Let’s use “+1” to denote best epistemic status, “−1” to denote worst epistemic status, and “0” to denote middling epistemic status. Our simplest, 2-valued scheme is:

<table>
<thead>
<tr>
<th></th>
<th>P</th>
<th>Q</th>
<th>Pennes</th>
<th>Q pne</th>
<th>Qqens</th>
</tr>
</thead>
<tbody>
<tr>
<td>P₁</td>
<td>T</td>
<td>T</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
</tr>
<tr>
<td>P₂</td>
<td>T</td>
<td>F</td>
<td>+1</td>
<td>−1</td>
<td>−1</td>
</tr>
<tr>
<td>P₃</td>
<td>F</td>
<td>T</td>
<td>−1</td>
<td>+1</td>
<td>−1</td>
</tr>
<tr>
<td>P₄</td>
<td>F</td>
<td>F</td>
<td>−1</td>
<td>−1</td>
<td>+1</td>
</tr>
</tbody>
</table>

If we’re going to use only 2-values (“correct/incorrect”), then it seems to me that this scheme is forced on us, by (†).

But, one might think that a 3-valued scheme makes more sense. David Christensen makes the following observation.

Suppose I’m going to flip a coin. Can I rationally be indifferent between heads (H) and tails (T)? It seems that H ≃ₜ T would be dominated by H >ₜ T (or T >ₜ H), since H ≃ₜ T is guaranteed to be “incorrect” and the latter aren’t.

Christensen is right. And, he suggests a 3-valued scheme.
**Theorem.** No 2 or 3-valued scoring scheme is such that:

1. \( S \) entails (at least some instances of) both transitivity and additivity as (weak) dominance norms.

and, the following eight (8) scoring desiderata are met:

1. Having a subset of judgments \( \{ p \succ_S q, p \succ_S r, q \sim_S r \} \) should not — in and of itself — ensure “incoherence”.
2. Ditto for subsets of the form \( \{ p \succ_S q, p \succ_S r, q \succ_S r \} \).
3. \( p \succ_S q \) should get a “worst” score when \( p \) is F and \( q \) is T.
4. \( p \succ_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) and \( q \) are both F.
5. \( p \sim_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) and \( q \) are both F.
6. \( p \sim_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) is F and \( q \) is T.
7. The score of \( p \succ_S q \) when \( p \) is T and \( q \) is F should not be strictly worse than the score of \( p \succ_S q \) when \( p, q \) are both T.
8. The score of \( p \succ_S q \) when \( p \) is T and \( q \) is F should be strictly better than the score of \( p \succ_S q \) when \( p \) is F and \( q \) is T.

**Extra.** Let’s suppose (arguendo) that \( S \) has a numerical credence function \( b : B \rightarrow \mathbb{R} \) (these \( b \)'s are opinionated, of course, and so we’re ignoring suspension of judgment here, once again).

- As usual, we need to settle on a way of scoring \( b \)'s for inaccuracy at each possible world \( w \) — call this \( I(b, w) \).
- For simplicity, I’ll assume \( I(b, w) \) is an additive function, which sums-up the inaccuracies of \( b \), for each \( p \in B \) at \( w \).
- If we associate the number 1 with T and the number 0 with F (at each world \( w \)), then the inaccuracy of \( b(p) \) at world \( w \) will be \( b \)'s “distance (\( d \)) from the 0/1-truth-value of \( p \)” at \( w \).

**Example.** Suppose \( S \) has just two (contingent) propositions \( \{ P, \sim P \} \) in their doxastic space. Then, there are two salient possible worlds (\( w_1 \) in which \( P \) is T, and \( w_2 \) in which \( P \) is F).

- \( I(b, w_1) = d(b(P), 1) + d(b(\sim P), 0) \).
- \( I(b, w_2) = d(b(P), 0) + d(b(\sim P), 1) \).

**Theorem.** No 2 or 3-valued scoring scheme is such that:

1. \( S \) entails (at least some instances of) both transitivity and additivity as (weak) dominance norms.

and, the following eight (8) scoring desiderata are met:

1. Having a subset of judgments \( \{ p \succ_S q, p \succ_S r, q \sim_S r \} \) should not — in and of itself — ensure “incoherence”.
2. Ditto for subsets of the form \( \{ p \succ_S q, p \succ_S r, q \succ_S r \} \).
3. \( p \succ_S q \) should get a “worst” score when \( p \) is F and \( q \) is T.
4. \( p \succ_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) and \( q \) are both F.
5. \( p \sim_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) and \( q \) are both F.
6. \( p \sim_S q \) should get the same score when \( p \) and \( q \) are both T as it does when \( p \) is F and \( q \) is T.
7. The score of \( p \succ_S q \) when \( p \) is T and \( q \) is F should not be strictly worse than the score of \( p \succ_S q \) when \( p, q \) are both T.
8. The score of \( p \succ_S q \) when \( p \) is T and \( q \) is F should be strictly better than the score of \( p \succ_S q \) when \( p \) is F and \( q \) is T.

**Extra.** Let’s suppose (arguendo) that \( S \) has a numerical credence function \( b : B \rightarrow \mathbb{R} \) (these \( b \)'s are opinionated, of course, and so we’re ignoring suspension of judgment here, once again).

- As usual, we need to settle on a way of scoring \( b \)'s for inaccuracy at each possible world \( w \) — call this \( I(b, w) \).
- For simplicity, I’ll assume \( I(b, w) \) is an additive function, which sums-up the inaccuracies of \( b \), for each \( p \in B \) at \( w \).
- If we associate the number 1 with T and the number 0 with F (at each world \( w \)), then the inaccuracy of \( b(p) \) at world \( w \) will be \( b \)'s “distance (\( d \)) from the 0/1-truth-value of \( p \)” at \( w \).

**Example.** Suppose \( S \) has just two (contingent) propositions \( \{ P, \sim P \} \) in their doxastic space. Then, there are two salient possible worlds (\( w_1 \) in which \( P \) is T, and \( w_2 \) in which \( P \) is F).

- \( I(b, w_1) = d(b(P), 1) + d(b(\sim P), 0) \).
- \( I(b, w_2) = d(b(P), 0) + d(b(\sim P), 1) \).
• **Theorem** (de Finetti). \( b \) is non-probabilistic if and only if there exists a probabilistic credence function \( b' \) such that (a) \( b' \) has a strictly lower Brier Score than \( b \) at some worlds, and (b) \( b' \) never has a greater Brier Score than \( b \) at any world.

• And, the proof of de Finetti’s theorem is constructive — it tells us precisely which functions \( b' \) “Brier-dominate” \( b \).

Joyce [6, 5] uses de Finetti’s Theorem (and generalizations of it) to ground an (epistemic) probabilistic coherence norm. (PC) \( S \)'s credences \( b \) should be probabilistic — on pain of being Brier-dominated by (specific) credence functions \( b' \).

Because Joyce thinks that Brier Score is a good measure of “credal inaccuracy”, he thinks this provides incoherent agents with some “epistemic reason” to be Pr-coherent.

Maher [10] points out that other prima facie plausible measures of “inaccuracy” do not undergird (PC). I’ll return to that issue below. But, first, a concrete toy example.

Suppose \( S \) adopts the Brier Score as their \( I \)-measure, and that \( S \)'s \( b \) is non-probabilistic. Then, there are alternative (coherent) credence functions \( b' \) that accuracy-dominate \( b \).

Intuitively, these \( b' \) functions should “look epistemically better” (in a precise sense) than \( S \)'s current credences \( b \).

But, our “evidentialist” (“Kolodny’s revenge”) worry lingers.

Consider a very simple toy agent \( S \) with one sentence \( P \) in their language. And, suppose \( S \)'s credence function assigns \( b(P) = 0.2 \) and \( b(\neg P) = 0.7 \). So, \( S \)'s \( b \) is non-probabilistic.

It follows from de Finetti/Joyce’s theorems that there is a specific set of credence functions \( b' \) that Brier-dominate \( b \).

The figure on the next slide depicts this situation. The red dot is \( S \)'s credence function \( b \). And, the shaded region depicts the credence functions \( b' \) that Brier-dominate \( b \).

[The black dot at \( (0.2, 0.8) \) depicts the only probabilistic credence function that is compatible with \( b(P) = 0.2 \).]

I don’t have the space to delve into the various other worries I have about Joyce’s argument(s) for probabilism. [But, in my lecture next week, I will discuss another worry.]

For now, I have a suggestion re the quantitative case.

Based on our experience from the qualitative and comparative cases, we should not expect an AD-justification of the full probabilistic norm(s) in the quantitative case...

Rather than trying to “justify” the use of \( s \) (or some other “distance measure” that yields the full probabilistic norms), why not start with desiderata for distance measures \( d \)? E.g.,

\[
\begin{align*}
&d(x, x) = 0, \\
&d(x, y) = d(y, x), \\
&d(x, y) \leq d(x, z) + d(z, y).
\end{align*}
\]

Once we settle on desiderata \( D \) for adequate measures of distance (in this context), then we could ask the following: (Q) What accuracy-dominance norms are entailed by \( D \)?

In other words, (Q) is asking what accuracy-dominance norms are agreed upon by all inaccuracy measures \( I_d(b, w) \), where all we assume about \( d \) is that it satisfies desiderata \( D \).

I don’t have an answer to (Q). But, I conjecture that this will lead to norms for \( b \) that are similar to those we saw in the comparative case — e.g., if \( p = q \), then \( b(p) \leq b(q) \), etc.

Idea: start with \( s(x, y) \) and Maher's \( S(x, y) = |x - y| \).
### References


---

- If $S$ violates **Monotonicity** (4), then $S$ is accuracy-dominated.

  (4) If $p$ entails $q$, then $p \succ_S q$.

<table>
<thead>
<tr>
<th>$w_1$</th>
<th>T</th>
<th>$Q \succ_S P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w_2$</td>
<td>T</td>
<td>$B$</td>
</tr>
<tr>
<td>$w_3$</td>
<td>F</td>
<td>$C$</td>
</tr>
<tr>
<td>$w_4$</td>
<td>F</td>
<td>$B$</td>
</tr>
</tbody>
</table>

- Indeed, as this table shows, any scoring scheme that satisfies our desiderata [viz., $(\dagger) \Rightarrow A < C$] entails Monotonicity.

- To see that de Finetti’s additivity axiom (3) does not have a dominance justification, one must look at all the possible ways of “fixing” a violation of (3), and show that none of these lead to a comparison set that dominates the original.

- There aren’t that many cases to check. [I won’t show them.]

- On the next slides, I’ll discuss the Scott Axiom...

---

#### Definition

For each state description $s$ and each sequence $(n$-tuple) of propositions $Z = \{z_1, \ldots, z_n\} \in \prod_n B$, let $c(s, Z)$ be the number of elements of $Z$ that are entailed by $s$.

#### OK, here’s the (dreaded) Scott Axiom:

(SA) Let $X, Y \in \prod_n B$ be (arbitrary) sequences of propositions, each having length $n > 0$. Let $(x_1, \ldots, x_n)$ denote the members of $X$, and $(y_1, \ldots, y_n)$ denote the members of $Y$.

**If** the following two conditions are satisfied

i. For every state description $s$, $c(s, X) = c(s, Y)$.

ii. For all $i \in (1, n)$, $x_i \succ_S y_i$.

**then**, the following must also be the case

iii. $y_1 \succ_S x_1$.

- Not only is (SA) *unintuitive*, it is also *quite strong*. It entails both de Finetti’s “additivity” (3) and (full) transitivity of $\succ_S$.

---

#### Theorem (Fishburn)

(SA) is true if and only if there exists a mass function $m$ on $B$ such that, for all propositions $p$ and $q$ in $B$, the following real-valued representation holds:

\[
(*) \quad p \succ_S q \text{ if and only if } \sum_{s \models p} m(s) > \sum_{s \models q} m(s).
\]

And, given de Finetti’s axiom (2), there will always be a probability mass function $m$ satisfying $(*)$.

Fishburn’s Theorem reveals that (SA) alone ensures a real-valued representation ($R\succ_S$) of the $\succ_S$-ordering.

Not only does this imply de Finetti’s additivity axiom (3), but it also implies axiom (1) as well ($\succ_R$ is a strict total order).

Thus, once we have (SA) on board, the only axiom of de Finetti that can do any work is his axiom (2), which just ensures that $R\succ_S$ is a probabilistic representation of $\succ_S$.