(S_{p'})\\ &= p'(X)(\lambda c + (1-\lambda)c')(S_{p'})\end{aligned}$$
(p.148) (3) If $c_1, c_2, \ldots$defer to $\mathscr{S}$, and $c = \lim_{n\to\infty} c_n$, then c defers to $\mathscr{S}$.
Proof. Suppose $c_1, c_2, \ldots$ defer to $\mathscr{S}$. Then
$$\begin{aligned}c(X \,\&\, S_{p'}) &= \lim_{n\to\infty} c_n(X \,\&\, S_{p'})\\ &= \lim_{n\to\infty} p'(X)c_n(S_{p'})\\ &= p'(X)\lim_{n\to\infty} c_n(S_{p'})\\ &= p'(X)c(S_{p'})\end{aligned}$$

This completes our proof. □

III.B Proof of Theorem 11.1.1

Theorem 11.1.1 Suppose ch is a possible current chance function that is self-undermining relative to the agent's evidence — that is, $ch(T_{ch}|E) < 1$. Then, if our agent has credence function c, T_{ch} is in $\mathscr{F}$, and c satisfies ETP, then $c(T_{ch}) = 0$.

Proof. Suppose $c(T_{ch}) > 0$. Then, since c satisfies ETP, it follows that $c(T_{ch}|T_{ch}) = ch(T_{ch}|E) < 1$. But, by the ratio definition of conditional probability, $c(T_{ch}|T_{ch}) = 1$. This gives a contradiction, as required. □

III.C Proof of Theorem 11.2.1

Theorem 11.2.1 If each possible current chance function is expectationally non-selfundermining in the presence of E, then ETP$^+$ entails GR.

Proof. It suffices to show that, if every member of $\mathscr{C}$ is expectationally non-selfundermining in the presence of E, then every element of $\mathrm{cl}(C_E^+)$ satisfies GR. Now, if each member of $\mathscr{C}$ is expectationally non-self-undermining in the presence of E, then each member of $\mathscr{C}_E$ satisfies GR. Moreover, if c and c' both satisfy GR, then so does $\lambda c + (1 - \lambda)c'$, for any $0 \leq \lambda \leq 1$. And the limit of an infinite sequence of credence functions each of which satisfies GR will also itself satisfy GR. It follows that every member of $\mathrm{cl}(C_E^+)$ satisfies GR. □

III.D Proof of Theorem 11.2.2

Theorem 11.2.2 Suppose ch is a possible current chance function; and suppose ch is self-undermining, but tolerant of non-self-undermining in the presence of E—that (p.149) is, ch' is self-undermining in the presence of E, but there is ch'' that is non-selfundermining in the presence of E, and $ch'(T_{ch''}|E) > 0$. Then, if credence function c satisfies GR, then $c(T_{ch'}) = 0$.

Proof. Suppose c satisfies GR. Thus, in particular,

$$c(T_{ch''}) = \sum_{ch \in C} c(T_{ch}) ch(T_{ch''}|E)$$

Then, subtracting $c(T_{ch''})$ from both sides, we get:

$$0 = \sum_{ch \neq ch''} c(T_{ch}) ch(T_{ch''}|E)$$

Thus, $c(T_{ch'}) ch'(T_{ch''}|E) = 0$. But $ch'(T_{ch''}|E) > 0$, by hypothesis. So $ch(T_{ch'}) = 0$, as required. □ (p.150)

Notes:

([1]) In fact, $ch_0^* = ch_0$ and $ch_4^* = ch_4$. But for i = 1,2,3, $ch_i^* \neq ch_i$.

([2]) Jenann Ismael has pursued this argument as a response to objections that I raised against her General Recipe (Ismael, 2013). More on that below.

([3]) In fact, in order to make the restricted-scope response work, even on its own terms, we would have to be quite careful. We would have to ensure that the demands made by this new, restricted version of the Temporal Principle, when combined with the demands made by the agent's evidence and the demand of Probabilism, does not end up making the demand on the agent that was problematic for the unrestricted version of the Temporal Principle in the first place. For instance, recall that $C_{ch_1} \equiv HTTT \vee THTT \vee TTHT \vee TTTH$ are in $\mathscr{F}'$, and if the agent has no evidence, and if $c(C_{ch_1}) > 0$, then the restricted version of the Temporal Principle will demand that $c(HTTT|C_{ch_1}) = ch_1(HTTT), \dots, c(TTTH|C_{ch_1}) = ch_1(TTTH)$. And then Probabilism will demand that $c(C_{ch_1}|C_{ch_1}) = ch_1(HTTT \vee \dots TTTH|C_{ch_1}) = c(HTTT|C_{ch_1}) + \dots + c(TTTH|C_{ch_1}) = ch_1(HTTT) + \dots + ch_1(TTTH) = ch_1(C_{ch_1})$. This gives a contradiction, as in the proof of Theorem 11.1.1. So, together the restricted version of the Temporal Principle and Probabilism demand that $c(C_{ch_1}) = 0$, just as the unrestricted version of the Temporal Principle demanded.