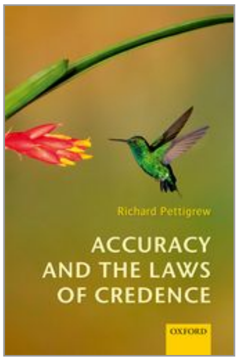


Accuracy and the Laws of Credence
Richard Pettigrew

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

# Measuring accuracy: a new account

Richard Pettigrew

DOI:10.1093/acprof:oso/9780198732716.003.0005

## Abstract and Keywords

This chapter presents the preferred characterization of the legitimate inaccuracy measures. It begins by characterizing all of the strictly proper inaccuracy measures. Later, a further condition is considered that narrows the field to a single inaccuracy measure, namely, the popular Brier score.

*Keywords:* Brier score, strictly proper scoring rule, Bregman divergence, calibration, additivity

I come at last to my own favoured characterization of the legitimate inaccuracy measures. It begins with one of the three postulates that conjoin to give Brier Alethic Accuracy. Recall, from the Introduction:

> **Perfectionism** The accuracy of a credence function at a world is its proximity to the ideal credence function at that world.

Thus, we assume that, if $\mathfrak{I}$ is a legitimate inaccuracy measure, then there is a measure $\mathfrak{D}$ of distance from one credence function to another credence function defined on the same set such that $\mathfrak{I}(c, w) = \mathfrak{D}(i_w, c)$, where $i_w$ is the ideal credence function at $w$. When mathematicians formalize the notion of distance, they usually make three assumptions:

- $\mathfrak{D}$ *is non-negative*. That is, $\mathfrak{D}(c, c') \geq 0$ with equality iff $c = c'$.
- $\mathfrak{D}$ *is symmetric*. That is, $\mathfrak{D}(c, c') = \mathfrak{D}(c', c)$ for all $c, c'$.
- $\mathfrak{D}$ *satisfies the triangle inequality*. That is, $\mathfrak{D}(c, c'') \leq \mathfrak{D}(c, c') + \mathfrak{D}(c', c'')$ for all $c, c', c''$.

If $\mathfrak{D}$ satisfies all three conditions, we say that $\mathfrak{D}$ is a *metric*. At least to begin with, we will be assuming only that $\mathfrak{D}$ is non-negative. We will follow statisticians in calling a non-negative function of two credence functions a *divergence*. In Section 4.4, we will consider whether we can assume that $\mathfrak{D}$ is also symmetric—indeed, this assumption will prove to be very powerful in the presence of our other assumptions. But we will never demand that $\mathfrak{D}$ satisfies the triangle inequality.[1] Thus, we make Perfectionism precise as follows:

> (p.48) **Perfectionism** If $\mathfrak{I}$ is a legitimate inaccuracy measure, there is a divergence $\mathfrak{D}$ such that $\mathfrak{I}(c, w) = \mathfrak{D}(i_w, c)$. Recall: $i_w$ is the ideal or vindicated credence function at $w$. We say that $\mathfrak{D}$ generates $\mathfrak{I}$ (relative to that notion of vindication).

The remaining two postulates that combine to give Brier Alethic Accuracy provided the identity of $i_w$ (Vindication) and the identity of $\mathfrak{D}$ (Squared Euclidean Distance). To begin with, we will assume neither of these further postulates. Indeed, we wish to argue, at least initially, for a more liberal account of the divergences that generate legitimate inaccuracy measures than that given by Squared Euclidean Distance. We will do this by laying down further conditions on such divergences. We would like to lay down these conditions irrespective of what we take to be the ideal credence function at a world.

As we will see, some divergences are legitimate for other purposes, but not for the purpose of measuring inaccuracy. For instance, the ‘taxicab’ divergence, which we encountered in Chapter 3, is legitimate if one’s purpose is to measure how far one credence function will have to move in order to match another credence function— it takes the amount that each credence will have to move, which is plausibly the difference between them, and sums these together. However, as we will see, it is not appropriate for measuring inaccuracy because it violates our fifth and central axiom, Decomposition, which we formulate below. First, however, there are two important axioms that we must motivate.

## 4.1 Additive divergences

We have met a version of the first of these two axioms already. In one of the characterizations that Leitgeb and I gave, we appealed to Agreement on Accuracy. Above, I noted that this entails Additivity, which says that the inaccuracy of a credence function at a world is the sum of the inaccuracies at that world of the individual credences it assigns. Thus, on this account, any legitimate inaccuracy measure on credence functions is generated by what we might think of as a 'local' inaccuracy measure: that is, a measure of the inaccuracy of an individual credence. When we stated it above, we were assuming Vindication: that is, we were assuming that the ideal credence function at a world is the omniscient credence function at that world. Thus, the postulate was stated as follows:

> **Additivity** If $\mathfrak{I}$ is a legitimate (global) measure of inaccuracy, then there is a local measure of inaccuracy $\mathfrak{s}$ such that
>
> $$\mathrm{I}(c, w) = \sum_{X \in \mathrm{F}} \mathrm{s}(v_w(X), c(X))$$

Here, we are not (yet) making that assumption about the ideal credence function. Thus, we state it more neutrally as follows:[2]

> (p.49) **Divergence Additivity** If $\mathfrak{I}$ is a legitimate inaccuracy measure generated by a divergence $\mathfrak{D}$, then ?? is additive.
>
> That is, if $\mathfrak{I}$ is a legitimate inaccuracy measure generated by $\mathfrak{D}$, then there is a one-dimensional divergence $\mathfrak{d}$ such that $\mathrm{D}(c, c') = \sum_{X \in \mathrm{F}} \mathrm{d}(c(X), c'(X))$. We say that $\mathfrak{d}$ *generates* $\mathfrak{D}$.

What motivates this assumption? The first thing to say is that summing the inaccuracy of individual credences to give the total inaccuracy of a credence function is the natural thing to do. When we say that we represent an agent by her credence function, it can sound as if we're representing her as having a single, unified doxastic state. But that's not what's going on. Really, we are just representing her as having an agglomeration of individual doxastic states, namely, the individual credences she assigns to the various propositions about which she has an opinion. A credence function is simply a mathematical way of representing this agglomeration; it is a way of collecting together these individual credences into a single object.

To illustrate the point, it might help to compare a credence function to a musical melody. Suppose I were to ask how far one melody lies from another. I would not simply treat each as a sequence of notes (pitches and durations) and measure the distance between each note in one and its counterpart in the other, and then sum them up. Rather, I would treat each melody as an integrated whole and I would ask how far the overall 'shape' of one lies from the overall 'shape' of the other. A credence function, on the other hand, is not an integrated whole—it is simply a mathematical representation of a list of credence-proposition pairings. Thus, we need not look to its ‘shape’ when we measure its distance from another credence function.

Thus, the situation is much like those situations in practical decision theory in which the outcome of an option at a given world is an agglomeration of commodities, each of which has a particular value for the agent: for instance, the outcome might consist of an orange, an apple, and a pear. If that list exhausts the constituents of the outcome that the agent values, it is natural to say that her utility for the outcome taken as a whole is given by the sum of the utilities she assigns to the individual commodities, namely, the orange, the apple, and the pear. The same is true in the case of the inaccuracy of a credence function. The outcome of having a credence function at a particular world can be viewed as a bundle of commodities: to each proposition to which the agent assigns a credence, there is a different commodity; and which commodity it is depends on the credence assigned by the agent and the credence assigned by the ideal credence function. Divergence Additivity simply says that we should measure the utility of the outcome taken as a whole in just the way we do in the practical case: that is, we ought to sum the utilities of its component commodities.

Of course, there are practical cases in which the outcome of an option is an agglomeration of commodities yet we don’t take its value to be simply the sum of the values of the commodities. These are cases in which some of the commodities in question are what economists call *dependent goods*. A dependent good is a good such that the value that it contributes to the overall value of a commodity bundle of (p.50) which it is a part depends upon the other commodities in the bundle. For instance, the value contributed to an outcome by a tin of beans depends on whether or not the outcome also includes a tin opener and a stove; the outcome contributed to an outcome by a television remote control depends on whether or not the outcome also includes a television that it can operate; and so on. Therefore, one consequence of Divergence Additivity is that accuracy is not a dependent good in this sense. That is, an inaccuracy measure that satisfies Divergence Additivity cannot assign less accuracy to a high credence in one true proposition on the basis that the agent already has a great deal of accuracy as a result of her high credence in another true proposition. In both cases, this consequence seems right. The badness of having a particular credence in a proposition should be the same and should contribute the same disutility to the overall badness of the credence function regardless of what values that credence function assigns to other propositions. Suppose Emmy is almost certain that 31 is prime and almost certain that 59 is prime (both numbers are, in fact, prime), while David is also almost cert ain that 31 is prime but not so sure about 59. We wouldn't want to say that the accuracy of David's credence in the former proposition contributes more to the overall accuracy of his credence function than the accuracy of Emmy's corresponding credence does to hers. Similarly, we wouldn't want to allow that, when Emmy and David both change their credence in the primality of 31 by exactly the same amount, then their overall inaccuracies might change by different amounts. In both cases, the reason is that accuracy is not a dependent good.

We might think that the following sort of case shows that accuracy is sometimes a dependent good. You and I both have credences in two propositions: Emmy Noether was German ($G$); Emmy Noether was German or Danish ($G \vee D$). Both propositions are true at the actual world. You have high credence in $G$; I have low credence in $G$. We both have high credence in the disjunction $G \vee D$. You might think that the accuracy contributed by this very accurate credence to my overall accuracy will be greater than the accuracy contributed by the same very accurate credence to your overall accuracy. After all, there is a sense in which you already have a strong proattitude towards the disjunction in virtue of your high credence in one of its disjuncts; thus, we seem to reason, you should receive fewer extra epistemic brownie points than I should. I think the mistake here arises again because we are thinking of credence functions as encoding a single, unified epistemic attitude. But they aren't. As the example shows, there is no unified attitude that my credence function encodes: it assigns high credence to $G$, but not to $G \vee D$. Thus, the accuracy that my high credence in $G \vee D$ contributes to my overall accuracy should be exactly the same as the accuracy that your same high credence contributes to yours.

It is a general consequence of Divergence Additivity that an inaccuracy measure cannot be sensitive to any irreducibly global features of a credence function, such as the minimum credence it assigns to a truth, for instance, or the maximum it assigns to a falsehood. Thus, as is often claimed in the case of practical decision theory, whatever the significance of irreducibly global features of options for our choices between them, (p.51) it should be reflected in the decision principles we adopt, not in the utilities we assign to the outcomes of those options. Just as some hold that risk aversion phenomena in practical decision theory are best understood as the result of doing something other than maximizing expected utility—minimizing regret, for instance, or maximizing the quantity favoured by one of the many non-expected utility theories—and not as having a concave utility function, so any sensitivity to global features of credence functions ought to be understood either as following from their local features or as following from the adoption of an alternative decision principle and not as having a non-additive inaccuracy measure.[3]

## 4.2 Continuity and the absence of jumps

According to Perfectionism, each legitimate inaccuracy measure is generated by a divergence. What sorts of divergences generate legitimate inaccuracy measures? In the previous section, we did not assume a particular account of vindication. We have argued that a legitimate inaccuracy measure is additive (Divergence Additivity) regardless of the notion of vindication we assume. One consequence of this is that a divergence that generates a legitimate inaccuracy measure must be additive as well—if it is not, there will be notions of vindication such that, relative to them, the divergence gives rise to an inaccuracy measure that is not additive. That is, putting together Perfectionism and Divergence Additivity, we obtain the following: if $\mathfrak{I}$ is a legitimate inaccuracy measure, then there is a divergence $\mathfrak{D}$ such that

(1) $\mathfrak{I}(c, w) = \mathfrak{D}(i_w, c)$. In this case, we say that $\mathfrak{I} = \mathfrak{I}_{\mathfrak{D}}$.

(2) There is a function $\mathfrak{d} : [0,1] \times [0,1] \to [0, \infty]$such that

(i) for all $x,y \in [0,1]$, $\mathfrak{d}(x, y) \geq 0$ with equality iff $x = y$;

(ii) $\mathrm{D}(c, c') = \sum_{X \in \mathrm{F}} \mathrm{d}(c(X), c'(X))$.

In this case, we call $\mathfrak{d}$ a *one-dimensional divergence* and we say that $\mathfrak{d}$ *generates* $\mathfrak{D}$.[4]

The next condition we will place on the legitimate inaccuracy measures presupposes this conclusion. It is a generalization of a condition that we have met already. In both of his characterizations of legitimate inaccuracy measures, Joyce assumed the following condition:

(p.52) **Continuity** If $\mathfrak{I}$ is a legitimate inaccuracy measure, then $\mathfrak{I}(c, w)$ is a continuous function of $c$, for all worlds $w$.

We generalize this requirement as follows:

**Divergence Continuity** If $\mathfrak{I}$ is a legitimate inaccuracy measure and there is a divergence $\mathfrak{D}$ generated by $\mathfrak{d}$ such that

$$\mathrm{I}(c, w) = \mathrm{I}_{\mathrm{D}}(c, w) = \mathrm{D}(i_w, c) = \sum_{X \in \mathrm{F}} \mathrm{d}(i_w(X), c(X))$$

then $\mathfrak{d}(x, y)$ is continuous in both of its arguments.

Let us begin by explaining what this means. To demand that $\mathfrak{d}$ is continuous in its second argument is to say that there are no ‘jumps’ in inaccuracy as credences change. That is, whatever credence you have and however little you wish its accuracy at a given world to change, there is some neighbourhood around your current credence such that, if you keep your credence in that neighbourhood, your accuracy at the world in question wont change any more than you wish it to. More precisely: Suppose $x$ is the ideal credence in $X$ and $y$ is a credence in $X$. And let $\varepsilon > 0$ be a small but positive number. Then, by demanding that $\mathfrak{d}$

is continuous in its second argument, we say that we can always find another small but positive number $\delta > 0$ such that, providing a credence $z$ is at most $\delta$-far from $y$, the inaccuracy of $z$ is at most $\varepsilon$-far from the inaccuracy of $y$. Here it is in symbols: $\mathfrak{d}$ is continuous in its second argument iff, for all $0 \leq x, y \leq 1$, the following holds:

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, z)[|y - z| < \delta \Rightarrow |\mathrm{d}(x, y) - \mathrm{d}(x, z)| < \varepsilon]$$

For instance, the one-dimensional squared Euclidean distance divergence $\mathfrak{d}^2(x, y) = |x - y|^2$ is continuous in its second argument. On the other hand, the following one-dimensional divergence, $\mathrm{d}^2_{0.9}$, is discontinuous in its second argument (for $x \neq 0.81$):

$$\mathrm{d}^2_{0.9}(x, y) := \begin{cases} |x - y|^2 & \text{if } y \leq 0.9 \\ 0 & \text{if } y > 0.9 \end{cases}$$

That is, $\mathrm{d}^2_{0.9}$ agrees with $\mathfrak{d}^2$ for credences up to and including 0.9; for higher credences, it is always 0. $\mathrm{d}^2_{0.9}$ is discontinuous at 0.9 (for $x \neq 0.81$). Suppose $x = 0$. Then the inaccuracy of $y = 0.9$ is $\mathrm{d}^2_{0.9}(0, 0.9) = 0.81$. Now there is no neighbourhood, however small, around $y = 0.9$ such that, if you keep your credences in that neighbourhood, your inaccuracy is guaranteed to be within, say, $\varepsilon = 0.5$ of $\mathrm{d}^2_{0.9}(0, 0.9) = 0.81$—after all, any such neighbourhood will include credences greater than 0.9, and their inaccuracy will be 0.

On the other hand, to demand that $\mathfrak{d}$ is continuous in its first argument is to say that there are no 'jumps' in inaccuracy as the ideal credence function changes. In symbols: $\mathfrak{d}$ is continuous in its first argument iff, for all $0 \leq x, y \leq 1$, the following holds:

$$(\forall\, \varepsilon > 0)(\exists\, \delta > 0)(\forall\, z)[|x - z| < \delta \Rightarrow |\mathrm{d}(x, y) - \mathrm{d}(z, y)| < \varepsilon]$$

(p.53) Why, then, must a one-dimensional divergence that generates a divergence that gives rise to a legitimate inaccuracy measure be continuous in both of its arguments? Again, it helps to begin by noting that this is a natural assumption. If inaccuracy is distance from ideal credence, as Perfectionism claims, it is natural to assume that inaccuracy will vary continuously. Thus, it seems to me that the burden of proof lies with anyone who wishes to deny Continuity. Or, rather, if there is a legitimate inaccuracy measure $\mathfrak{I}$ that is based on a divergence $\mathfrak{D}$, which is in turn generated by a one-dimensional divergence $\mathfrak{d}$, then if $\mathfrak{d}$ is discontinuous at some credence or ideal credence, then there must be some reason for the discontinuity—the assumption, after all, is on the side of continuity.

What might this reason be? Here is one possibility that arises if we follow Vindication and take the ideal credences ($i_w$) to be the omniscient credences ($\nu_w$). If you accept the *Lockean Thesis* concerning the relationship between full belief and credence, you will think that there is a threshold $\frac{1}{2} < t \leq 1$ such that:[5]

(i) An agent has a belief in $X$ iff her credence in $X$ is at least $t$;
(ii) An agent has a disbelief in $X$ iff her credence in $X$ is at most $1 - t$.

Thus, you might expect a discontinuity in a local inaccuracy measure (or scoring rule) at $1 - t$ and at $t$. That is, if $\mathfrak{s}$ is a local inaccuracy measure, you might expect $\mathfrak{s}(1, x)$ and $\mathfrak{s}(0, x)$ to be discontinuous at $x = 1 - t$ and at $x = t$. And if you expect that, then you will expect a discontinuity in the second argument of the one-dimensional divergence $\mathfrak{d}$ that gives rise to it—that is, you will expect $\mathfrak{d}(1, x)$ and $\mathfrak{d}(0, x)$ to be discontinuous at $x = 1 — t$ and $x = t$ as well. After all, suppose I move from credence $t - \varepsilon$ to $t$ in a falsehood. Then I will increase my inaccuracy by moving my credence away from the omniscient credence of 0 in this proposition. But I will also have acquired a new full belief in this falsehood, which I didn't have before. And you might expect this newly acquired and inaccurate doxastic state—my new full belief in a falsehood—to give a 'bump' to my inaccuracy; you might expect it to give rise to a 'jump' in inaccuracy. That is, you might expect my inaccuracy to increase discontinuously over the threshold $t$ because my doxastic state changes discontinuously over that threshold—it moves from involving no full beliefs or full disbeliefs to involving a single full belief. Up to $t$, I have only my degrees of belief in the proposition; from $t$ upwards, I have those as well as a full belief in a falsehood. And similarly for the threshold $1 - t$, which leads me from neither belief nor disbelief to disbelief.

The thought is tempting, I can see, but I think it's wrong. Before we say why it's wrong, let's consider a bad argument against it. This bad argument protests that our job here is to characterize legitimate measures of the inaccuracy only of the credal part of a doxastic state, not the total doxastic state, which might include full beliefs and disbeliefs as well. Thus, even if other doxastic states come and go in accordance with certain features of our credences, our inaccuracy measure on a credence function need (p.54) not reflect that. So, according to the bad argument, we accept that the inaccuracy of a total doxastic state may differ from the inaccuracy of its credal part, perhaps because it includes a separate measure of the inaccuracy of the full belief part as well; but we note that here we are only interested in characterizing legitimate inaccuracy measures for credence functions, not total doxastic states; so, we conclude, it is acceptable to assume Continuity.

To see why this is a bad argument, we note that, if that is all that we are doing here, our argument for Probabilism is in danger of being severely weakened. After all, it may be that, although all non-probabilistic credence functions are dominated when only the inaccuracy of the credence function is measured, there are non-probabilistic credence functions that give rise (via the Lockean Thesis) to total doxastic states— credences along with full beliefs and disbeliefs —that are not dominated when we measure the inaccuracy of the credence function and the full beliefs to which they give rise (*via the Lockean Thesis*). Indeed, on one very natural way of spelling out the inaccuracy of a total doxastic state, this is exactly what happens, as we will now show.[6]

Consider the following situation. The threshold for full belief is 0.7. Rozy has credences only in the propositions $Rain$ and $\overline{Rain}$. She has credence 0.7 in $Rain$ and 0.6 in $\overline{Rain}$. Call her credence function $c$. Thus, by the Lockean Thesis, Rozy has a full belief in $Rain$, but no full belief or disbelief in $\overline{Rain}$. Now, suppose we consider the inaccuracy of the credal part of her total doxastic state at each world, the rainy world and the dry world. Then she is Brier dominated by credence functions that are not themselves Brier dominated and that indeed are not moderately Brier-modest. This, according to our accuracy argument for Probabilism, renders her irrational.

However, suppose we now measure the inaccuracy of Rozy's whole doxastic state. How are we to do this? A natural suggestion is that we add the inaccuracy of the credal part to the inaccuracy of the full belief part. So let's try that. How are we to measure the inaccuracy of a set of beliefs and disbeliefs? Again, we follow the natural suggestion, which has been pursued by Hempel (1962), Easwaran (to appear), and Fitelson (ms):

(i) True beliefs and false disbeliefs get a reward of $R$ (we assume $R > 0$, so we say that the inaccuracy of a true belief or false disbelief is $-R$)
(ii) False beliefs and true disbeliefs get a penalty of $W$ (we assume $W > 0$, so we say that the inaccuracy of a false belief or true disbelief is $W$)

And then we take the inaccuracy of a set of beliefs and disbeliefs to be simply the sum of the inaccuracy of the individual states (thus, in the appropriate sense, it is additive).

On the basis of this measure of the inaccuracy of the full beliefs and disbeliefs that an agent has, we can define the inaccuracy of her total doxastic state. Now, since the Lockean Thesis allows us to determine the total doxastic state just by looking at its credal part, we can define an inaccuracy measure for a total doxastic state as an inaccuracy measure on a credence function. So suppose that we measure the credal (p.55) part of a doxastic state using the Brier score, which is generated by the quadratic scoring rule; and we measure the inaccuracy of the full beliefs and disbeliefs as suggested above. Then we can define the following measure of the inaccuracy of the total state:

$$\mathfrak{q}_t(1, x) := \begin{cases} \mathfrak{q}(1, x) - R & \text{if } t \leq x \leq 1 \\ \mathfrak{q}(1, x) & \text{if } 1 - t < x < t \\ \mathfrak{q}(1, x) + W & \text{if } 0 \leq x \leq t \end{cases}$$

$$\mathfrak{q}_t(0, x) := \begin{cases} \mathfrak{q}(0, x) + W & \text{if } t \leq x \leq 1 \\ \mathfrak{q}(0, x) & \text{if } 1 - t < x < t \\ \mathfrak{q}(0, x) - R & \text{if } 0 \leq x \leq t \end{cases}$$

Thus, for example, suppose the threshold for belief is $t = 0.7$. And suppose my credence in a proposition is 0.8. And suppose that the proposition is true. Then the inaccuracy of my total doxastic state is given by the inaccuracy of the credal component, which is q (1,0.8), added to the inaccuracy of the belief components, which is $-R$, because my credence in the proposition gives rise to a belief via the Lockean Thesis, and the proposition is true. Thus, my total inaccuracy is $q(1{,}0.8) - R$, just as qt says. Notice that, as we anticipated, $\mathfrak{q}_t$ is discontinuous at $t$ and at $1 - t$.

Having defined the local inaccuracy of the total doxastic state corresponding to a particular credence $x$, I define the global inaccuracy for the total doxastic state corresponding to an entire credence function $c$ as follows:

$$\mathrm{B}_t(c, w) := \sum_{X \in \mathrm{F}} q_t(v_w(X), c(X))$$

Now let us return to Rozy’s doxastic state. We can show the following:[7]

> **Proposition 4.2.1** *There is no credence function c* that weakly* $\mathfrak{B}_t$*-dominates c.*

The reason, roughly, is that, while $c$ is $\mathfrak{B}$-dominated, the credence functions that $\mathfrak{B}$-dominate it do not give rise to a full belief in *Rain*. It is the accuracy of this full belief in *Rain* that boosts the accuracy of Rozy’s credence function and prevents those $\mathfrak{B}$-dominators from $\mathfrak{B}_t$-dominating her. A little extra work shows that no other credence function $\mathfrak{B}_t$-dominates her either. Thus, Rozy’s non-probabilistic credence function is not $\mathfrak{B}_t$-dominated. The result is illustrated in Figure 4.1. Since we

presumably care ultimately about the accuracy of our total doxastic state, this result is very worrying for the accuracy argument for Probabilism. It seems wrong to say that Rozy is irrational on the grounds that the credal part of her doxastic state is accuracy dominated, when her total doxastic state is not accuracy dominated.
(p.56)

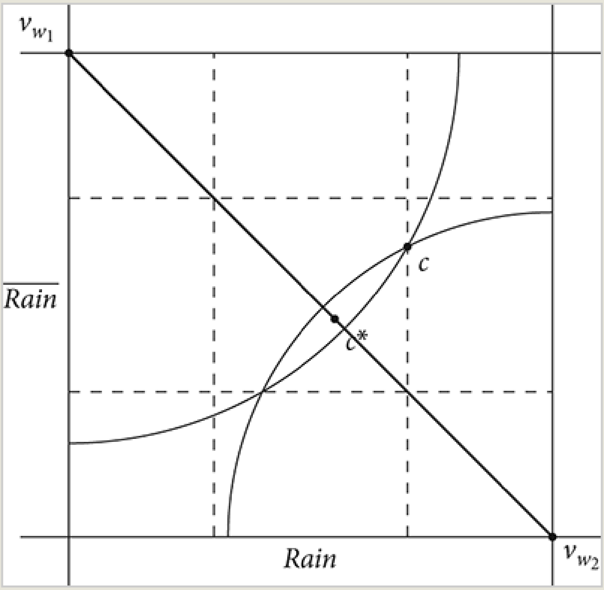


***Figure 4.1*** The dotted lines represent the Lockean thresholds: the horizontal lines mark the thresholds for belief and disbelief in $\overline{Rain}$; the vertical lines mark the thresholds for belief and disbelief in *Rain*. Note that all credence functions that dominate *c*—that is, those that lie between the two arcs—give rise (via the Lockean Thesis) to neither full beliefs nor full disbeliefs on *Rain* and on $\overline{Rain}$. On the other hand, *c* gives rise (via the Lockean Thesis) to a belief in *Rain*. And thus *c* receives a boost in its accuracy at world $w_2$ from the accuracy of the belief to which it gives rise that its dominators do not enjoy. And this is enough to ensure that it is not accuracy dominated when we consider the inaccuracy of the total doxastic state.

We considered the accuracy of the total doxastic state because we were wondering how to respond to the suggestion that the inaccuracy of a credence at a world ought to be discontinuous at the two Lockean thresholds (the belief threshold and the disbelief threshold). The foregoing result shows that we mustn't respond to this suggestion by accepting that way of measuring the inaccuracy of a total doxastic state if we are to retain our argument for Probabilism.

How else are we to measure it? On the view that motivates the above suggestion, full beliefs and disbeliefs are *sui generis* doxastic states, whose occurrence is closely connected to high or low credence, but which are something over and above those.[8] Thus, while the inaccuracy of the credal part of a doxastic state may be a continuous function of credences, the inaccuracy of the total state is not, because those *sui generis* states pop in and out of existence discontinuously. However, this is not my view of full beliefs and disbeliefs. Rather, for me, to say that someone has a belief is just to say that they have high enough credence (and saying they have a disbelief is just to say that they have a low enough credence). It is not to attribute to them some state over and above the high credence (or low credence). It is analogous to saying that someone (p.57) is tall or saying that they are far from London. In neither case would we say that we are ascribing to them some property over and above the property of being at least 6 ft in height, say, or being more than 500 miles from London.[9] Rather, our assertion is merely a shorthand for these properties. Thus, if we are measuring how close a person is to 7 ft tall, for instance, we don't use a distance measure that is discontinuous at 6 ft on the grounds that, when they move from 5 ft 11 in to 6 ft, they not only move an inch closer to 7 ft, but also become tall and thus, since 7 ft is also tall, take a discontinuous leap closer. Rather, we say that the move closer to 7 ft is exhausted by the change in height in inches—moving into the category 'tall' does nothing extra to affect it.

The same goes for credences and full beliefs and disbeliefs: moving closer to the omniscient credence by moving across a Lockean threshold may make it appropriate to ascribe a state of belief or disbelief, but these states are nothing over and above the credences that give rise to them; so their inaccuracy adds nothing to the inaccuracy of a total doxastic state that is not already contributed by the inaccuracy of the credences. Thus, I conclude, our measures of inaccuracy should be continuous.

## 4.3 Calibration and accuracy

We come now to the final condition that we will place on a divergence that generates a legitimate inaccuracy measure. To motivate this condition, we look at a different account of the 'ideal' credence function from the one offered by Vindication. We begin with a quotation from Ramsey:

> Granting that [an agent] is going to think always in the same way about all yellow toadstools, we can ask what degree of confidence it would be best for him to have that they are unwholesome. And the answer is that it will in general be best for his degree of belief that a yellow toadstool is unwholesome to be equal to the proportion of yellow toadstools that are unwholesome. (This follows from the meaning of degree ofbelief.) (Ramsey, 1931, 195)

In this passage, Ramsey seems to identify and endorse an alternative candidate for the role of *perfect* or *ideal* credence function; that is, he seems to disagree with Vindication that the omniscient credence function is the ideal credence function, and he proposes an alternative that it “would be best” for an agent’s credences to match. For Ramsey, it seems, the ideal credence function is the one that assigns to a proposition the proportion of true propositions amongst all propositions that are relevantly similar to it. Thus, the ideal credence in the proposition *Yellow toadstool t is unwholesome* is the proportion of all propositions of that type—that is, propositions of the form *Yellow toadstool x is unwholesome*—that are true: in other words, it is the proportion of yellow toadstools that are unwholesome.

(p.58) Now, it is clear that, in many cases, this credence function, which Ramsey seems to take to be ideal, is different from the omniscient credence function, which Joyce and I take to be ideal. The omniscient credence function at a world will assign credence 1 to the proposition *Yellow toadstool t is unwholesome* if, and only if, *t* really is an unwholesome yellow toadstool at that world. Ramsey’s alternative ideal credence function, in contrast, will assign 1 to that proposition if, and only if, *all* yellow toadstools are unwholesome, since then and only then will the proportion of true propositions amongst all propositions of the same type be 1.

However, this difference relies on us taking all propositions of the form *Yellow toadstool x is unwholesome* to be of the same type as the proposition *Yellow toadstool t is unwholesome*. Suppose instead that we take each proposition to be *sui generis*—that is, for each *x*, we take *Yellow toadstool x is unwholesome* to be the only proposition of the same type as itself. Then, since on Ramsey’s account of ideal credences, the ideal credence in *Yellow toadstool t is unwholesome* is the proportion of true propositions amongst those of the same type as that proposition, and since it is the only proposition of the same type as itself, the ideal credence is 1 if the proposition is true and 0 if it is false. Thus, if we treat all propositions as *sui generis*, Ramsey’s ideal credence function coincides with the omniscient credence function, which Vindication takes to be the ideal credence function. The same is true if the only propositions of the same type as a given proposition are those that are logically equivalent to it. However, if the types are any more permissive than this—that is, if there are two propositions of the same type that are not equivalent—then Ramsey’s ideal credence function at some worlds will be different from the corresponding omniscient credence function.

The upshot is this: on Ramsey's account, the ideal credence function depends on the reference classes into which we partition the set $\mathscr{F}$ of propositions to which the agent assigns credences. Which is the correct partition of $\mathscr{F}$? A natural suggestion is that the world fixes this: that is, propositions just naturally fall into distinct types, which somehow carve the space of propositions at its joints. Thus, the world fixes that *Yellow toadstool t is unwholesome* is of the same type as *Yellow toadstool t′ is unwholesome* but not of the same type as *Red toadstool t″ is unwholesome*. I rather doubt that the world can do this. Even if we grant the Lewisian ideology of natural properties, it isn't obvious to me how to move from this to a natural taxonomy of propositions (Lewis, 1983).

Another suggestion, which has been pursued by Bas van Fraassen and Abner Shimony, is that the reference classes are fixed by the agent's credence function (van Fraassen, 1983; Shimony, 1988). The idea is this: Suppose $c$ is a credence function defined on the opinion set $\mathscr{F}$. Then, relative to $c$, two propositions $X$ and $X'$ in $\mathscr{F}$ belong to the same reference class iff $c(X) = c(X')$. This departs from Ramsey's suggestion a little. There is no longer a single credence function that is ideal at any given world: rather, a credence function is ideal or perfect at a world if, for each credence $x$ that it assigns, the proportion of true propositions amongst all propositions (p.59) to which it assigns $x$ is $x$. In this case, we say that the credence function is *well calibrated*:

*Definition 4.3.1* **(Well calibrated)** *Suppose c is a credence function defined on $\mathscr{F}$ and w is a possible world relative to $\mathscr{F}$. Then c is well calibrated at w if, for each x in* $\mathrm{ran}(c)$,[10]

$$x = \frac{|\{X \in \mathscr{F} : c(X) = x \text{ and } v_w(X) = 1\}|}{|\{X \in \mathscr{F} : c(X) = x\}|}$$

Note that, for each world $w$, the omniscient credence function $v_w$ at $w$ is well calibrated. After all, $v_w$ assigns credence 1 to a proposition iff that proposition is true, so 100% of the propositions to which it assigns 1 are true; and $v_w$ assigns credence 0 to a proposition iff that proposition is false, so 0% of the propositions to which it assigns 0 are true, as required. Indeed, van Fraassen and Shimony contend that it is this feature of $v_w$ that gives it the maximal accuracy it enjoys at world $w$. But other credence functions enjoy this maximal accuracy at $w$ as well, according to their account. For instance, if a credence function is defined on only 10 propositions, and if it assigns credence 0.9 to each of those 10 propositions, then it is well calibrated at a world at which 9 out of 10 of those propositions are true. So, if I have a credence of 0.9 that I will lose a 10-ticket lottery, I will be well calibrated whichever ticket wins. According to van Fraassen and Shimony, these credence functions—and all others that share this feature—have maximal accuracy. For them, what it means for a credence function to match a world perfectly is that it assigns credences that match certain frequencies that obtain at that world. Note that the frequencies that these credences must match are partly determined by the agent and partly determined by the world. The agent fixes the reference class with respect to which the frequencies are calculated. She does this by distributing her credences in a particular way—the reference class of a particular proposition is the set of propositions to which the agent assigns the same credence. The world then fixes the frequency of true propositions amongst those in each reference class. According to van Fraassen and Shimony, the agent matches the world in the sense that gives rise to maximal accuracy if her credences match the frequency of true propositions in the reference classes created by her credences.

> Vindication I shah explicate in terms of calibration …. Calibration plays the conceptual roles that truth, or empirical adequacy play in other contexts of discussion. (van Fraassen, 1983, 301)

If this is correct, we can use a divergence $\mathfrak{D}$ to define an inaccuracy measure as follows. On van Fraassen's and Shimony's account, we want to say that an agent's inaccuracy is how far she lies from being well calibrated. But of course there are many well calibrated credence functions for a given set of propositions. So we must determine from which well calibrated credence function we should take the divergence to $c$ in order to obtain a measure of how far $c$ lies from calibration (and thus the inaccuracy of $c$)? What we need is the notion of a *well calibrated counterpart*:

(p.60) **Definition 4.3.2 (Well calibrated counterpart)** *Suppose $c$ is a credence function defined on $\mathcal{F}$ and $w$ is a possible world relative to $\mathcal{F}$. Then the* well calibrated counterpart of $c$ at $w$ (*written $c^w$*) *is defined as follows:*

$$c^w(Z) := \frac{|\{X \in \mathcal{F} : c(X) = c(Z) \text{ and } v_w(X) = 1\}|}{|\{X \in \mathcal{F} : c(X) = c(Z)\}|}$$

*for each $Z$ in $\mathcal{F}$.*

Thus, the well calibrated counterpart of $c$ at $w$ assigns to a proposition $Z$ in $\mathcal{F}$ the proportion of propositions that are true at $w$ amongst all propositions in $\mathcal{F}$ to which $c$ assigns the same credence that it assigns to $Z$. In other words: the well calibrated counterpart of $c$ at $w$ is the unique credence function that is well calibrated at $w$ and which assigns the same credence to two propositions whenever $c$ does. For example: suppose my credence function $c$ is defined on three propositions, $G$ (*Noether is German*), $D$ (*Noether is Danish*), and $G \vee D$ (*Noether is*

*German or Danish*). Suppose I'm exactly as confident that Noether is German as I am that she's Danish; and just a little more confident that she is German or Danish. Thus, perhaps, $c(G) = c(D) = 0.3$ and $c(G \vee D) = 0.4$. Then, at the actual world, where $G$ and $G \vee D$ are true and $D$ is false, the well calibrated counterpart of my credence function $c$—namely, $c^{@}$,where @ is the actual world—assigns credences as follows:

- $c^{@}(G) = c^{@}(D) = 0.5$, since $c$ assigns the credence 0.3 only to $G$ and $D$, and exactly one out of these two propositions is true;
- $c^{@}(G \vee D) = 1$, since $c$ assigns the credence 0.4 only to $G \vee D$ and exactly one out of this one proposition is true.

Note that, if $c$ is well calibrated at $w$, then it is its own well calibrated counterpart—that is, $c^{w} = c$. Thus, we define the *calibration inaccuracy of c at w (relative to a divergence* $\mathfrak{D}$) to be the divergence from $c^{w}$ to $c$. And we write it $\mathfrak{C}_{\mathfrak{D}}$. That is,

$$C_{\mathrm{D}}(c, w) := \mathrm{D}(c^{w}, c)$$

Suppose $\mathfrak{D}$ is an additive divergence generated by the one-dimensional divergence $\mathfrak{d}$—as we argued above they must be, regardless of the account of vindication at play. Then $\mathfrak{C}_{\mathfrak{D}}$ $(c, w)$ is obtained by taking each proposition $X$ in $\mathscr{F}$, asking how far the credence $c(X)$ lies from the proportion of truths amongst all propositions to which $c$ assigns credence $c(X)$—where the distances here are measured by the one-dimensional divergence $\mathfrak{d}$—and summing together the results.

However, there are well-known problems with this account of inaccuracy. Here, briefly, are two; our main focus will be on a third, which we discuss below. See (Seidenfeld, 1985) and (Joyce, 1998) for further discussion of the problems facing calibration inaccuracy.

A measure of calibration inaccuracy is usually not continuous, even if the divergence that generates it is. To see this, consider the following example, where $\mathrm{F} = \{X, \overline{X}\}$ and $w$ is the world relative to $\mathscr{F}$ at which $X$ is true:

| | $X$ | $\overline{X}$ |
|---|---|---|
| $c$ | 0.5 | 0.5 |
| $c_{\varepsilon}$ | $0.5 + \varepsilon$ | $0.5 - \varepsilon$ |
| $c^{w}$ | 0.5 | 0.5 |
| $c_{\varepsilon}^{w}$ | 1 | 0 |

(p.61)

For all $0 < \varepsilon < 0.5$, the calibration inaccuracy of $c_\varepsilon$ is the divergence from $c^w_\varepsilon$ to $c_\varepsilon$. But that divergence will typically increase as $\varepsilon$ tends to 0. After all, the credences assigned to $X$ and $\overline{X}$ by $c_\varepsilon$ both move away from those assigned by $c^w_\varepsilon$ as $\varepsilon$ tends to 0. However, the calibration of c—which is the limit of $c_\varepsilon$ as $\varepsilon$ tends to 0—is 0. Thus, the calibration inaccuracy of a credence function at a world is not guaranteed to be a continuous function of that credence function.[11]

The second problem with the calibrationist account of inaccuracy is that, for many agents, it is far too easy to come by maximal calibration accuracy. To see this, consider again the example involving $c_\varepsilon$ and $c$ above. $c$ is well calibrated at both worlds, and thus has maximal calibration accuracy (minimal calibration inaccuracy) at both worlds, the world in which $X$ is true and the world in which $\overline{X}$ is true. Thus, an agent can be guaranteed to achieve maximal calibration accuracy without conducting any investigation of the world—it is knowable a priori that $c$ achieves it. And indeed this often happens when the set $\mathcal{F}$ over which $c$ is defined has a certain logical structure. If $\mathcal{F}$ is closed under negation, for instance, then a credence function that assigns credence 0.5 to every proposition is guaranteed to be well calibrated and thus have maximal calibration accuracy. If $\mathcal{F}$ is a partition, then a credence function that assigns $\frac{1}{|\mathbb{F}|}$ to each proposition in $\mathcal{F}$ is well calibrated at every world. Whatever accuracy is, it is not a virtue that can be gained so easily. Accuracy is something for which one strives over the course of an epistemic life. It is a goal such that one collects evidence in order to achieve it better. On the calibrationist account, accuracy is something many agents can acquire maximally at the beginning of their epistemic life, prior to acquiring any evidence.[12]

These two objections to the calibrationist account of accuracy are serious, but the third and final one I will consider is more illuminating for our purposes. Recall the first claim we made about accuracy in the Introduction to this book: the accuracy (p.62) of a credence in a truth is greater the higher it is; the accuracy of a credence in a falsehood is greater the lower it is. At the time, we noted that this does not give us enough information to compare the accuracies of every pair of credence functions. But it does allow us to do this in certain cases. For instance, suppose $c, c'$ are credence functions and $w$ is a world; and suppose that, for each proposition that is true at $w$, $c$ assigns a higher credence than $c'$, whereas for each proposition that is false at $w$, $c$ assigns a lower credence; then it must be that $c$ is less inaccurate than $c'$—after all, each of the credences assigned by $c$ is less inaccurate than the corresponding credence assigned by $c'$. Recall from Section 3.3, if this is true of a candidate inaccuracy measure—if it always assigns lower inaccuracy to one credence function than to another when each credence assigned by the first is more accurate than the corresponding credence assigned by the second—we say that the measure is *truth-directed*. Demanding that an inaccuracy measure has this property seems an extremely minimal and uncontroversial requirement. However, it is not satisfied by $\mathfrak{C}_{\mathfrak{D}}$, as we will now show.

Consider again the example of $c$ and $c_\varepsilon$ from above. Then:

- $c$ is well calibrated at $w$. After all, $c^w = c$. So $\mathfrak{C}_{??}(c, w) = \mathfrak{D}(c^w, c) = 0$.
- $c_\varepsilon$ is not well calibrated at $w$ (unless $\varepsilon = 0.5$). After all, $c_\varepsilon^w = v_w \neq c_\varepsilon$. So $C_{\mathrm{D}}(c_\varepsilon, w) = \mathrm{D}(c_\varepsilon^w, c_\varepsilon) > 0$.

So $c$ has greater calibration accuracy than $c_\varepsilon$ at $w$: that is, $\mathfrak{C}_{\mathfrak{D}}(c, w) < \mathfrak{C}_{\mathfrak{D}}(c_\varepsilon, w)$. But, at $w$, $c_\varepsilon$ assigns a higher credence to each truth than $c$ does, and a lower credence to each falsehood. Thus, $\mathfrak{C}_{\mathfrak{D}}$ is not truth-directed.

When Joyce considers calibration accuracy, he recognizes the force of the thought that the correct notion of accuracy for a set of credences should involve matching certain frequencies:

> What can it mean … to assign degree of belief $x$ to $X$ if not to think something like, 'Propositions like $X$ are true about $x$ proportion of the time'? (Joyce 1998)

However, in the end, he rejects $\mathfrak{C}_{\mathfrak{D}}$ as an inaccuracy measure because it is not truth-directed. In what follows, I hope to provide an account of accuracy that retains the intuition that motivates the calibrationist account of accuracy whilst also respecting the importance of having an inaccuracy measure that is truth-directed.

To be clear, I do not wish to retain the motivation for the calibrationist account that Joyce moots in the quoted passage. It is certainly too strong to say that it is constitutive of my having credence $x$ in $X$ that I think something like 'Propositions like X are true about $x$ proportion of the time'. For one thing, there are clear cases in which I quite reasonably assign a particular credence only to propositions in a given set of propositions whilst knowing that the frequency of truths amongst the propositions in that set is not given by that credence. For instance, I might know that Keith and Sara have the same colour of eyes: either both blue or both brown. Thus, I know that the frequency of truths amongst *Keith has blue eyes and Sara has blue eyes* is either 0 or 1. But I may still assign credence 0.5 to each. Thus, thinking that $x$ is a well calibrated (p.63) credence is neither constitutive of having that credence, nor rationally required by it. The motivating intuition for calibrationist accounts of accuracy that I would like to retain as far as possible is that credences are better the closer they are to being well calibrated. Thus, in Richard Jeffrey's terminology, according to this motivating intuition, a credence is not a guess of the frequency—after all, one incorrect guess is as inaccurate as any other, so guessing a frequency that isn't a possible frequency, such as in the case of Keith and Sara's eye colour, is guaranteed to be maximally inaccurate and thus always weakly dominated by a guess that is a possible frequency. Rather, a credence is an *estimate* of the frequency. According to Jeffrey, estimates are distinguished from guesses by the way in which we assess their success; that is, an estimate of a frequency is better the closer it is to the true frequency, so it can make sense to estimate a frequency that you know not to be a possible frequency (Jeffrey, 1986).

To retain this motivating intuition, we say that, while Joyce is right that calibration accuracy cannot be the whole story about accuracy because it is not truth-directed, it is nonetheless part of the story—calibration accuracy is a component of accuracy. The other component of accuracy, I claim, is directly motivated by the desideratum of truth-directedness. The calibration accuracy of a credence function at a world is given by the proximity of that credence function to its well calibrated counterpart at that world. Such a measure is not truth-directed because one credence function can be closer to its well calibrated counterpart than another is to its while the first well calibrated counterpart is further from the omniscient credence function than the second is. We see this in the example above: the credence function $c$ is closer to its well calibrated counterpart $c^w$ than the credence function $c_\varepsilon$ is to its well calibrated counterpart $c_\varepsilon^w$. But $c^w$ is further from the omniscient credence function $\nu_w$ than $c_\varepsilon^w$ is. So, intuitively, proximity to $c^w$ will result in greater inaccuracy than the same proximity to $c^w$. This suggests that we might preserve the intuition behind calibrationism and the intuition behind truth-directedness by saying that the accuracy of a credence function at a world is partly determined by its proximity to its well calibrated counterpart, but partly by the proximity of its well calibrated counterpart to the omniscient credence function. This retains both intuitions: the truth-directedness intuition is that proximity to the omniscient credence function is important; and the calibrationist intuition is that proximity to the well calibrated counterpart is important. We retain both by saying that a divergence can only generate a legitimate inaccuracy measure if the divergence from the omniscient credence function to a given credence function is determined by the divergence from the well calibrated counterpart of that credence function to the credence function itself and the divergence from the omniscient credence function to the well calibrated counterpart. That is, we impose the following condition on a divergence that generates an inaccuracy measure:

> **Decomposition** If an inaccuracy measure $\mathfrak{I}$ is generated by a divergence $\mathfrak{D}$, there are real numbers *α, β* such that:
>
> $$\mathrm{D}(v_w, c) = \alpha \mathrm{D}(c^w, c) + \beta \mathrm{D}(v_w, c^w)$$
>
> (p.64) That is, writing $\mathfrak{I}_{\mathfrak{D}}$ for $\mathfrak{I}$, we have:
>
> $$I_{\mathrm{D}}(c, w) = \alpha C_{\mathrm{D}}(c, w) + \beta I_{\mathrm{D}}(c^w, w)$$

By imposing this condition, we obtain a *ceteris paribus* version of the calibrationist's account of accuracy. Recall: On the calibrationist's account, you are more accurate the closer you are to your well calibrated counterpart. On the *ceteris paribus* version, *if other things are equal*, you are more accurate the closer you are to your well calibrated counterpart. In the presence of Decomposition, we now know exactly when other things are equal—it is when the inaccuracy of your well calibrated counterpart does not change. If the inaccuracy of your well calibrated counterpart does not change— that is, $\mathfrak{I}_{\mathfrak{D}}$ $(c^w, w)$ remains constant—then Decomposition entails that your inaccuracy $\mathfrak{I}_{\mathfrak{D}}$ $(c, w)$ varies with your distance from your well calibrated counterpart—that is, your calibration inaccuracy, $\mathfrak{C}_{\mathfrak{D}}$ $(c, w) = \mathfrak{D}$ $(c^w, c)$.

Let's look again at the example of a 10-ticket lottery. Let $L_i$ be the proposition *Ticket i will lose* and let $W_i$ be the proposition *Ticket i will win*. Let us suppose that I have credences only in these twenty propositions. And let us suppose that I have the same credence $x$ in each $L_i$ and I have the same credence $y$ in each $W_i$, but $x \neq y$. Then, amongst credence functions of this sort, Decomposition entails that the most accurate will be the well calibrated counterpart that they all share, namely, the credence function that assigns 0.1 to each $W_i$ and 0.9 to each $L_i$. Of course, there are more accurate credence functions available, but they do not assign the same credences to each $W_i$ and to each $L_i$. Of those credence functions, the one mentioned is most accurate. Using this observation, I think we can give an error theory for the intuitions that drive Ramsey, van Fraassen, and Shimony to propose the calibrationist account of accuracy. When we think about accuracy and ask which credence functions seem more accurate than which others, we often compare credence functions that have the same well calibrated counterpart; and we notice that the ones that we take to be most accurate are those that are closest to that well calibrated counterpart. This leads us to think that all that matters for accuracy is proximity to the well calibrated counterpart— after all, in these comparisons, that is all that matters! Doing so, we fail to notice that the proximity of the well calibrated counterpart to the omniscient credence function also plays a role. We can see this in the quotation from Ramsey above. He assumes that we assign the same credence to all propositions concerning the unwholesomeness of particular yellow toadstools—'Granting that [an agent] is going to think always in the same way about all yellow toadstools'. If we assume that, and compare only credence functions that satisfy that assumption, we compare credence functions that share the same well calibrated counterpart. This leads us to take proximity to that well calibrated counterpart to be the only component of accuracy. But, as we saw above, that leads to a measure of inaccuracy—the calibrationist's measure—that is not truth-directed. Decomposition saves as much of this intuition as possible, recovering a *ceteris paribus* version of it. I submit that this justifies Decomposition, the final condition we will impose on measures of inaccuracy.

(p.65) Thus, we impose the following five conditions on an inaccuracy measure:

- **Alethic Vindication** The omniscient credence function at a world is the ideal credence function to have at that world. Thus, $\nu_w$ is the ideal credence function at $w$.
- **Perfectionism** If $\mathfrak{I}$ is a legitimate inaccuracy measure, there is a divergence $\mathfrak{D}$ such that $\mathrm{I}(c, w) = \mathrm{I}_{\mathrm{D}}(c, w) = \mathrm{D}(v_w, c)$. We say that $\mathfrak{D}$ generates $\mathfrak{I}$.
- **Divergence Additivity** If $\mathfrak{I}$ is a legitimate inaccuracy measure generated by $\mathfrak{D}$, there is a one-dimensional divergence $\mathfrak{d}$ such that $\mathrm{D}(c, c') = \sum_{X \in \mathrm{F}} \mathrm{d}(c(X), c'(X))$. We say that $\mathfrak{d}$ generates $\mathfrak{D}$.
- **Divergence Continuity** If $\mathfrak{I}$ is a legitimate inaccuracy measure generated by an additive divergence $\mathfrak{D}$ that is generated by $\mathfrak{d}$, then $\mathfrak{d}$ is continuous in its first and second argument.
- **Decomposition** If $\mathfrak{I}$ is a legitimate inaccuracy measure generated by a divergence $\mathfrak{D}$, then there are $\alpha, \beta$ such that
$$\mathrm{D}(v_w, c) = \alpha \mathrm{D}(c^w, c) + \beta \mathrm{D}(v_w, c^w)$$

The following theorem shows that if these five assumptions all hold, then every legitimate inaccuracy measure is generated by a certain sort of divergence called an *additive Bregman divergence* (see Appendix to Part I for the definition). These divergences are well known to statisticians.

**Theorem 4.3.3** *Suppose Alethic Vindication, Perfectionism, Divergence Additivity, Divergence Continuity, and Decomposition. Then, if $\mathcal{I}$ is a legitimate inaccuracy measure, there is an additive Bregman divergence $\mathfrak{D}$ such that $\mathfrak{I}(c, w) = \mathfrak{D}(\nu_w, c)$.*

The following theorem then shows that we can mount an accuracy-based argument for Probabilism on the basis of this characterization of the legitimate inaccuracy measures together with the decision-theoretic principle of Immodest Dominance. The theorem is closely related to the main theorem of (Predd et al., 2009).

**Theorem 4.3.4** *Suppose $\mathcal{F}$ is a finite opinion set, $\mathfrak{D}$ is an additive Bregman divergence, and $\mathfrak{I}(c, w) = \mathfrak{D}(\nu_w, c)$. Then*

(i) *Each non-probabilistic credence function is strongly $\mathcal{I}$-dominated by a probabilistic credence function.*
*That is: If $c \notin \mathcal{P}_{\mathcal{F}}$, then there is $p \in \mathcal{P}_{\mathcal{F}}$ such that, for all worlds $w \in \mathcal{W}_{\mathcal{F}}$, $\mathfrak{I}(c, w) < \mathfrak{I}(p, w)$.*
(ii) *Every probabilistic credence function is not even moderately $\mathcal{I}$-modest.*
That is: If *$p \in \mathcal{P}_{\mathcal{F}}$ and $c \in \mathcal{B}_{\mathcal{F}}$ and $p \neq c$, then*

$$\mathrm{Exp}_I(p|p) < \mathrm{Exp}_I(c|p)$$

*where* $\mathrm{Exp}_J(c|p) := \sum_{w \in W_F} p(w) J(c, w)$.

(p.66) One might wonder how this characterization of the legitimate inaccuracy measures compares to others in the literature. In fact, the answer is straightforward: the inaccuracy measures generated by additive Bregman divergences in this way are exactly the so-called *additive and continuous strictly proper inaccuracy measures*. Thus:

**Theorem 4.3.5** *Suppose $\mathcal{I}$ is an inaccuracy measure. Then the following two propositions are equivalent*:

(1) *There is an additive Bregman divergence $\mathfrak{D}$ such that $\mathfrak{I}(c, w) = \mathfrak{D}(\nu_w, c)$*
(2) *There is a scoring rule $\mathfrak{s} : \{0,1\} \times [0,1] \to [0, \infty]$ such that*
 (a) *$\mathfrak{s}(i, x)$ is a continuous function of $x$.*
 (b) *$\mathfrak{s}$ is strictly proper. That is, for all $0 \leq p \leq 1$,*

$$p\mathrm{s}(1, x) + (1 - p)\mathrm{s}(0, x)$$

 *is uniquely minimized as a function of $x$ at $x = p$.*
 (c) $I(c, w) = \sum_{X \in F} \mathrm{s}(v_w(X), c(X))$

*If $\mathfrak{I}$ satisfies (2), we say that it is an additive and continuous strictly proper inaccuracy measure*.

Before we give the accuracy-based argument for Probabilism that is based on this characterization together with Immodest Dominance, there is some unfinished business to deal with in the following chapter. We must address an objection to the accuracy-based argument for Probabilism that arises if our characterization of the legitimate inaccuracy measures permits two or more inaccuracy measures that disagree on which credence functions accuracy-dominate a given credence function. However, before that, I'd like to ask whether we can in fact narrow the class of legitimate inaccuracy measures so that it doesn't permit two or more such measures. In fact, I'd like to consider a way of narrowing it down so that only the Brier score and linear transformations of it are legitimate: since exactly the same credence functions dominate a given credence function regardless of which of these measures we use, the problem that we will address in the next chapter will not arise if we can narrow the field in this way.

## 4.4 Symmetry

Perfectionism says that the inaccuracy of a credence function at a world is given by the distance between it and the ideal credence function at that world. This clearly requires a measure of the distance between two credence functions. At the beginning of this chapter, we assumed that this measure of distance is given by a divergence. That is, there is a divergence $\mathfrak{D}$ such that the inaccuracy of $c$ at $w$ is the divergence from the ideal credence function at $w$ to $c$—that is, $\mathfrak{D}(i_w, c)$. Vindication then identifies $i_w$ while Divergence Addivity, Divergence Continuity, and Decomposition narrow down the candidates for $\mathfrak{D}$. However, as we noted above, divergences are not necessarily (p.67) symmetric: that is, there are divergences $\mathfrak{D}$ and credence functions $c, c'$ such that $\mathfrak{D}(c, c') = \mathfrak{D}(c', c)$. Moreover, none of the further conditions that we imposed above entail symmetry: that is, there are non-symmetric divergences that satisfy Divergence Addivity, Divergence Continuity, and Decomposition.

Having noticed this, we might wish to impose symmetry as a further condition on the divergences that generate inaccuracy measures. After all, it seems strange that there might be two credence functions such that the distance of the first from the second is different from the distance of the second from the first. Thus, we may wish to impose the following condition:

**Symmetry** If $\mathfrak{I}$ is a legitimate inaccuracy measure generated by a divergence $\mathfrak{D}$, then $\mathfrak{D}$ is symmetric: that is, $\mathfrak{D}(c, c') = \mathfrak{D}(c', c)$.

The following theorem shows how powerful this condition is:

**Theorem 4.4.1** *Suppose Alethic Vindication, Perfectionism, Divergence Additivity, Divergence Continuity, Decomposition, and Symmetry. Then, if $\mathfrak{I}$ is a legitimate inaccuracy measure, then $\mathfrak{I}$ is the Brier score or some linear transformation of it.*

That is, there is only one symmetric additive Bregman divergence and it is Squared Euclidean Distance, the divergence that gives rise to the Brier score. Thus, Alethic Vindication, Perfectionism, Divergence Additivity, Divergence Continuity, Decomposition, and Symmetry together characterize the Brier score uniquely (up to positive linear transformation).

How compelling is Symmetry? We might worry that its intuitive force derives from an over-reliance on the analogy with spatial distance. Intuitively, spatial distance measures must be symmetric: indeed, as we noted above, when spatial distance is formalized in mathematics, it is as a metric, which is assumed to be symmetric. But why think that a measure of distance between two credence functions must have the same abstract properties as a measure of distance between two spatial points?

In fact, I think that the analogy with spatial distance is not the source of our strong inclination towards Symmetry. Rather, I think, we reason to Symmetry as follows: We have a strong intuition that the inaccuracy of an agent's credence function at a world is the distance between that credence function and the ideal credence function at that world. But we have no strong intuition that this distance must be the distance *from* the ideal credence function *to* the agent's credence function rather than the distance *to* the ideal credence function *from* the agent's credence function; nor have we a strong intuition that it is the latter rather than the former. But if there were non-symmetric divergences that gave rise to measures of inaccuracy, we would expect that we would have intuitions about this latter question, since, for at least some accounts of the ideal credence function at a world and for some agents, this would make a difference to the inaccuracies to which such a divergence gives rise. Thus, there cannot be such divergences. Symmetry follows.

(p.68) I think this reasoning is sound, and I will assume Symmetry at some points in the remainder of the book. But it is worth noting that it is not always required. As I mentioned above, in the next chapter, I consider an objection to the accuracy-based argument for Probabilism that arises if we permit two or more inaccuracy measures that disagree on which credence functions accuracy-dominate which others. If we accept Symmetry, and thereby narrow the field of legitimate inaccuracy measures to linear transformations of the Brier score, this problem is solved, since all of these inaccuracy measures agree on the accuracy-domination facts. So, in my final version of the accuracy argument for Probabilism, I will assume Symmetry. The same goes for the accuracy argument for the Principal Principle that I give in Part II. On the other hand, the arguments in Part III for the Principle of Indifference and some of its variations do not require us to assume Symmetry. The arguments in Part IV for Conditionalization are a mixed bag—some arguments require Symmetry; others don't. I will make clear in all cases what we are assuming.

In any case, at the end of this chapter, we have at least narrowed down the divergences that generate legitimate inaccuracy measures to the additive Bregman divergences; and we have thereby narrowed down the legitimate inaccuracy measures to the continuous and additive strictly proper ones. We have also argued that the divergences must also be symmetric; and we have seen that this narrows down the legitimate inaccuracy measures to just the positive linear transformations of the Brier score.

Notes:

([1]) Such a demand would rule out squared Euclidean distance: while Euclidean distance is a metric and satisfies the triangle inequality, squared Euclidean distance is not and does not. Furthermore, the triangle inequality is only intuitively appealing when we are measuring physical distance. In those cases, its intuitive appeal arises from the thought that the distance between two points is the length of the shortest path between them (together with the thought that the length of a path that divides into two parts is obtained by summing the length of the first path with the length of the second path). If the distance from point $x$ to point $z$ were longer than the sum of the distance from $x$ to y and from $y$ to $z$, then the distance from $x$ to $z$ could not be the length of the shortest path from $x$ to $z$, since the path that runs first from $x$ to $y$ and then from $y$ to $z$ would be shorter.

([2]) Note that Additivity and Additivity* entail the Strong Extensionality axiom that we considered in Section 3.3 above.

([3]) See Buchak (2014a,b) for a well worked out non-expected utility theory as well as a clear survey of the various alternatives. Note, however, that Buchak allows that a utility function might reflect some risk aversion: she allows, for instance, that money might be a dependent good and that this might account for some of the risk-averse behaviour we witness. But she also holds that much rational risk averse behaviour must be accounted for by appealing to non-expected utility decision rules as well. In Part IV, we will consider the consequences of risk-sensitive decision principles when the options are credence functions and utility is accuracy.

([4]) Note here that we assume only that the range of $\mathfrak{d}$ is contained in $[0, \infty]$, which is the set of nonnegative real numbers together with infinity. Thus, we reject the Finiteness axiom of Section 3.2. We allow divergences and thus inaccuracy measures to take an infinite value.

([5]) See, for instance, (Foley 1993), (Kyburg 1961), (Kyburg 1970), Fitelson (ms).

([6]) For a fuller treatment of this problem, see (Pettigrew, 2015).

([7]) Cf. (Pettigrew, 2015, Theorem 3).

([8]) For an account of full beliefs that explicitly takes them to be *sui generis* states, see Leitgeb (2015), Fitelson (ms).

([9]) Let us indulge in the fiction that these predicates are not vague and are, in fact, coextensive with these definitions.

([10]) $\mathrm{ran}(c)$ is the range of $c$. It is the set ofvalues that $c$ takes. Thus, $\mathrm{ran}(c) := \{x : (\exists X \in \mathscr{F})[c(X) = x]\}$.

([11]) How is this possible when we explicitly assumed that the divergences that generate inaccuracy measures are continuous in their second argument in order to ensure that inaccuracy measures are continuous regardless of the notion of vindication that we assume? The answer is that the continuity of the divergence guarantees the continuity of the inaccuracy measure only if the vindicated credence function at a world depends only on the world itself and not on anything else. In the case of calibration accuracy, that isn't the case. When we measure the inaccuracy of $c$ at $w$, we measure the divergence from the ideal credence function $c^w$, which depends not only on the world $w$ *but also on* $c$.

([12]) This complaint has the same structure as Goldman's objection to pure evidentialism. Goldman points out that, if responding appropriately to evidence is the goal of belief, it is something one can achieve perfectly without acquiring any evidence as long as one adopts whatever is the appropriate response to a lack of evidence. In both cases, the objection turns on the following claim: whatever is the goal of belief, it is not something that is so easily achieved (Goldman, 2002, Sections 3, 6).