# Degrees of Incoherence, Dutch Bookability & Guidance Value

Jason Konek

## 1  Introduction

What's so good about having probabilistically coherent credences? Some answer this question by showing that incoherent credences are needlessly bad at playing some key functional or theoretical role, *e.g.*, guiding action or encoding accurate truth-value estimates. Coherent credences, in contrast, lack this defect. For example, de Finetti (1974), building on the seminal work of Ramsey (1931), showed that incoherent credences lead to *sure loss*. That is, for any agent with incoherent credences there is some collection of gambles that (i) the agent is required to judge fair even though (ii) jointly those gambles are *guaranteed* to result in negative utility. That seems foolish. Coherent credences never lead to such foolishness. This seems to show that incoherent credences are needlessly bad at guiding action.

Joyce (1998, 2009) and Pettigrew (2016), in contrast, show that incoherent credences are needlessly bad at playing a key *epistemic* role. In particular, they encode truth-value estimates that are *accuracy-dominated*. More carefully, for any agent with incoherent credences and any reasonable measure of accuracy, there is some coherent set of credences that encode truth-value estimates that are *guaranteed* to be strictly more accurate. Coherent credences are never accuracy-dominated in this way. This seems to show that incoherent credences are needlessly bad at "representing the world" (encoding accurate truth-value estimates).[1]

There are, of course, other answers to the question that kicked us off. For example, Koopman (1940b,a), Good (1950), and Krantz et al. (1971) all treat credences as mere numerical measures of *comparative belief*. Paired with appropriate

[1]Joyce (1998, 2009) and Pettigrew (2016) build on the work of de Finetti (1974), Savage (1971) and many others. De Finetti showed that avoiding sure loss is equivalent to avoiding Brier score dominance. That is, credences avoid sure loss if and only if there is no other set of credences that is *guaranteed* to incur a strictly smaller loss when penalized by the Brier score (*i.e.*, mean squared error). (Hence they avoid both defects if and only if they are probabilistically coherent.) While de Finetti thought of these penalties as *practical*—the agent is docked in some quantity with linear utility (lottery tickets perhaps)—Joyce and Pettigrew think of these penalties as *epistemic*. The Brier score and other strictly proper scoring rules matter for their purposes because they are reasonable measures of *accuracy* and accuracy is the principal determinant of epistemic value.

representation theorems, this view entails that an agent has probabilistically coherent credences just in case her comparative beliefs satisfy various rationality constraints.[2,3] You better have probabilistically coherent credences, then, lest you violate these constraints.

Julia Staffel, in her excellent book *Unsettled Thoughts*, takes the on/off, binary question that we started with and makes it gradational: Why is it good to be *less*, rather than *more* incoherent? She argues that the functional-role-style answer—incoherent credences are needlessly bad at playing a key functional/theoretical role—extends naturally to this gradational variant. If your credences are incoherent, then there is some way of nudging them toward coherence that is guaranteed to make them more accurate *and* reduces the extent to which they are Dutch-bookable. This seems to show that such nudging makes your credences better at both representing the world and guiding action. This is crucially important for non-ideal agents like us. We will always fall short of perfect coherence however much we nudge. If all of the reward is on the other side of the finish line and that finish line is too far away to cross, then why try? If a miss by an inch is no better than a miss by a mile, then why waste the effort to miss by an inch? Staffel argues that, luckily for us, we are not in this unfortunate predicament. There's a path toward the finish line (*i.e.*, probabilistic coherence) that allows us to pick up increasing shares of the reward (*i.e.*, practical and epistemic value) as we move along it.

As promising as this sounds, I will argue that Staffel's answer to the gradational question needs some tweaking. Staffel's general strategy is as follows:

> First, identify some value that justifies some requirement of rationality, in the sense that complying with the requirement best promotes this value, and violating the requirement precludes optimal promotion of the value. Then identify a way in which this value can be had to greater or lesser degrees. Lastly, select a distance measure or divergence such that approximating ideal compliance with the requirement of rationality delivers increasing amounts of the value. (Staffel, 2019, p. 94)

The basic problem is this. Staffel identifies a plausible class of epistemic value measures, *viz., accuracy* measures, but not practical value measures. Accuracy may well be the cardinal epistemic good-making feature of credences. (So say veritists, anyway.) So numerical measures of accuracy may well be fit to serve as epistemic value/utility functions. But on the practical end, Staffel focuses on measures of Dutch bookability. The problem: degree of immunity to Dutch books is definitely *not* the cardinal practical good-making feature of credences. Susceptibility to a Dutch book is a practical defect, no doubt. It is a surefire *sign* that one's credences are needlessly bad at guiding action. But one's degree of immunity to Dutch books is not *itself* a good measure of how well your credences guide action. Indeed, credences that are *less* immune from Dutch books are often *better* at guiding action. Hence numerical measures of Dutch bookability are not fit to serve as measures of the practical value of one's credences.

Luckily, all Staffel's strategy needs is a *small* tweak. While Joyce (1998, 2009), Pettigrew (2016) and others argue that strictly proper scoring rules

---

[2] *Cf.* Konek (2019).

[3] Cox (1961) takes a similar approach. He shows that (conditional) credences that satisfy some plausible axioms fully agree with a probability function. *Cf.* Joyce (2009).

are reasonable measures of accuracy, Schervish (1989), Levinstein (2017) and Pettigrew (2020) show that they are also reasonable measures of guidance value. In section 3, I will exploit this fact to provide a new version of Staffel's argument. I will show that if the scoring rules that measure accuracy and guidance value are "sufficiently similar", then there will be some way of approximating coherence that is guaranteed to yield more of both. The upshot: we can retain Staffel's general strategy and fix the basic problem.

## 2 The Problem

Consider an example from Easwaran and Fitelson (2012) and Joyce (2013):

> Imagine a believer, Joshua, who has opinions about whether a certain coin will come up heads or tails when next tossed, and who also has evidence about the coin's bias. We may think of Joshua's credences as assigning real numbers to atomic events $[\pm H \& ch(\pm H) = x]$, where $\pm H$ might be $H$ or $\neg H$ and where $ch(\pm H) = x$ says that the coin's objective chance of landing $\pm H$ is $x \in [0, 1]$. Let's suppose further that Joshua knows that the coin's bias toward heads is 0.2, so that $b(ch(H) = 0.2) = 1$, and that this is all the relevant evidence he has about the coin. (Joyce, 2013, p. 9)

Suppose that Joshua's credence for $H$ is 0.2 (*i.e.*, $b(H) = 0.2$) and his credence $\neg H$ is also 0.2 (*i.e.*, $b(\neg H) = 0.2$). So Joshua's credences are probabilistically incoherent. Now, if Joshua had these incoherent credences but knew nothing about the bias of the coin (and had no other information suggesting one credence deserves more revision than another), then Staffel suggests that "moving directly towards the closest coherent credence assignment seems like a reasonable way to improve coherence" (Staffel, 2019, p. 84). For concreteness, let's measure inaccuracy by the Brier score, $\mathcal{B}$:

$$\mathcal{B}(b, H) = (1 - b(H))^2 + (0 - b(\neg H))^2 \tag{1}$$

$$\mathcal{B}(b, \neg H) = (0 - b(H))^2 + (1 - b(\neg H))^2 \tag{2}$$

And let's measure the proximity from one credence function $b$ to another $c$ by the "Bregman divergence" associated with the Brier score, $\mathcal{D}_{\mathcal{B}}$, which is just squared Euclidean distance:

$$\mathcal{D}_{\mathcal{B}}(b, c) = (b(H) - c(H))^2 + (b(\neg H) - c(\neg H))^2 \tag{3}$$

The closest coherent credence assignment to Joshua's when measuring closeness using squared Euclidean distance is:

$$c(H) = 0.5, c(\neg H) = 0.5 \tag{4}$$

And moving "directly towards" $c$ just means adopting some convex combination of $b$ and $c$:

$$b_\lambda(H) = \lambda c(H) + (1 - \lambda)b(H), b_\lambda(\neg H) = \lambda c(\neg H) + (1 - \lambda)b(\neg H) \tag{5}$$

So Joshua could nudge toward coherence, for example, by first adopting a credence of 0.21 for $H$ and $\neg H$, respectively, and then adopting a credence of

0.22, and so on. This seems reasonable, according to Staffel, for a number of reasons. Firstly, as you nudge $b$ directly toward $c$, you end up with credences that are *guaranteed* to be increasingly accurate (*i.e.*, have a strictly lower Brier score):

|  | $\mathcal{B}(b, H)$ | $\mathcal{B}(b, \neg H)$ |
|---|---|---|
| $b_0(H) = b_0(\neg H) = 0.2$ | 0.68 | 0.68 |
| $b_{0.0333}(H) = b_{0.0333}(\neg H) = 0.21$ | 0.6682 | 0.6682 |
| $b_{0.0666}(H) = b_{0.0666}(\neg H) = 0.22$ | 0.6568 | 0.6568 |
| $\vdots$ | $\vdots$ | $\vdots$ |
| $b_{0.9666}(H) = b_{0.9666}(\neg H) = 0.49$ | 0.5002 | 0.5002 |
| $b_1(H) = b_1(\neg H) = 0.5$ | 0.5 | 0.5 |

On top of this, nudging $b$ directly toward $c$ reduces the extent to which Joshua's credences are Dutch-bookable according to a number of measures of Dutch book vulnerability. For example, according to the *neutral/sum measure* of Dutch bookability considered by Schervish et al. (2003), we calculate Joshua's Dutch-bookability by buying and/or selling bets to/from him in a way that maximizes the amount he is guaranteed to lose (normalizing the loss by the sum of the total stakes) ((Schervish et al., 2003, p.5), (Staffel, 2019, p. 64)). In the case at hand, we maximize the (normalized) amount that Joshua is guaranteed to lose by buying the gamble [$1 *if H*, $0 *otherwise*] from him for $0.2 (which he considers fair) and buying the gamble [$1 *if $\neg H$*, $0 *otherwise*] from him for $0.2 (which he also considers fair). The result is a normalized sure loss of $0.3 (=total loss/sum of stakes=0.6/2). If Joshua nudges his credences from 0.2 to $x \in (0.2, 0.5]$, he incurs a normalized sure loss of $(1 - 2x)/2$, which gets smaller as $x$ approaches 1/2. So his degree of Dutch bookability decreases as he nudges toward coherence.

To recap: if Joshua had the incoherent credences $b(H) = b(\neg H) = 0.2$, knew nothing about the bias of the coin, and had no other information suggesting one credence deserves more revision than another, then this seems as good a way to go as any. Given his lack of information, the only evaluative facts he can rely on to revise his credences are ones that are *guaranteed* to hold. Approximating coherence by nudging directly toward $c(H) = c(\neg H) = 0.5$ is guaranteed to yield more accuracy and less Dutch bookability, as we just saw (at least relative to one measure of accuracy and Dutch-bookability, respectively). And no other way of revising his credences is guaranteed to do better. So no alternative is obviously preferable.

But in our example Joshua is *not* in such an informationally impoverished situation. He *does* have more to go on. In particular, Joshua knows that the coin has a 0.2 bias toward heads. In that case, he need not rely only on evaluative facts that are *guaranteed* to hold. He might also rely on facts about what the objective chance function *expects* to hold. This is what leads Jim Joyce to respond to Joshua's case as follows:

> I do not praise Franco when I say that Hitler was worse along every dimension of dictatorial evil. I am not recommending Northern Manitoba's climate when I tell you that the weather in Churchill is better than the weather in Vostok in every season. Likewise, when I point out that Joshua's credences are accuracy-dominated by $\langle 0.5, 0.5 \rangle$ I do not imply that he should adopt the latter beliefs. In

fact, as we will see below, I should be quite certain that a person who knows what Joshua does about the chances should not adopt any of the credences that dominate his own. What he should do, instead, is to reflect more carefully on his total evidence with the goal of finding a credal state that strikes the optimal balance between the good of being confident in truths and the evil of doubting them, it being understood that this optimal state might well not be found among the dominating credences. When he does he will see that the optimal credences are $\langle 0.2, 0.8 \rangle$. (Joyce, 2013, p. 15)

The true chance function expects

$$ch(H) = 0.2, ch(\neg H) = 0.8 \tag{6}$$

to be maximally accurate (relative to the Brier score or any other strictly proper scoring rule). It also expects moving along the straight line from $b$ to $ch$ to strictly improve accuracy (on any strictly proper scoring rule). Even better, it expects moving along the straight line from $b$ to $ch$ to improve accuracy *more* than moving along the straight line from $b$ to $c$.
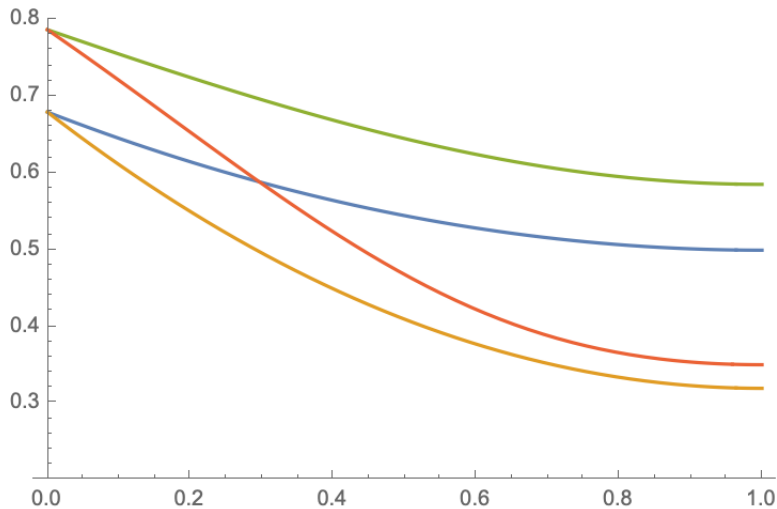


Figure 1: *Objective expected Brier inaccuracy of* $\lambda c + (1 - \lambda)b$ *(blue) and* $\lambda ch + (1 - \lambda)b$ *(orange) for* $\lambda \in [0, 1]$; *Objective expected Spherical inaccuracy of* $\lambda c + (1 - \lambda)b$ *(green) and* $\lambda ch + (1 - \lambda)b$ *(red).*

In light of this fact, approximating coherence by nudging "directly towards" $ch$ improves the *justification* of Joshua's credence more than moving directly towards $c$, on Joyce's view (Joyce, 2013, p. 17). So there is a clear epistemic rationale for setting aside Staffel's default advice about how to improve coherence.[4] The

---

[4]Staffel very openly does *not* provide general advice about how to improve coherence. Though "we would ultimately like to answer questions about how particular irrational thinkers should reason or change their credences," Staffel says, "taken by itself, the evaluative theory I have developed cannot (and is not intended to) answer these questions" (Staffel, 2019, p. 160). In general, how we should improve coherence will depend on a range of epistemic factors—*e.g.*, facts about objective expected accuracy, if you happen to have evidence about the chances—as well as practical factors, *e.g.*, whether spending mental energy improving coherence in some domain is worthwhile, or whether that mental energy might be better put to another purpose (Staffel, 2019, p. 161).

rationale is this: another path toward coherence is objectively expected to yield more epistemic value (accuracy *and* justification if Joyce is right) than the path toward the closest dominator. And Joshua knows what the true chance function is.

There is also a clear *practical* rationale for Joshua to improve coherence by nudging $b$ towards $ch$ rather than $c$. Consider an arbitrary bet on $H$ that Joshua might face.

|   | $H$ | $\neg H$ |
|---|---|---|
| $F$ | $a$ | $b$ |
| $G$ | $c$ | $d$ |

For concreteness, let $a = 3, b = 4, c = -1$ and $d = 6$. So Joshua faces the following decision problem, $\mathbb{D}$:

| $\mathbb{D}$ | $H$ | $\neg H$ |
|---|---|---|
| $F$ | 3 | 4 |
| $G$ | -1 | 6 |

Suppose that Joshua uses his credence $b(H) = 0.2$ to make his choice. Then he will choose $G$, since the expected utility of $G$ is higher than the expected utility of $F$ according to $b(H) = 0.2$:

$$b(H)(a) + (1 - b(H))(b) = (0.2)(3) + (1 - 0.2)(4) \tag{7}$$
$$= 3.8 \tag{8}$$
$$< 4.6 \tag{9}$$
$$= (0.2)(-1) + (1 - 0.2)(6) \tag{10}$$
$$= b(H)(c) + (1 - b(H))(d) \tag{11}$$

Would it be better for Joshua to use some other credence $x$ for $H$ to decide between $F$ and $G$? According to the true chance function, *no*. In decision problem $\mathbb{D}$, any $x < 1/3$ uniquely recommends $G$ and any $x > 1/3$ uniquely recommends $F$. (If $x = 1/3$, then both $F$ and $G$ are choiceworthy.) So chance expects Joshua to end up with the following utility if he uses any credence $x < 1/3$ to guide his choice:

$$ch(H)u(G\&H) + ch(\neg H)u(G\&\neg H) = (0.2)(-1) + (1 - 0.2)(6) = 4.6 \tag{12}$$

If $x > 1/3$ then the objective expected utility is

$$ch(H)u(F\&H) + ch(\neg H)u(F\&\neg H) = (0.2)(3) + (1 - 0.2)(4) = 3.8 \tag{13}$$

If Joshua flips a fair coin to decide when both $F$ and $G$ are choiceworthy, *i.e.*, when $x = 1/3$, then chance expects him to end up with utility 4.2. In any case, Joshua's credence of 0.2 for $H$ maximizes objective expected utility (as does any $x < 1/3$).

What about Joshua's credence for $\neg H$? Would it be better for him to use some other credence $y$ for $\neg H$ to decide between $F$ and $G$? Yes! In decision problem $\mathbb{D}$, any $y < 2/3$ recommends $F$ and its objective expected utility is 3.8. Any $y > 2/3$ recommends $G$ and its objective expected loss is 4.6. And if $y = 2/3$ and Joshua flips a fair coin to break the tie between $F$ and $G$, then

his objective expected utility is 4.2. The upshot: Joshua's credence of 0.2 for $\neg H$ *minimizes* rather than maximizes objective expected utility (as does any $y < 2/3$). He would be better off using a credence $y > 2/3$ for $\neg H$.

Now suppose that we calculate the objective expected *guidance value* in $\mathbb{D}$ of any pair of credences $\langle x, y \rangle$ by the sum of their individual objective expected utilities in $\mathbb{D}$. Then $ch = \langle 0.2, 0.8 \rangle$ maximizes objective expected guidance value in $\mathbb{D}$. Moreover, the true chance function expects moving along the straight line from $b$ to $ch$ to (weakly) increase guidance value in $\mathbb{D}$. Interestingly, it expects moving along the straight line from $b$ to $c$ to (weakly) *decrease* guidance value in $\mathbb{D}$, even though it improves Dutch bookability.
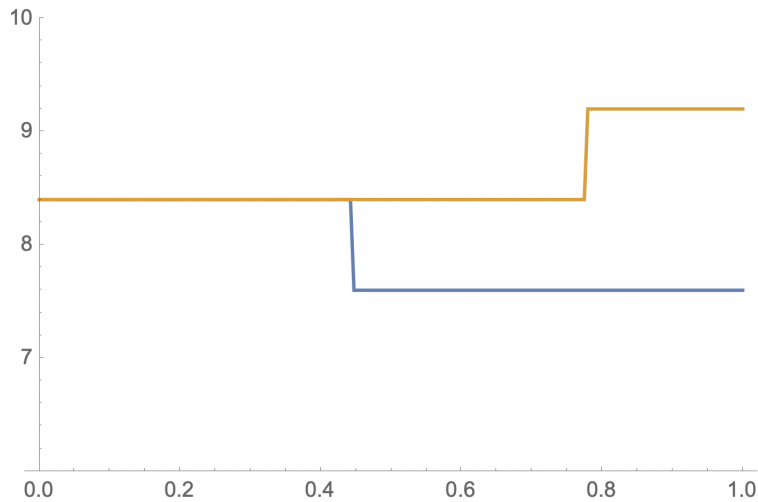


*Figure 2: Objective expected guidance value of $\lambda c + (1 - \lambda)b$ (blue) and $\lambda ch + (1 - \lambda)b$ (orange) in $\mathbb{D}$ for $\lambda \in [0, 1]$.*

What's going on here? Well, moving along the straight line from $b$ to $c$ raises Joshua's credence in $H$ from 0.2 to 0.5. From the perspective of the true chance function, this is a bad idea. If Joshua uses his current credence of 0.2 in $H$ to guide his choice in $\mathbb{D}$, he will choose $G$, which is what chance prefers. But once his credence crosses the $1/3$ threshold, he will will choose $F$. And chance thinks that choosing $F$ is likely to lead to greater loss.

On the other hand, chance is no fan of Joshua's 0.2 credence in $\neg H$. If Joshua uses his 0.2 credence in $\neg H$ to guide his choice in $\mathbb{D}$, he will choose $F$, which gets a *Booooo!* from chance. But moving along the straight line from $b$ to $c$ does not push his credence in $\neg H$ high enough. Joshua needs to cross the $2/3$ threshold to switch from $F$ to $G$. Moving along the straight line from $b$ to $c$ will never get him there. So, from chance's perspective, approximating coherence by moving from $b$ directly towards $c$ is all risk and no reward.

Moving along the straight line from $b$ to $ch$ is another story. Approximating coherence in *this* way will leave his current credence of 0.2 in $H$ intact. So if Joshua uses this credence to guide his choice in $\mathbb{D}$, he will continue to choose $G$— the right answer from chance's perspective. Even better, moving from $b$ directly towards $ch$ will eventually push his credence in $\neg H$ across the $2/3$ threshold. Once it crosses the threshold, then using his credence in $\neg H$ to guide his choice

will also lead him to choose $G$. So, from chance's perspective, approximating coherence by moving from $b$ directly towards $ch$ is all reward and no risk.

Of course, objective expected guidance value in $\mathbb{D}$ is not objective expected guidance value *tout court*. But that's no problem. Something similar is true for *any* binary decision on $H$ and $\neg H$. In any such problem, there is some threshold $q$ that divides the credences that recommend $F$ from the credences that recommend $G$. Any credence $x$ on one side of $q$ will recommend $F$ (or $G$, depending on the utilities in play). Any $x$ on the other side of $q$ will recommend the other option. (If $x = q$, then it recommends both.) And chance will always expect credences $x$ for $H$ on the same side of $q$ as $ch(H) = 0.2$ to maximize utility. Credences on the wrong side of $q$ minimize expected utility. Likewise, chance will always expect credences $y$ for $\neg H$ on the same side of $1 - q$ as $ch(\neg H) = 0.8$ to maximize utility. Credences on the wrong side minimize expected utility.

But the only credence $x$ for $H$ that is on the same side of $q$ as $ch(H) = 0.2$ *for all* $q \in [0, 1]$ is $x = 0.2$. So the only credence for $H$ that maximizes objective expected utility for *every possible bet on $H$ and $\neg H$* is the chance of $H$ itself, *viz.*, 0.2. Ditto for $\neg H$. The only credence for $\neg H$ that maximizes objective expected utility for *every possible bet on $H$ and $\neg H$* is the chance of $\neg H$ itself, *viz.*, 0.8. Hence $ch = \langle 0.2, 0.8 \rangle$ is the unique credal state that maximizes objective expected guidance value *in all binary decision problems on $H$ and $\neg H$*.

This gives us a grip on the best way for Joshua to approximate coherence from the practical perspective. The true chance function expects moving along the straight line from $b$ to $ch$ to (weakly) increase guidance value in all binary decision problems on H and $\neg H$. The reason is simple. Moving along the straight line from $b$ to $ch$ will never shift any of your credences from the right side of the decision-threshold to the wrong side, according to the true chance function (since both $ch(H) = 0.2$ and $ch(\neg H) = 0.8$ are on the right side of their respective thresholds from chance's perspective). So it cannot strictly decrease objective expected guidance value. But if $b(\neg H) = 0.2$ is on the wrong side of the decision threshold, then moving it toward $ch(\neg H) = 0.8$ will eventually pull it onto the right side. So it *can* strictly increase objective expected guidance value). (Indeed, this true whenever you approximate coherence by moving $b$ uniformly toward $ch$, whether you follow the straight line or not.) The moral: approximating coherence by moving from $b$ directly (or even just uniformly) towards $ch$ is all reward and no risk, from chance's perspective, regardless of which binary decision problem on $H$ and $\neg H$ you happen to face. But as we saw earlier, approximating coherence by moving $b$ directly towards its closest dominator $c$ sometimes decreases objective expected guidance value.

Where does this leave us? There is a clear epistemic and practical rationale for Joshua to improve coherence by nudging $b$ towards $ch$ rather than $c$. Even though nudging $b$ towards $c$ is *guaranteed* to improve accuracy and nudging $b$ towards $ch$ is not, the latter is nevertheless preferable to the former from the epistemic perspective. The reason: nudging $b$ towards $ch$ is objectively expected to yield more accuracy than nudging $b$ toward $c$. And Joshua knows the chances! Likewise, even though nudging $b$ towards $ch$ reduces one's degree of *guaranteed* loss (Dutch bookability) no more or less than a comparable nudge towards $c$, the latter is nevertheless preferable to the former from the practical perspective. The reason: nudging $b$ towards $ch$ is objectively expected to yield more guidance value than nudging $b$ toward $c$.

All of this reveals a slight oddity at the heart of Staffel's approach. Staffel

answers the gradational question—Why is it good to be *less*, rather than *more* incoherent?—as follows. There is some way of nudging incoherent credences toward coherence that both reduces the extent to which they are Dutch-bookable *and* is guaranteed to make them more accurate. This, according to Staffel, shows that there is a way of improving coherence that delivers increasing amounts of practical and epistemic value. But measures of Dutch bookability are not good measures of practical value. Credences are practically valuable in virtue of guiding action well. And guiding action well is a matter of recommending practically valuable actions—ones that produce as much utility as possible. Conspiracy theorists and superforecasters may both be probabilistically coherent. So they may both be perfectly immune to Dutch books. But the credences of the latter are nevertheless better at guiding action than the former (from the perspective of both *truth* and *chance*). And if Dutch bookability can remain fixed while guidance value varies, then Dutch bookability is not a good measure of practical value. Similarly, in our running example,

$$d(H) = 0.2, d(\neg H) = 0.794 \tag{14}$$

is Dutch bookable and

$$c(H) = 0.5, c(\neg H) = 0.5 \tag{15}$$

is not. But chance expects $d$'s recommended action to produce at least as much utility as $c$'s in every binary decision problem, and sometimes strictly more (since $d$'s credences are uniformly closer to the true chances than $c$'s). So chance expects $d$ to be better than $c$ at guiding action. But chance also expects (indeed is *certain*) that $c$ is less Dutch-bookable than $d$. If chance's view of guidance value and Dutch bookability can come apart like this, then so much the worse for Dutch bookability as a measure of practical value.

Of course, if your credences are Dutch bookable, then they must be incoherent. And in that case, they are *guaranteed* to be worse at guiding action than some other credences (see section 3). So Dutch bookability is a surefire *sign* that one's credences are needlessly bad at guiding action. But your degree of immunity to Dutch books is not *itself* a good measure of how practically valuable your credences are.

The lesson, I think, is that Staffel's approach needs a small tweak. Following Schervish (1989), Levinstein (2017) and Pettigrew (2020), I propose using strictly proper scoring rules to measure *both* accuracy *and* guidance value. If the principal epistemic role of credences is to encode accurate truth-value estimates, then strictly proper scoring rules are fit to serve as measures of epistemic value. Likewise, if the principal practical role of credences is to recommend actions that produce as much utility as possible, then strictly proper scoring rules are fit to serve as measures of practical value. The upshot: we ought to replace Dutch bookability measures in Staffel's framework with strictly proper scoring rules. I will argue that this allows us to give a more satisfying version of Staffel's answer to the gradational question.

## 3   The Solution

One last time, consider an arbitrary binary decision problem $D$ on $X$ and $\neg X$ that Joshua might face.

|   | $X$ | $\neg X$ |
|---|-----|----------|
| $F$ | $a$ | $b$ |
| $G$ | $c$ | $d$ |

For any credence $x$ in $X$, there is some act that $x$ recommends in $D$, *viz.*, the act that maximises expected utility. If

$$(x)(a) + (1 - x)(b) > (x)(c) + (1 - x)(d) \qquad (16)$$

then $x$ uniquely recommends $F$ in $D$. In that case, we call $F$ the *Bayes act* (relative to $x$ in $D$). If

$$(x)(a) + (1 - x)(b) < (x)(c) + (1 - x)(d) \qquad (17)$$

then $x$ uniquely recommends $G$ ($G$ is the Bayes act). If

$$(x)(a) + (1 - x)(b) = (x)(c) + (1 - x)(d) \qquad (18)$$

then $x$ recommends both $F$, $G$, and any coin flip between the two (any "mixed act"). In that case, we can simply *pick* one of $F$, $G$ or a mixture as "the" Bayes act, since all are equally choice*worthy*. How we pick won't matter.[5]

Using the Bayes act, we can define the *loss* of credence $x$ in decision problem $D$ and state $X$, which we denote $\mathcal{L}_D(x, X)$. $\mathcal{L}_D(x, X)$ is the difference between the maximum possible utility achievable in state $X$ and the utility that $x$'s recommended action produces in $X$. For example, if $a > c$ and $x$ recommends $G$ in $D$, then $\mathcal{L}_D(x, X) = a - c$, since $a$ is the maximum possible utility achievable in state $X$ and $c$ is the utility that $G$ produces in $X$. $\mathcal{L}_D(x, X)$ represents how far $x$'s recommended action falls short of the objectively best action to perform in state $X$. (For formal details, see the appendix.) Similarly, $\mathcal{L}_D(x, \neg X)$ is the difference between the maximum possible utility achievable in state $\neg X$ and the utility that $x$'s recommended action produces in $\neg X$.

Finally, we can use these losses to measure the guidance value of credence $x$ if $X$ is true or false, respectively. Firstly, choose a measure $\mu$ on the space of all possible binary decision problems. The measure $\mu$ might reflect: (i) how likely it is that your next decision problem will fall in one class or another from the perspective of the true chance function (objective doxastic interpretation); (ii) how likely this is from your own perspective (subjective doxastic interpretation); (iii) how much you care about choosing well in any given class of choice problems (subjective bouletic interpretation). We will return to this question in section 4. In any case, given a measure $\mu$, we can measure the guidance value of credence $x$ in state $X$ by integrating the losses $\mathcal{L}_D(x, X)$ relative to $\mu$:

$$g_1(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) \, \mathrm{d}\mu$$

$g_1(x)$ represents how far $x$'s recommended action falls short of the objectively best action (maximum possible utility) in state $X$ *on average* (according to $\mu$).

---

[5]It won't matter for two reasons. Firstly, the class of decision problems with non-unique Bayes acts will be a set of measure zero relative to any measure that is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^4$ (*i.e.*, the space of possible decision problems). Secondly, relative to *any* measure on $\mathbb{R}^4$ and *any* choice of Bayes act in this case, the measure of guidance value that we construct will turn out to be strictly proper.

It is worth emphasising that $g_1$ *penalises* $x$ for recommending actions that are suboptimal in state $X$. The lower the penalty, the better $x$ does (at guiding action, on average, in state $X$).

Similarly, we can measure the guidance value of $x$ in state $\neg X$ by integrating the losses $\mathcal{L}_D(x, \neg X)$ relative to $\mu$:

$$g_0(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) \, d\mu$$

$g_0(x)$ represents how far $x$'s recommended action falls short of the objectively best action in state $\neg X$ on average (according to $\mu$).

What's more, we can prove that any pair of guidance value measures $\langle g_0, g_1 \rangle$ that are constructed in this way are *strictly proper* (given fairly weak assumptions about $\mu$; see appendix):

**Proposition 1.** For any $x, p \in [0, 1]$ with $x \neq p$

$$pg_1(x) + (1-p)g_0(x) > pg_1(p) + (1-p)g_0(p)$$

We can then calculate the guidance value of an entire credence function $c : \mathcal{F} \to \mathbb{R}$ (where $\mathcal{F}$ is a $\sigma$-algebra on a finite sample space $\Omega$) at a world $w \in \Omega$ by taking a weighted average of the guidance value of $c(X)$ for each $X \in \mathcal{F}$:

$$\mathcal{G}(c, w) = \sum_{X \in \mathcal{F} \, : \, w \in X} \lambda_X g_1(c(X)) + \sum_{Y \in \mathcal{F} \, : \, w \notin Y} \lambda_Y g_0(c(Y))$$

As (Pettigrew, 2020, pp. 76-77) notes, we can use such guidance value measures to provide a new pragmatic argument for probabilism. For any agent with incoherent credences $b$ and any reasonable measure of guidance (dis)value $\mathcal{G}$ (which is strictly proper according to proposition 1), there is some coherent set of credences $c$ that is *guaranteed* to guide action strictly better than $b$, *i.e.*,

$$\mathcal{G}(c, w) < \mathcal{G}(b, w)$$

for all $w \in \Omega$ Predd et al. (2009). Coherent credences are never guidance-value-dominated in this way. This shows that incoherent credences are needlessly bad at guiding action.

In addition, we can use these guidance value measures to provide a more satisfying version of Staffel's answer to the gradational question. Why is it good to be less, rather than more incoherent? The answer: the cardinal *epistemic* good-making feature of credences is *accuracy*. And strictly proper scoring rules are the right tools to measure accuracy. Similarly, the cardinal *practical* good-making feature of credences is *guiding action well*, or put more perspicuously, *recommending actions that produce as much utility as possible*. Strictly proper scoring rules are the right tools for measuring this as well. Moreover, when one's epistemic scoring rule and practical scoring rule are "sufficiently similar," then there is some way of nudging incoherent credences toward coherence that is guaranteed to yield more of both types of value, relative to those scoring rules. More carefully:

1. **Veritistic Epistemic Value**: Credences have *epistemic* value in virtue of being *accurate* (*i.e.*, encoding accurate truth-value estimates). So the

epistemic value of a credence function $c : \mathcal{F} \to \mathbb{R}$ at a world $w$ is given by $-\mathcal{I}(c, w)$, where $\mathcal{I}$ is some reasonable measure of inaccuracy. We assume that $\mathcal{I}$ is additive, $i.e.$,

$$\mathcal{I}(c, w) = \sum_{X \in \mathcal{F} \,:\, w \in X} \lambda_X f_1(c(X)) + \sum_{Y \in \mathcal{F} \,:\, w \notin Y} \lambda_Y f_0(c(Y))$$

and that $f_1$ and $f_0$ are continuous and strictly proper.

2. **Causalist Practical Value**: Credences have *practical* value in virtue of *recommending practically valuable actions*. And actions have practical value in virtue of *producing utility*. So the practical value of a credence function $c : \mathcal{F} \to \mathbb{R}$ at a world $w$ is given by $-\mathcal{G}(c, w)$, where $\mathcal{G}$ is some reasonable measure of *guidance value*. $\mathcal{G}$ is a reasonable measure of guidance value just in case $-\mathcal{G}(c, w)$ captures the extent to which $c$ recommends actions that produce maximal utility in $w$ (on average, in binary decision problems). We assume that $\mathcal{G}$ is additive, $i.e.$,

$$\mathcal{G}(c, w) = \sum_{X \in \mathcal{F} \,:\, w \in X} \lambda_X g_1(c(X)) + \sum_{Y \in \mathcal{F} \,:\, w \notin Y} \lambda_Y g_0(c(Y))$$

and that $g_1(x)$ and $g_0(x)$ take the following form:

$$g_1(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) \, \mathrm{d}\mu$$

and

$$g_0(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) \, \mathrm{d}\mu$$

3. **Proposition 1 (appendix).** Any reasonable measure of guidance value $\mathcal{G}$ is strictly proper.

4. **Similarity Postulate**: If $\mathcal{I}$ and $\mathcal{G}$ are reasonable measures of the epistemic and practical value of a single agent's credences, then they must be "sufficiently similar" to one another in the following sense: If $x + y > 1$, then

$$\frac{xy}{(1-x)(1-y)} > \frac{m(x)/n(x)}{m(y)/n(y)}, \frac{m(y)/n(y)}{m(x)/n(x)}$$

and if $x + y < 1$, then

$$\frac{xy}{(1-x)(1-y)} < \frac{m(x)/n(x)}{m(y)/n(y)}, \frac{m(y)/n(y)}{m(x)/n(x)}$$

where $m$ and $n$ are the Schervish density functions of $\mathcal{I}$ and $\mathcal{G}$ (see below).

5. **Nudge Toward Coherence.** If $\mathcal{I}$ and $\mathcal{G}$ are "sufficiently similar," in the sense of the Similarity Postulate, then for any incoherent credence function $b : \mathcal{F} \to \mathbb{R}$, there is some coherent credence function $c : \mathcal{F} \to \mathbb{R}$ such that nudging $b$ in the direction of $c$ is guaranteed to yield more epistemic and practical value, respectively, $i.e.$,

$$\nabla_c \mathcal{I}(b, w) < 0$$

and

$$\nabla_c \mathcal{G}(b, w) < 0$$

for all $w \in \Omega$. (Propositions 3 and 4 in the appendix prove this for the special case of $\mathcal{F} = \{X, \neg X\}$.)

12

6. **Dominance.** If credences $c$ are guaranteed to be more epistemically and practically valuable than credences $b$, then any agent ought to strictly prefer $c$ to $b$.

C. Any agent with incoherent credences $b$ ought to strictly prefer the result of nudging $b$ toward some coherent $c$ over $b$ itself, *i.e.,* she ought to strictly prefer $(1 - \epsilon)b + \epsilon c$ to $b$ (for some $\epsilon > 0$).

I will not defend premises 1 and 2 further here. Premise 3 is a slight variant of a result in (Pettigrew, 2020, 6.3.2). Proofs of premises 3 and 5 are in the appendix. Premise 6 is uncontroversial.[6] That just leaves premise 4 as the odd duckling.

To get a sense of what premise 4 amounts to, consider three popular strictly proper rules:

- **Brier Score**: $\mathcal{I}(c, w) = \sum_{X \in \mathcal{F}} (w(X) - c(X))^2$

- **Log Score**: $\mathcal{I}(c, w) = \sum_{X \in \mathcal{F}} -log(|1 - w(X) - c(X)|)$

- **Spherical Score**: $\mathcal{I}(c, w) = \sum_{X \in \mathcal{F}} \left( 1 - \frac{|1 - w(X) - c(X)|}{\sqrt{c(X)^2 + (1 - c(X))^2}} \right)$

Building on the work of (Savage, 1971, pp. 786-7), (Schervish, 1989, Thm 4.2) shows that a pair of scoring rules $f_1$ and $f_0$ are continuous and strictly proper if and only if there is some non-negative density function $m : [0, 1] \to \mathbb{R}_{\geqslant 0}$ such that

$$f_1(x) = \int_x^1 (1 - t)m(t) \, \mathrm{d}t$$

and

$$f_0(x) = \int_0^x tm(t) \, \mathrm{d}t$$

and $\int_a^b m(t) \, \mathrm{d}t > 0$ for any $0 \leqslant a < b \leqslant 1$.[7] This density function $m$ is what I called the *Schervish density function* in premise 4. These density functions are closely related to the measure $\mu$ that we used to construct our guidance value measures.[8] $\mu$ reflects either (i) how likely it is that your next decision problem will fall in one class or another (objectively or subjectively), or (ii) how much you care about choosing well in any given class of choice problems. The Schervish density function can be thought of as condensing the information in $\mu$. It lumps each (binary) decision problem together with all of the other decision problems that share the same decision threshold $q$ (section 2), which are sometimes called $q$-problems. It reflects either (i) how likely it is that your next decision problem

---

[6](Pettigrew, 2016, Ch. 2) argues that we ought to weaken Dominance to a principle he calls "Immodest Dominance." But Pettigrew casts Dominance as a principle governing *rationality* rather than *strict preference*. None of Pettigrew's concerns impugn the Dominance principle governing strict preference.

[7](Schervish, 1989, Thm 4.2) actually only requires that $f_1$ and $f_0$ are left continuous and that their values at 0 and 1 are the limits of their values as you approach 0 and 1 from the right and left, respectively.

[8]The density $m$ is the Radon–Nikodym derivative (with respect to the Lebesgue measure) of a certain transformation of $\mu$.

will be a $q$-problem (objectively or subjectively), or (ii) how much you care about choosing well in a $q$-problem.

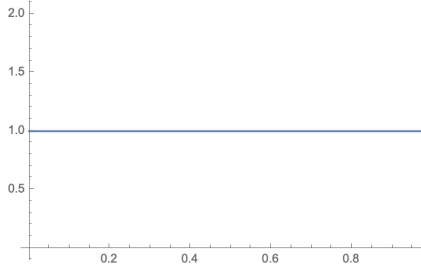In any case, here are the Schervish density functions for the Brier, Log and Spherical scores, respectively.
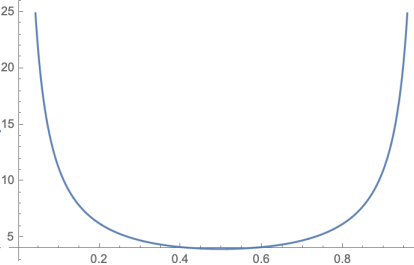
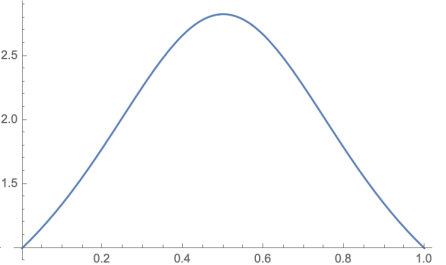*Fig 3: Brier score density.*     *Fig 4: Log score density.*     *Fig 5: Spherical score density.*

To help illustrate what premise 4 means, observe: The Brier score is "sufficiently similar" to both the Log and Spherical scores. That is, the Brier density and Log density satisfy the Similarity Postulate. The Brier density and the Spherical density also satisfy the similarity postulate. But the Log and Spherical scores are just too different to satisfy the Similarity Postulate. To see this, let $m$ be the Log density:

$$m(t) = \frac{1}{t(1-t)}$$

Let $n$ be the Spherical density:

$$n(t) = \frac{\frac{(3-2t)t-1}{(1+2(t-1)t)^{3/2}} + \frac{1}{(1+2(t-1)t)^{1/2}}}{t}$$

And consider the incoherent credences $x = 0.9$ for $X$ and $y = 0.2$ for $\neg X$. Then we have

$$\frac{xy}{(1-x)(1-y)} = 2.25, \ \frac{m(x)/n(x)}{m(y)/n(y)} \approx 2.35, \ \frac{m(y)/n(y)}{m(x)/n(x)} \approx 0.425$$

So the Log score and Spherical score jointly violate the Similarity Postulate (since $2.25 \not> 2.35$).

Here's what this means. If your epistemic scoring rule is the Spherical score and your practical scoring rule is the Brier score, then they are "sufficiently similar" for there to be single way of nudging $\langle 0.9, 0.2 \rangle$ toward coherence that is guaranteed to yield more of both types of value. But if your epistemic scoring rule is the Spherical score and your practical scoring rule is the Log score, then they violate the Similarity Postulate and our main argument guarantees no such thing. Indeed, in this case *any* way of nudging $\langle 0.9, 0.2 \rangle$ toward coherence that is guaranteed to increase epistemic value *fails* to deliver a guaranteed increase in practical value (and vice versa).

Hopefully this gives a little insight into our mysterious little duckling. Let's take stock. I argued that Staffel's approach to the gradational question is promising, but needs a small tweak. We *can* show that it's better to be more rather than less coherent. To do so, we need to identify reasonable measures of epistemic and practical value, respectively. Then we need to show that we can always nudge incoherent credences toward coherence in a way that is

guaranteed to yield more of both types of value. Staffel's problem, I argued, was that while she identified appropriate measures of epistemic value, she did not identify appropriate measures of practical value. Staffel uses measures of Dutch-bookability as her preferred measures of practical value. But credences have practical value in virtue of recommending actions that produce as much utility as possible. And while susceptibility to a Dutch book is a surefire *sign* that one's credences are needlessly bad at this task, one's degree of Dutch-bookability is not *itself* a good measure of how well they recommend practically valuable actions. Strictly proper scoring rules are the right tools, I argued, for measuring both epistemic and practical value. Luckily, we can rerun Staffel's strategy swapping in strictly proper scoring rules for Dutch-bookability measures and end up with a very similar conclusion.

How satisfying is this new version of Staffel's answer to the gradational question? That depends primarily on what sort of case can be made for the Similarity Postulate. Perhaps disappointingly, I will not attempt anything like a proper defence of it here (though I will say a *bit* more in section 4). It is worth noting, however, that Staffel guarantees that there is sufficient similarity between measures of epistemic and practical value by requiring the former to be *convex*. While Joyce (2009) offers some considerations in favour of Convexity, it is nevertheless controversial. Defending Convexity is not obviously easier or harder than defending the Similarity Postulate. So I think we have successfully defused the basic problem for Staffel's approach without incurring much dialectical cost. And ditching Convexity in favour of the Similarity Postulate comes with some advantages. For example, our main argument establishes that it's better to be more rather than less coherent even when your epistemic scoring rule is the Spherical score. But the Spherical score is not convex. So Staffel's original argument does not establish this.

Before wrapping up, let's return one more time to Joshua. Joshua's case led us to the realisation that we ought to replace Dutch bookability measures in Staffel's framework with measures of guidance value (which turn out to take a familiar form: they are strictly proper scoring rules). And that led to a more satisfying version of Staffel's answer to the gradational question. But Joshua is no mere stepping stone! Let's use what we have learned about measures of guidance value to more carefully diagnose why Joshua should nudge toward coherence in one way rather than another.

Recall, Joshua has the incoherent credences

$$b(H) = 0.2, b(\neg H) = 0.2$$

The closest coherent credences to Joshua's (when measuring accuracy by the Brier score and closeness by squared Euclidean distance) are:

$$c(H) = 0.5, c(\neg H) = 0.5$$

And the chance of heads and tails—which Joshua knows—are:

$$ch(H) = 0.2, ch(\neg H) = 0.8$$

We said that there is a clear epistemic rationale for Joshua to nudge towards *ch* rather than *c*. Even though nudging *b* towards *c* is *guaranteed* to improve accuracy and nudging *b* towards *ch* is not, nudging towards *ch* is *objectively*

*expected* to yield more accuracy than nudging towards $c$. Using the Brier score, $\mathcal{B}$, to measure inaccuracy, we can make the point even more explicitly. Nudge $b$ towards $c$. Let's go half way between the two, just for concreteness. Then we end up with:

$$p(H) = 0.35, p(\neg H) = 0.35$$

If we slide $b$ half way towards $ch$ we end up with:

$$q(H) = 0.2, q(\neg H) = 0.5$$

Observe: $p$ is guaranteed to have a lower Brier score (be less inaccurate) than $b$; $q$ is not.

|   | $\mathcal{B}(\cdot, H)$ | $\mathcal{B}(\cdot, \neg H)$ |
|---|---|---|
| $b$ | 0.68 | 0.68 |
| $p$ | 0.5 | 0.5 |
| $q$ | 1.28 | 0.08 |

Nevertheless, chance *expects* $q$ to have a lower Brier score than $p$:

$$E_{ch}(\mathcal{B}(q)) = 0.41, E_{ch}(\mathcal{B}(p)) = 0.545$$

And Joshua knows the chances! This gives him good epistemic reason to prefer $q$ (the result of nudging $b$ towards $ch$) over $p$ (the result of nudging $b$ towards $c$). Of course, Joshua may not know that the Brier score is the "true" measure of accuracy, as we have assumed. But

$$E_{ch}(\mathcal{I}(q)) < E_{ch}(\mathcal{I}(p))$$

for any strictly proper scoring rule $\mathcal{I}$ (because $q$'s probabilities are uniformly closer to $ch$'s than $p$'s). So Joshua knows that $q$ has higher objective expected accuracy than $p$ simply because this is true for *all* reasonable inaccuracy measures.

We also said that there is a clear practical rationale for Joshua to nudge $b$ towards $ch$ rather than $c$. Even though nudging $b$ towards $ch$ reduces one's degree of *guaranteed* loss (Dutch bookability) no more or less than a comparable nudge towards $c$, nudging towards $ch$ is *objectively expected* to yield more guidance value than nudging towards $c$. Using a particular measure of Dutch book vulnerability, *e.g.*, the neutral/sum measure, and a particular scoring rule to measure guidance value, *e.g.*, the log score, $\mathcal{L}$, we can put the point more explicitly. Nudging $b$ towards $c$ and $ch$, respectively (resulting in $p$ and $q$), improves Joshua's Dutch book vulnerability to exactly the same degree, according to the neutral/sum measure:

|   | Neutral/sum measure |
|---|---|
| $b$ | 0.3 |
| $p$ | 0.15 |
| $q$ | 0.15 |

Nevertheless, chance *expects* $q$ to have a lower Log score (higher guidance value) than $p$:

$$E_{ch}(\mathcal{L}(p)) = 1.48, E_{ch}(\mathcal{L}(q)) = 1.19$$

Since Joshua knows the chance of $H$ and $\neg H$, this gives him good reason to prefer making decisions with $q$ rather than $p$. Of course, Joshua may not know that the Log score is the "true" measure of guidance value, as we have assumed. But

$$E_{ch}(\mathcal{G}(q)) < E_{ch}(\mathcal{G}(p))$$

for any strictly proper scoring rule $\mathcal{G}$. So Joshua knows that $q$ has higher objective expected guidance value than $p$ simply because this is true for *all* reasonable guidance value measures.

You might still wonder: what does it really mean to say that $q$ has higher objective expected guidance value than $p$? What sort of reason does this give Joshua to prefer making decisions with $q$ rather than $p$? Again, the reason roughly is that chance expects $q$ to recommend actions that produce more utility on average than $q$ does. But what "on average" amounts to depends on how we interpret $\mathcal{G}$. Recall

$$\mathcal{G}(c, w) = \sum_{X \in \mathcal{F} \,:\, w \in X} g_1(c(X)) + \sum_{Y \in \mathcal{F} \,:\, w \notin Y} g_0(c(Y))$$

where

$$g_1(x) = \int_x^1 (1 - q) m(q) \, \mathrm{d}q$$

$$g_0(x) = \int_0^x q m(q) \, \mathrm{d}q$$

and $m$ is a Schervish density function. This density function might reflect the *chance* that next binary decision problem Joshua will face on $H$ and $\neg H$ will be a $q$-problem (*i.e.*, a decision problem where all the credences on one side of $q$ recommend one option and all the credences on the other side of $q$ recommend the other option). In that case, $E_{ch}(\mathcal{G}(p))$ is just chance's best unconditional estimate of how far $p$'s recommended action will fall short of optimal in that next problem. Ditto for $E_{ch}(\mathcal{G}(q))$. Knowing these best estimates seems like excellent reason to prefer making decisions with $q$ rather than $p$.

Alternatively, the density function $m$ might reflect Joshua's *degree of belief* that next binary decision problem he will face on $H$ and $\neg H$ will be a $q$-problem. In that case, $E_{ch}(\mathcal{G}(p))$ is what Joshua should adopt, given his knowledge of the chance of $H$ and $\neg H$, as *his own* best estimate of how far $p$'s recommended action will fall short of optimal in the next decision problem (according to $\mathcal{G}$).[9] Ditto for $E_{ch}(\mathcal{G}(q))$. This too gives him excellent reason to prefer making decisions with $q$ rather than $p$.

Finally, the density function $m$ might reflect how much Joshua *cares* about choosing well in different types of decision problems. On this interpretation (roughly speaking), the larger $m(q)$ is, the more Joshua cares about choosing well in $q$-problems. In that case, $g_1(x)$ represents how far $x$'s recommended action will fall short of optimal on average if $X$ is true. But the decision problems that Joshua cares more about figure more heavily into this average. The ones he cares less about figure in less heavily. Similarly for $g_0(x)$. $E_{ch}(\mathcal{G}(p))$, then,

---

[9]More carefully, it is the estimate that he would adopt if (i) he adopted the known chances of $H$ and $\neg H$ as his own credences and (ii) used his expectation of $p$'s (and $q$'s) guidance value as his best estimate(s), which is what coherence requires.

can be glossed roughly as chance's best estimate of how far $p$'s recommended action will fall short of optimal in the decision problems that Joshua cares most about. (Better but more opaque: $E_{ch}(\mathcal{G}(p))$ is chance's best estimate of how far $p$'s recommended action will fall short of optimal on priority-weighted-average.)

Does this give Joshua good reason to prefer making decisions with $q$ rather than $p$? That's not so clear. If, for example, Joshua was sure that he was going to face the decision problems that he cares *least* about, then chance's best estimate of how good $p$'s (and $q$'s) recommended actions will be in the problems he cares *most* about would be neither here nor there. Preferring to make decisions with $q$ rather than $p$ on the basis of such estimates would be no better than wishful thinking (*i.e.*, wishfully thinking that he will face one type of decision problem when he will likely face another). Even if Joshua has *no idea* which decision problem he will face, it seems like wishful thinking to prefer $q$ to $p$ on this basis. Joshua doxastic state should not *preclude* the possibility that his most preferred decision problems are highly unlikely. But in that case, the mere fact that chance expects $q$ to do better than $p$ in those problems is not decisive reason to prefer $q$ to $p$.[10]

This is symptomatic of a deeper problem with the subjective bouletic interpretation. Credences have practical value *at a world* in virtue of recommending actions which are practically valuable *at that world*. And actions have practical value *at a world* in virtue of producing utility *at that world*. Now, as we have set things up, "worlds" are just elements $w$ of a finite sample space $\Omega$. They determine the truth-value of any proposition $X$ in $\mathcal{F}$ (which is just a $\sigma$-algebra on $\Omega$): $X$ is true if $w \in X$ and false otherwise. But they do not determine which binary decision problems on $X$ and $\neg X$ any agent will face at that "world". So we cannot, strictly speaking, say how practically valuable actions—and hence credences—are at such "worlds". But we can come close if we independently specify how likely (either objectively or subjectively) she is to face different classes of decision problems. In that case, we can specify the *expected* practical value of *credences* at such "worlds" and treat this expectation or estimate as a proxy for their final practical value. This is precisely how measures of guidance value function on the doxastic interpretation (both objective and subjective). But on the bouletic interpretation, $g_1(x)$ and $g_0(x)$ are not *estimates* at all; certainly not estimates of how close to optimal $x$'s recommended actions will be (in $X$ and $\neg X$, respectively). Rather, they reflect how close to optimal those actions will be in the decision problems that the agent cares most about. But in general this tells us *nothing* about how much utility those actions will *actually* produce (or even produce in expectation). So are these "priority weighted averages" fit to serve as proxies for final practical value? I see no reason to think so. That is, I see no reason to think so *unless* we treat an agent's level of *concern* for $q$-problems (represented by $m(q)$) as her *degree of belief* that she will face a $q$-problem. But that is just wishful thinking!

So much, then, for the subjective bouletic interpretation. The long and short

---

[10]If Joshua applied the Principle of Indifference, then you might be tempted to interpret $g_1(x) = \int_x^1 (1-q) \cdot m(q) \cdot 1 \, dq$ as follows: Joshua's degree of concern for $q$-problems, $m(q)$, scales the loss that $x$ incurs in $q$-problems if $X$ is true, $(1-q)$. And $n(q) = 1$ reflects his uniform subjective probability density over $q$-problems. Likewise for $g_0$. In that case, $E_{ch}(\mathcal{G}(p))$ reflects Joshua's own best estimate of $p$'s scaled losses in the next decision problem. But the POI is no requirement of rationality. So this way forward grounds no *general* reason for Joshua to prefer $q$ to $p$.

of it is this: there is indeed a clear practical rationale for Joshua to nudge $b$ towards $ch$ rather than $c$. The rationale is that either (i) Joshua himself should expect (on the subjective doxastic interpretation) or (ii) Joshua knows that chance expects (on the objective doxastic interpretation) that $q$ will recommend better actions than $p$; ones that will produce more utility in the next decision problem that Joshua faces.

# 4    Open Questions

Let's wrap up by considering a few open and pressing concerns about our new version of Staffel's answer to the gradational question.

1. **The Similarity Postulate.** A proper defence of our answer requires a proper defence of the Similarity Postulate. Recall that the Similarity Postulate says: If $\mathcal{I}$ and $\mathcal{G}$ are reasonable measures of the epistemic and practical value of a single agent's credences, then they must be "sufficiently similar" to one another. More carefully, their Schervish densities must be "sufficiently similar" to one another. But why in the world would we expect this to be true? Perhaps if those densities each reflected some type of *preference* that the agent has, we could work up an argument that they should not be too dissimilar (maybe they are preferences over similar types of options, or something of the sort). But we have already pooh-poohed the subjective bouletic interpretation! So it seems that we will have to hope for some tight(ish) correspondence between the chance-facts or belief-facts that determine the density of $\mathcal{G}$ and whatever it is that determines the density of $\mathcal{I}$. Perhaps this can be done, but it is not obvious how.

2. **Binarity.** In our main argument, we assume that reasonable measures of guidance value $\mathcal{G}$ are additive, *i.e.*,

$$\mathcal{G}(c, w) = \sum_{X \in \mathcal{F} \,:\, w \in X} \lambda_X g_1(c(X)) + \sum_{Y \in \mathcal{F} \,:\, w \notin Y} \lambda_Y g_0(c(Y))$$

and that $g_1(x)$ and $g_0(x)$ take the following form:

$$g_1(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) \, \mathrm{d}\mu$$

and

$$g_0(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) \, \mathrm{d}\mu$$

$\mathcal{D}$ is the space of binary decision problems on $X$ and $\neg X$, *i.e.*, decision problems of the form:

|   | $X$ | $\neg X$ |
|---|-----|----------|
| $F$ | $a$ | $b$ |
| $G$ | $c$ | $d$ |

The result is that $\mathcal{G}(c, w)$ only captures the extent to which $c$'s recommended actions *in binary decision problems* (actions like $F$ or $G$) produce optimal results in $w$ (in expectation). But not all decision problems are binary! Sometimes our menu has more than two options. And sometimes those

options have more than two possible outcomes. Why ignore all of these other decision problems when calculating the guidance value of $c$ at $w$? The reason is this: $c$ only encodes truth-value estimates and truth-value estimates are informationally limited. When an agent's credences $c$ are probabilities, then coherence requires her to estimate any other variable $V$ (*e.g.*, how much utility action $A$ will produce in non-binary decision problem $D$) with $E_c(V)$—the expected value of $V$ relative to $c$. (If $c$ is probabilistic, then $E_c$ is the unique coherent extension of $c$ to the space of all variables $V : \Omega \to \mathbb{R}$.) But when $c$ is non-probabilistic, there is no way to extract estimates of non-binary variables from $c$. ($c$ is already incoherent. So we cannot extract an estimate of $V$ from $c$ by looking at the unique coherent extension of $c$. There is no such thing.) As a result, $c$ simply fails to provide any guidance value in non-binary decision problems. The upshot: if we want to assess both coherent and incoherent credences for guidance value using a common standard, then we must focus myopically on binary decision problems.

But there is another way forward. Rather than assessing *credences* or *truth-value estimates* and focusing myopically on binary decision problems, we could directly specify an agent's best estimates of *all variables* and score these for accuracy and guidance value, respectively. If we go this route, then we do not need to extract an estimate of $V$ from $c$. We take it is an input, just like credences/estimates of truth-values. This gives us recommended actions in *all* decision problems: $A$ is choiceworthy iff $A$ is amongst the options with maximal estimated utility. Schervish et al. (2014a,b) explore scoring rules for such infinite sets of estimates/forecasts.

3. **Single-Shot Guidance Value**. Scoring rules

$$g_1(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) \, \mathrm{d}\mu$$

and

$$g_0(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) \, \mathrm{d}\mu$$

are *expectations* of the loss of $x$ in state $X$ and $\neg X$, respectively, only if the decision problems in $\mathcal{D}$ are mutually exclusive and jointly exhaustive. And they *must* be expectations (either objective or subjective) for $\mathcal{G}$ to be a statistic worth caring about. Hence we must think of the decision problems in $\mathcal{D}$ as different possible descriptions of a *single* decision problem that our agent will face (*cf.* (Schervish, 1989, p. 1859)). But in that case, $\mu$—and hence $g_0$, $g_1$ and $\mathcal{G}$—will change from one decision context to the next. Consequently, whether the Similarity Postulate is satisfied will change from one decision context to the next. You might worry that this renders our version of Staffel's answer to the gradational question rather brittle. Whether or not this is so will depend on how concern #1 shakes out.

4. **Nudge Toward Coherence**. In the appendix, I prove that if one's epistemic score rule, $\mathcal{I}$, and practical scoring rule, $\mathcal{G}$, are "sufficiently similar," in the sense of the Similarity Postulate, then for any incoherent credence function $b : \mathcal{F} \to \mathbb{R}$, there is some coherent credence function $c : \mathcal{F} \to \mathbb{R}$ such that nudging $b$ in the direction of $c$ is guaranteed to yield more epistemic and practical value. But I only prove this for the

special case of $\mathcal{F} = \{X, \neg X\}$, *i.e.*, when you just have a credence for one proposition and its negation, nothing more. A proper defence of our answer to the gradational question requires generalising propositions 2-4 in the appendix.

# 5 Conclusion

Why is it good to be *less*, rather than *more* incoherent? Julia Staffel answers this question by showing that if your credences are incoherent, then there is some way of nudging them toward coherence that is guaranteed to make them more accurate *and* reduce the extent to which they are Dutch-bookable. This seems to show that such a nudge toward coherence makes them better fit to play their key epistemic and practical roles: representing the world and guiding action. I argued that Staffel's strategy needs a small tweak. While she identifies appropriate measures of epistemic value, she does not identify appropriate measures of practical value. Staffel measures practical value using Dutch-bookability scores. But credences have practical value in virtue of recommending actions that produce as much utility as possible. And while susceptibility to a Dutch book is a surefire *sign* that one's credences are needlessly bad at this task, one's degree of Dutch-bookability is not *itself* a good measure of how well they recommend practically valuable actions. Strictly proper scoring rules are the right tools, I argued, for measuring both epistemic and practical value. I then showed that we can rerun Staffel's strategy swapping in strictly proper scoring rules for Dutch-bookability measures. So long as one's epistemic scoring rule and practical scoring rule are "sufficiently similar," there is some way of nudging incoherent credences toward coherence that is guaranteed to yield more of both types of value.

# 6 Appendix

## 6.1 Guidance Value and Strictly Proper Scoring Rules

Let the space of all binary decision problems be $\mathcal{D} = \mathbb{R}^4$. A point $\langle a, b, c, d \rangle$ in $\mathcal{D}$ represents the decision problem

|   | $X$ | $\neg X$ |
|---|-----|----------|
| $F$ | $a$ | $b$ |
| $G$ | $c$ | $d$ |

Let $\mu$ be a measure on $\mathcal{D}$ such that

$$\int_{\mathcal{D}} a \, \mathrm{d}\mu, \int_{\mathcal{D}} b \, \mathrm{d}\mu, \int_{\mathcal{D}} c \, \mathrm{d}\mu, \int_{\mathcal{D}} d \, \mathrm{d}\mu < \infty$$

and

$$\int_{\mathcal{R}} \mathrm{d}\mu > 0$$

for any non-degenerate region $\mathcal{R} \subseteq \mathcal{D}$. For any $x \in [0, 1]$ and any $D = \langle a, b, c, d \rangle \in \mathcal{D}$, let $\mathcal{B}(x, D)$ be the *Bayes act in D according to x*, *i.e.*, the act that maximises expected utility in $D$:

$$\mathcal{B}(x, D) = \arg\max_{H \in \{F, G\}} \left( x u(H, X) + (1 - x) u(H, \neg X) \right)$$

If $\mathcal{B}(x, D)$ is not unique (*i.e.*, if $(x)(a) + (1-x)(b) = (x)(c) + (1-x)(d)$), then we stipulate that $\mathcal{B}(x, D)$ is the mixed act $\lambda F + (1-\lambda)G$ for some $\lambda \in [0,1]$. The choice of $\lambda$ is unimportant since the expected utility of $F$, $G$, and any mixture of the two is the same when $(x)(a) + (1-x)(b) = (x)(c) + (1-x)(d)$.

Let the loss of $x$ in decision problem $D$ and state $X$ be

$$\mathcal{L}_D(x, X) = \max\{a, c\} - u(\mathcal{B}(x, D), X)$$

So the loss of $x$ in state $X$ is the difference between the utility that $x$'s recommended action $\mathcal{B}(x, D)$ produces in $X$ and the maximum possible utility achievable in state $X$. It represents how far $x$'s recommended action falls short of the objectively best action to perform in state $X$. Likewise, let

$$\mathcal{L}_D(x, \neg X) = \max\{b, d\} - u(\mathcal{B}(x, D), \neg X)$$

Finally, let

$$g_1(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) \, d\mu$$

$$g_0(x) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) \, d\mu$$

$g_1(x)$ represents how far $x$'s recommended action falls short of the objectively best action in state $X$ *on average*, across all possible decision problems, where the weight given to any class of decision problems is specified by $\mu$. Ditto for $g_0(x)$ (but in state $\neg X$). Note that given our choice of $\mu$, $g_1(x), g_0(x) < \infty$.

**Proposition 1.** *For any $x, p \in [0, 1]$ with $x \neq p$*

$$pg_1(x) + (1-p)g_0(x) > pg_1(p) + (1-p)g_0(p)$$

*Proof.* Observe that

$$g_1(x) - g_1(p) = \int_{\mathcal{D}} \mathcal{L}_D(x, X) - \mathcal{L}_D(p, X) \, d\mu \tag{19}$$

$$= \int_{\mathcal{D}} u(\mathcal{B}(p, D), X) - u(\mathcal{B}(x, D), X) \, d\mu \tag{20}$$

and

$$g_0(x) - g_0(p) = \int_{\mathcal{D}} \mathcal{L}_D(x, \neg X) - \mathcal{L}_D(p, \neg X) \, d\mu \tag{21}$$

$$= \int_{\mathcal{D}} u(\mathcal{B}(p, D), \neg X) - u(\mathcal{B}(x, D), \neg X) \, d\mu \tag{22}$$

So

$$p(g_1(x) - g_1(p)) + (1-p)(g_0(x) - g_0(p)) \tag{23}$$

$$= p \left( \int_{\mathcal{D}} u(\mathcal{B}(p, D), X) - u(\mathcal{B}(x, D), X) \, d\mu \right) \tag{24}$$

$$+ (1-p) \left( \int_{\mathcal{D}} u(\mathcal{B}(p, D), \neg X) - u(\mathcal{B}(x, D), \neg X) \, d\mu \right)$$

$$= \int_{\mathcal{D}} (p \, u(\mathcal{B}(p, D), X) + (1-p) \, u(\mathcal{B}(p, D), \neg X)) \tag{25}$$

$$- (p \, u(\mathcal{B}(x, D), X) + (1-p) \, u(\mathcal{B}(x, D), \neg X)) \, d\mu$$

By the definition of $\mathcal{B}(p, D)$, the inside term is always non-negative. And since $\mu$ places positive measure on every non-degenerate region, it is strictly positive on some measurable set. Hence

$$p(g_1(x) - g_1(p)) + (1 - p)(g_0(x) - g_0(p)) > 0 \tag{26}$$

$\square$

## 6.2   Local Dominance Relative to Distinct Scoring Rules

Let

$$f_1(x) = \int_x^1 (1 - t)\, m(t)\, \mathrm{d}t$$

$$f_0(x) = \int_0^x t\, m(t)\, \mathrm{d}t$$

$$\mathcal{I}_1(x, y) = f_1(x) + f_0(y)$$

$$\mathcal{I}_0(x, y) = f_0(x) + f_1(y)$$

and

$$g_1(x) = \int_x^1 (1 - t)\, n(t)\, \mathrm{d}t$$

$$g_0(x) = \int_0^x t\, n(t)\, \mathrm{d}t$$

$$\mathcal{I}_1^*(x, y) = g_1(x) + g_0(y)$$

$$\mathcal{I}_0^*(x, y) = g_0(x) + g_1(y)$$

Note that

$$\frac{\partial \mathcal{I}_1}{\partial x}(x, y) = (x - 1)m(x)$$

$$\frac{\partial \mathcal{I}_1}{\partial y}(x, y) = y\, m(y)$$

$$\frac{\partial \mathcal{I}_0}{\partial x}(x, y) = x\, m(x)$$

$$\frac{\partial \mathcal{I}_0}{\partial y}(x, y) = (y - 1)m(y)$$

$$\frac{\partial \mathcal{I}_1^*}{\partial x}(x, y) = (x - 1)n(x)$$

$$\frac{\partial \mathcal{I}_1^*}{\partial y}(x, y) = y\, n(y)$$

$$\frac{\partial \mathcal{I}_0^*}{\partial x}(x, y) = x\, n(x)$$

$$\frac{\partial \mathcal{I}_0^*}{\partial y}(x, y) = (y - 1)n(y)$$

Define

$$\mathrm{posi}(\{g_1, \dots, g_m\}) = \left\{ \sum_{i \leqslant m} a_i g_i \mid a_1, \dots, a_m \geqslant 0, \sum_{i \leqslant m} a_i > 0 \right\}$$

23

and

$$\text{span}(\{g_1, \dots, g_m\}) = \left\{ \sum_{i \leqslant m} a_i g_i \mid a_1, \dots, a_m \in \mathbb{R} \right\}$$

**Proposition 2.** *The following two conditions are equivalent:*

*1.*
$$0 \notin \text{posi}(\left\{ \left\langle \frac{\partial \mathcal{I}_1}{\partial x}(x, y), \frac{\partial \mathcal{I}_1}{\partial y}(x, y) \right\rangle, \left\langle \frac{\partial \mathcal{I}_0}{\partial x}(x, y), \frac{\partial \mathcal{I}_0}{\partial y}(x, y) \right\rangle \right\})$$

*2.* $y \neq 1 - x$

*Proof.* Suppose condition 1 holds but 2 does not. So $y = 1 - x$. Let $\alpha = x$ and $\beta = 1 - x$. Then

$$\alpha \frac{\partial \mathcal{I}_1}{\partial x}(x, y) + \beta \frac{\partial \mathcal{I}_0}{\partial x}(x, y) = x(x-1)m(x) + (1-x)x\, m(x) = 0 \qquad (27)$$

and

$$\alpha \frac{\partial \mathcal{I}_1}{\partial y}(x, y) + \beta \frac{\partial \mathcal{I}_0}{\partial y}(x, y) = x(1-x)m(1-x) + (1-x)(-x)m(1-x) = 0$$
$$(28)$$

which contradicts condition 1. Conversely, suppose that condition 2, *i.e.*,

$$0 \in \text{posi}(\left\{ \left\langle \frac{\partial \mathcal{I}_1}{\partial x}(x, y), \frac{\partial \mathcal{I}_1}{\partial y}(x, y) \right\rangle, \left\langle \frac{\partial \mathcal{I}_0}{\partial x}(x, y), \frac{\partial \mathcal{I}_0}{\partial y}(x, y) \right\rangle \right\})$$
$$= \text{posi}(\{\langle (x-1)m(x), ym(y) \rangle, \langle xm(x), (y-1)m(y) \rangle\})$$

For any $\alpha, \beta > 0$,

$$\text{posi}(\{\langle (x-1)m(x), ym(y) \rangle, \langle xm(x), (y-1)m(y) \rangle\})$$
$$= \text{posi}(\{\langle \alpha(x-1)m(x), \alpha ym(y) \rangle, \langle \beta xm(x), \beta(y-1)m(y) \rangle\})$$

Let

$$\alpha = \frac{1}{(1-x)m(x) + ym(y)}$$
$$\beta = \frac{1}{xm(x) + (1-y)m(y)}$$

Note that

$$\langle \alpha(x-1)m(x), \alpha ym(y) \rangle + \langle \beta xm(x), \beta(y-1)m(y) \rangle = \langle \kappa, \kappa \rangle \qquad (29)$$

where

$$\kappa = \frac{(x+y-1)m(x)m(y)}{((1-x)m(x) + ym(y))(xm(x) + (1-y)m(y))} \qquad (30)$$

Clearly then

$$0 \in \text{posi}(\{\langle \alpha(x-1)m(x), \alpha ym(y) \rangle, \langle \beta xm(x), \beta(y-1)m(y) \rangle\})$$

iff $\kappa = 0$ which holds iff $y = 1 - x$.

$\square$

**Proposition 3.** *The following two conditions are equivalent:*

1. *There is some $0 \leqslant p \leqslant 1$ s.t.*

$$\nabla_{\langle p-x,(1-p)-y\rangle}\mathcal{I}_1(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_1((1-\epsilon)x+\epsilon\, p,(1-\epsilon)y+\epsilon\,(1-p))-\mathcal{I}_1(x,y)\right] < 0$$

$$\nabla_{\langle p-x,(1-p)-y\rangle}\mathcal{I}_0(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_0((1-\epsilon)x+\epsilon\, p,(1-\epsilon)y+\epsilon\,(1-p))-\mathcal{I}_0(x,y)\right] < 0$$

$$\nabla_{\langle p-x,(1-p)-y\rangle}\mathcal{I}_1^*(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_1^*((1-\epsilon)x+\epsilon\, p,(1-\epsilon)y+\epsilon\,(1-p))-\mathcal{I}_1^*(x,y)\right] < 0$$

$$\nabla_{\langle p-x,(1-p)-y\rangle}\mathcal{I}_0^*(x,y) = \lim_{\epsilon\to 0}\frac{1}{\epsilon}\left[\mathcal{I}_0^*((1-\epsilon)x+\epsilon\, p,(1-\epsilon)y+\epsilon\,(1-p))-\mathcal{I}_0^*(x,y)\right] < 0$$

2.

$$0 \notin posi\left(\left\{\left\langle\frac{\partial\mathcal{I}_1}{\partial x}(x,y),\frac{\partial\mathcal{I}_1}{\partial y}(x,y)\right\rangle,\left\langle\frac{\partial\mathcal{I}_0}{\partial x}(x,y),\frac{\partial\mathcal{I}_0}{\partial y}(x,y)\right\rangle,\right.\right.$$
$$\left.\left.\left\langle\frac{\partial\mathcal{I}_1^*}{\partial x}(x,y),\frac{\partial\mathcal{I}_1^*}{\partial y}(x,y)\right\rangle,\left\langle\frac{\partial\mathcal{I}_0^*}{\partial x}(x,y),\frac{\partial\mathcal{I}_0^*}{\partial y}(x,y)\right\rangle\right\}\right)$$
$$= posi\left(\left\{\langle(x-1)m(x),ym(y)\rangle,\langle xm(x),(y-1)m(y)\rangle,\right.\right.$$
$$\left.\left.\langle(x-1)n(x),yn(y)\rangle,\langle xn(x),(y-1)n(y)\rangle\right\}\right)$$

*Proof.* Suppose that condition 1 holds. Since in general $\nabla_{\langle a,b\rangle}\phi(x,y) = a\frac{\partial\phi}{\partial x}(x,y)+b\frac{\partial\phi}{\partial y}(x,y)$, condition 1 holds iff

$$(p-x)(x-1)m(x) + ((1-p)-y)ym(y) < 0 \tag{31}$$
$$(p-x)xm(x) + ((1-p)-y)(y-1)m(y) < 0 \tag{32}$$
$$(p-x)(x-1)n(x) + ((1-p)-y)yn(y) < 0 \tag{33}$$
$$(p-x)xn(x) + ((1-p)-y)(y-1)n(y) < 0 \tag{34}$$

Suppose that condition 2 does not hold. Then there are $\alpha,\beta,\epsilon,\delta \geqslant 0$ with $\alpha+\beta+\epsilon+\delta > 0$ such that

$$\alpha(x-1)m(x) + \beta xm(x) + \epsilon(x-1)n(x) + \delta xn(x) = 0 \tag{35}$$
$$\alpha ym(y) + \beta(y-1)m(y) + \epsilon yn(y) + \delta(y-1)n(y) = 0 \tag{36}$$

But (31)-(34) imply

$$\alpha(p-x)(x-1)m(x) + \alpha((1-p)-y)ym(y) \leqslant 0 \tag{37}$$
$$\beta(p-x)xm(x) + \beta((1-p)-y)(y-1)m(y) \leqslant 0 \tag{38}$$
$$\epsilon(p-x)(x-1)n(x) + \epsilon((1-p)-y)yn(y) \leqslant 0 \tag{39}$$
$$\delta(p-x)xn(x) + \delta((1-p)-y)(y-1)n(y) \leqslant 0 \tag{40}$$

with at least one of (37)-(40) strict. This implies

$$(p-x)\left(\alpha(x-1)m(x) + \beta xm(x) + \epsilon(x-1)n(x) + \delta xn(x)\right) \tag{41}$$
$$((1-p)-y)\left(\alpha ym(y) + \beta(y-1)m(y) + \epsilon yn(y) + \delta(y-1)n(y)\right) < 0$$

which contradicts (35)-(36).

Now suppose that condition 2 holds. Let

$$A = \mathrm{posi}\left(\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x),(y-1)m(y)\rangle,\right.$$
$$\left.\langle(x-1)n(x), yn(y)\rangle, \langle xn(x),(y-1)n(y)\rangle\}\right) \cup \{0\}$$
$$B = \{0\}$$

Let $A' = A \cap -A$. By condition 2, $A' = B = \{0\}$. So by Klee's separation theorem, there are $\alpha, \beta \in \mathbb{R}$ such that $\alpha\,a + \beta\,b < 0$ for all $\langle a, b\rangle \in A \setminus A'$ (Klee, 1955, Theorem 2.5). In particular then

$$\alpha(x-1)m(x) + \beta ym(y) < 0 \tag{42}$$
$$\alpha xm(x) + \beta(y-1)m(y) < 0 \tag{43}$$
$$\alpha(x-1)n(x) + \beta yn(y) < 0 \tag{44}$$
$$\alpha xn(x) + \beta(y-1)n(y) < 0 \tag{45}$$

Since condition 2 holds, we know from proposition 2 that $y \neq 1 - x$. If $x + y > 1$, then the proof of proposition 2 establishes that

$$\{\langle a, b\rangle \mid a, b \geqslant 0, a + b > 0\} \subseteq \mathrm{posi}(\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x),(y-1)m(y)\rangle\})$$
$$\subseteq \mathrm{posi}\left(\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x),(y-1)m(y)\rangle,\right.$$
$$\left.\langle(x-1)n(x), yn(y)\rangle, \langle xn(x),(y-1)n(y)\rangle\}\right)$$

In that case, we must have $\alpha, \beta \leqslant 0$ with at least one strict. Similarly, if $x + y < 1$, then the proof of proposition 2 establishes that

$$\{\langle a, b\rangle \mid a, b \leqslant 0, a + b < 0\} \subseteq \mathrm{posi}(\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x),(y-1)m(y)\rangle\})$$
$$\subseteq \mathrm{posi}\left(\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x),(y-1)m(y)\rangle,\right.$$
$$\left.\langle(x-1)n(x), yn(y)\rangle, \langle xn(x),(y-1)n(y)\rangle\}\right)$$

In that case, we must have $\alpha, \beta \geqslant 0$ with at least one strict. Either way, let

$$p = \frac{\alpha(1-y) + \beta\,x}{\alpha + \beta} \in [0, 1]$$

Then (31) holds iff

$$\frac{1 - x - y}{\alpha + \beta}\left(\alpha(x-1)m(x) + \beta ym(y)\right) < 0 \tag{46}$$

(32) holds iff

$$\frac{1 - x - y}{\alpha + \beta}\left(\alpha xm(x) + \beta(y-1)m(y)\right) < 0 \tag{47}$$

(33) holds iff

$$\frac{1 - x - y}{\alpha + \beta}\left(\alpha(x-1)n(x) + \beta yn(y)\right) < 0 \tag{48}$$

And (34) holds iff

$$\frac{1 - x - y}{\alpha + \beta}\left(\alpha xn(x) + \beta(y-1)n(y)\right) < 0 \tag{49}$$

And whether we have (i) $x + y > 1$, $\alpha, \beta \leqslant 0$ and $\alpha + \beta < 0$ or (ii) $x + y < 1$, $\alpha, \beta \geqslant 0$ and $\alpha + \beta > 0$, we end up with

$$\frac{1 - x - y}{\alpha + \beta} > 0$$

Hence (42)-(45) and proposition 2 jointly imply (31)-(34).

$\square$

**Proposition 4.** *If $x + y > 1$ then*

$$\frac{xy}{(1-x)(1-y)} > \frac{m(x)/n(x)}{m(y)/n(y)}, \frac{m(y)/n(y)}{m(x)/n(x)}$$

*iff*

$$0 \notin posi\left(\left\{\left\langle \frac{\partial \mathcal{I}_1}{\partial x}(x,y), \frac{\partial \mathcal{I}_1}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0}{\partial x}(x,y), \frac{\partial \mathcal{I}_0}{\partial y}(x,y)\right\rangle, \right.\right.$$
$$\left.\left.\left\langle \frac{\partial \mathcal{I}_1^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_1^*}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_0^*}{\partial y}(x,y)\right\rangle\right\}\right)$$

*Similarly if $x + y < 1$ then*

$$\frac{xy}{(1-x)(1-y)} < \frac{m(x)/n(x)}{m(y)/n(y)}, \frac{m(y)/n(y)}{m(x)/n(x)}$$

*iff*

$$0 \notin posi\left(\left\{\left\langle \frac{\partial \mathcal{I}_1}{\partial x}(x,y), \frac{\partial \mathcal{I}_1}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0}{\partial x}(x,y), \frac{\partial \mathcal{I}_0}{\partial y}(x,y)\right\rangle, \right.\right.$$
$$\left.\left.\left\langle \frac{\partial \mathcal{I}_1^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_1^*}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_0^*}{\partial y}(x,y)\right\rangle\right\}\right)$$

*Proof.* For any $\alpha, \beta, \epsilon, \delta > 0$,

$$\text{posi}\left(\left\{\left\langle \frac{\partial \mathcal{I}_1}{\partial x}(x,y), \frac{\partial \mathcal{I}_1}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0}{\partial x}(x,y), \frac{\partial \mathcal{I}_0}{\partial y}(x,y)\right\rangle, \right.\right.$$
$$\left.\left.\left\langle \frac{\partial \mathcal{I}_1^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_1^*}{\partial y}(x,y)\right\rangle, \left\langle \frac{\partial \mathcal{I}_0^*}{\partial x}(x,y), \frac{\partial \mathcal{I}_0^*}{\partial y}(x,y)\right\rangle\right\}\right)$$
$$= \text{posi}\left(\left\{\langle(x-1)m(x), ym(y)\rangle, \langle xm(x), (y-1)m(y)\rangle, \right.\right.$$
$$\left.\left.\langle(x-1)n(x), yn(y)\rangle, \langle xn(x), (y-1)n(y)\rangle\right\}\right)$$
$$= \text{posi}\left(\left\{\langle\alpha(x-1)m(x), \alpha ym(y)\rangle, \langle\beta xm(x), \beta(y-1)m(y)\rangle, \right.\right.$$
$$\left.\left.\langle\epsilon(x-1)n(x), \epsilon yn(y)\rangle, \langle\delta xn(x), \delta(y-1)n(y)\rangle\right\}\right)$$

Let

$$\alpha = \frac{1}{(1-x)m(x) + ym(y)}$$
$$\beta = \frac{1}{xm(x) + (1-y)m(y)}$$
$$\epsilon = \frac{1}{(1-x)n(x) + yn(y)}$$
$$\delta = \frac{1}{xn(x) + (1-y)n(y)}$$

27

Note then that

$$\langle \alpha(x-1)m(x), \alpha ym(y) \rangle + \langle \beta xm(x), \beta(y-1)m(y) \rangle = \langle \kappa_1, \kappa_1 \rangle \quad (50)$$

$$\langle \alpha(x-1)m(x), \alpha ym(y) \rangle + \langle \delta xn(x), \delta(y-1)n(y) \rangle = \langle \kappa_2, \kappa_2 \rangle \quad (51)$$

$$\langle \epsilon(x-1)n(x), \epsilon yn(y) \rangle + \langle \beta xm(x), \beta(y-1)m(y) \rangle = \langle \kappa_3, \kappa_3 \rangle \quad (52)$$

$$\langle \epsilon(x-1)n(x), \epsilon yn(y) \rangle + \langle \delta xn(x), \delta(y-1)n(y) \rangle = \langle \kappa_4, \kappa_4 \rangle \quad (53)$$

for some constants, $\kappa_1, \kappa_2, \kappa_3, \kappa_4$. Hence

$$0 \notin \mathrm{posi}\left(\{\langle \alpha(x-1)m(x), \alpha ym(y) \rangle, \langle \beta xm(x), \beta(y-1)m(y) \rangle, \right.$$
$$\left. \langle \epsilon(x-1)n(x), \epsilon yn(y) \rangle, \langle \delta xn(x), \delta(y-1)n(y) \rangle\}\right)$$

iff either (i) $\kappa_1, \kappa_2, \kappa_3, \kappa_4 > 0$ or (ii) $\kappa_1, \kappa_2, \kappa_3, \kappa_4 < 0$. And

$$\kappa_1 = \frac{(x+y-1)m(y)}{((1-x)m(x)+ym(y))(xm(x)+(1-y)m(y))} \quad (54)$$

$$\kappa_2 = \frac{(x-1)m(x)}{(1-x)m(x)+ym(y)} + \frac{xn(x)}{xn(x)+(1-y)n(y)} \quad (55)$$

$$\kappa_3 = \frac{xm(x)}{xm(x)+(1-y)m(y)} + \frac{(x-1)n(x)}{(1-x)n(x)+yn(y)} \quad (56)$$

$$\kappa_4 = \frac{(x+y-1)n(y)}{((1-x)n(x)+yn(y))(xn(x)+(1-y)n(y))} \quad (57)$$

Now, if $x+y > 1$, then $\kappa_1, \kappa_4 > 0$. So we must have $\kappa_2, \kappa_3 > 0$ as well. And $\kappa_2 > 0$ iff

$$\frac{xn(x)}{xn(x)+(1-y)n(y)} > \frac{(1-x)m(x)}{(1-x)m(x)+ym(y)} \quad (58)$$

iff

$$xyn(x)m(y) > (1-x)(1-y)m(x)n(y) \quad (59)$$

iff

$$\frac{xy}{(1-x)(1-y)} > \frac{m(x)/n(x)}{m(y)/n(y)} \quad (60)$$

Similarly $\kappa_3 > 0$ iff

$$\frac{xy}{(1-x)(1-y)} > \frac{m(y)/n(y)}{m(x)/n(x)} \quad (61)$$

Likewise, if $x+y < 1$, then $\kappa_1, \kappa_4 < 0$. So we must have $\kappa_2, \kappa_3 < 0$ as well, which holds iff

$$\frac{xy}{(1-x)(1-y)} < \frac{m(x)/n(x)}{m(y)/n(y)}, \frac{m(y)/n(y)}{m(x)/n(x)} \quad (62)$$

$\square$

# References

R. T. Cox. *The Algebra of Probable Inference*. John Hopkins Press, Baltimore, 1961.

Bruno de Finetti. *Theory of Probability. A Critical Introductory Treatment*. John Wiley & Sons, 1974.

Kenny Easwaran and Branden Fitelson. An 'evidentialist' worry about joyce's argument for probabilism. *Dialectica*, 66(3):425–433, 2012.

I. J. Good. *Probability and the weighing of evidence*. Griffin, London, 1950.

J Joyce. A nonpragmatic vindication of probabilism. *Philosophy of Science*, 65 (4):575–603, 1998.

James Joyce. The development of subjective bayesianism. In Dov Gabbay, Stephan Hartmann, and John Woods, editors, *Handbook of the History of Logic*, volume 10. Elsevier, 2009.

James Joyce. Why evidentialists need not worry about the accuracy argument for probabilism. *Ms.*, pages 1–29, 2013.

James M Joyce. Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In Franz Huber and Christoph Schmidt-Petri, editors, *Degrees of Belief*, volume 342. Springer, Dordrecht, 2009.

V. L. Klee. Separation properties of convex cones. *Proceedings of the American Mathematical Society*, 6(2):313–318, 1955.

Jason Konek. Comparative probabilities. In Richard Pettigrew and Jonathan Weisberg, editors, *The Open Handbook of Formal Epistem- ology*, pages 267–348. PhilPapers Foundation, 2019.

B. O. Koopman. The bases of probability. *Bulletin of the American Mathematical Society*, 46:763–774, 1940a.

B. O. Koopman. The axioms of algebra of intuitive probability. *Annals of Mathematics*, 41(2):269–292, 1940b.

David Krantz, R. D. Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement Vol. I: Additive and Polynomial Representations*. Academic Press, New York and London, 1971.

B.A. Levinstein. A pragmatist's guide to epistemic utility. *Philosophy of Science*, 84:613–638, 2017.

Richard Pettigrew. *Accuracy and the Laws of Credence*. Oxford University Press, Oxford, 2016.

Richard Pettigrew. *The Dutch Book Arguments*. Cambridge University Press, 2020.

Joel Predd, Robert Seiringer, Elliott Lieb, Daniel Osherson, H. Vincent Poor, , and Sanjeev R. Kulkarni. Probabilistic coherence and proper scoring rules. *IEEE Transaction on Information Theory*, 55(10):4786–4792, 2009.

F. P. Ramsey. Truth and probability. In R. B. Braithwaite, editor, *The Foundations of Mathematics and other Logical Essays*, chapter VII, pages 156–198. Kegan, Paul, Trench, Trubner & Co., London, 1931. Reprinted in **?** and **?**.

L.J. Savage. Elicitation of personal probabilities and expectations. *Journal of the American Statistical Association*, 66(336):783–801, 1971.

Mark Schervish. A general method for comparing probability assessors. *The Annals of Statistics*, 17(4):1856–1879, 1989.

Mark Schervish, Teddy Seidenfeld, and Joseph Kadane. Measures of incoherence: How not to gamble if you must. In Jose Bernardo et al., editor, *Bayesian Statistics 7*, pages 385–401. Oxford University Press, Oxford, 2003.

Mark Schervish, Teddy Seidenfeld, and Joseph Kadane. Infinite previsions and finitely additive expectations. *arXiv: Statistics Theory*, 2014a.

Mark Schervish, Teddy Seidenfeld, and Joseph Kadane. Dominating countably many forecasts. *The Annals of Statistics*, 42(2):728–756, 2014b.

Julia Staffel. *Unsettled Thoughts: A Theory of Degrees of Rationality*. Oxford University Press, 2019.