

## An ‘Evidentialist’ Worry About Joyce’s Argument for Probabilism

Kenny EASWARAN<sup>†</sup> and Branden FITELSON<sup>‡</sup>

### ABSTRACT

Joyce (1998) argues that for any credence function that doesn’t satisfy the probability axioms, there is another function that dominates it in terms of accuracy. But if some potential credence functions are ruled out as violations of the Principal Principle, then some non-probabilistic credence functions fail to be dominated. We argue that to fix Joyce’s argument, one must show that all epistemic values for credence functions derive from accuracy.

### 1. Some background on arguments for probabilism

*Probabilism* asserts that every epistemically rational agent  $S$ ’s ‘degrees of confidence’ (a.k.a., ‘degrees of credence’ or ‘*credences*’) should be faithfully representable via some *probability function*  $b$  from the set of propositions in  $S$ ’s doxastic space to the real numbers. All of the arguments for probabilism that one finds in the literature are of the following (rough-and-ready) general form:

- An agent  $S$  has a non-probabilistic degree of belief function  $b$  iff ( $\Leftrightarrow$ )  $S$  has some ‘bad’ property  $B$  – in virtue of the fact that their credence function  $b$  has a ‘bad’ formal property  $F$ .<sup>1</sup>

These arguments rest on *Theorems* ( $\Rightarrow$ ) and *Converse Theorems* ( $\Leftarrow$ ), which establish that  $b$  is non-probabilistic iff ( $\Leftrightarrow$ )  $b$  has some specific (‘bad’) formal property  $F$ . Here are two well-known examples:

- **Dutch Book Arguments.**  $B$  is *susceptibility to sure monetary loss* (in a gambling scenario), and  $F$  is the formal role played by non-probabilistic  $b$ ’s in the Dutch Book Theorem (DBT) and its Converse.
- **Representation Theorem Arguments.**  $B$  is *having preferences that violate some (Savage-style) rational constraints* (and/or *being unrepresentable as an expected utility maximizer*), and  $F$  is the formal role played by non-probabilistic  $b$ ’s in some (Savage-style) Representation Theorem.

To the extent that we have reasons to avoid these ‘bad  $B$ -properties’, these arguments provide reasons not to have an incoherent credence function  $b$  – and perhaps

<sup>†</sup> School of Philosophy, University of Southern California; Email: easwaran@usc.edu;

<sup>‡</sup> Department of Philosophy, Rutgers University; Email: branden@fitelson.org.

<sup>1</sup> This characterization of arguments for probabilism is adapted from Hájek (2008).

even reasons *to* have a coherent one. But, note that these two traditional arguments for probabilism involve what might be called ‘pragmatic’ reasons (not) to be (in)coherent. In the case of the Dutch Book argument, the ‘bad’ property is *pragmatically* bad (to the extent that one values money). But, it is not clear whether the DBA pinpoints any *epistemic* defect of incoherent agents. The same can be said for Representation Theorem arguments, since they involve the structure of an agent’s *preferences*.

A recent argument for probabilism, put forward by James Joyce, can be cast in similar terms. That is, it can be seen as identifying a ‘bad property’ of incoherent agents, which arises in virtue of a ‘bad formal property’ of their credence function. However, Joyce’s argument aims to identify an *epistemic* defect shared by all (and only) incoherent agents. In the next section, we’ll give a brief primer on Joyce’s argumentative strategy. We will keep the discussion as simple and non-technical as possible, since our present worry is rather basic and fundamental.

## 2. Joyce’s argument for probabilism – the basic ideas

For present purposes, we can think of Joyce’s argument for probabilism as a variant of an older argument due to de Finetti. What de Finetti (1974) showed is that if a credence function  $b$  is incoherent, then there exists a coherent function  $b'$  that is (in one precise sense) *strictly more accurate – in all possible worlds*.<sup>2</sup> Let’s keep things maximally simple. Consider a toy agent  $S$  whose language  $\mathcal{L}$  contains only a single atomic sentence  $P$ . We will assume that  $S$  is ‘logically omniscient’ in the sense that  $S$  recognizes all logical equivalences in  $\mathcal{L}$ . Thus,  $S$ ’s doxastic space contains only four propositions  $\{P, \neg P, \top, \perp\}$ , where  $\top$  represents an arbitrary tautological statement, and  $\perp$  represents an arbitrary contradictory statement. We will also assume that  $S$  assigns credence 1 to  $\top$  and credence 0 to  $\perp$ . Thus, the question of  $S$ ’s coherence reduces to the question of whether  $S$ ’s credence function  $b$  satisfies the following pair of constraints:

- $b(P) \in [0, 1]$  and  $b(\neg P) \in [0, 1]$ .
- $b(P) + b(\neg P) = 1$ .

Next, let’s think about how we might ‘score’ a credence function, in terms of its ‘distance from the truth (or inaccuracy) in a possible world’. For our toy agent, there are only two relevant possible worlds:  $w_1$  in which  $P$  is false, and  $w_2$  in which  $P$  is true. If we use the number 1 to ‘numerically represent’ the truth-value true (at a world) and the number 0 to ‘numerically represent’ the truth-value false (at a

<sup>2</sup> Strictly speaking, de Finetti did not interpret the Brier score (his favored scoring rule) as a measure of ‘inaccuracy’. Joyce (1998; 2009) was the first to give this interpretation to de Finetti’s argument. But, our objection is applicable to any argument with the *formal* structure of de Finetti’s (or Joyce’s). So, this issue of interpretation is actually inessential for our purposes here.

world), then we can 'score' a credence function  $b$  using a *scoring rule* which is some function of (i) the values  $b$  assigns to  $P$  and  $\neg P$ , and (ii) the 'numerical truth-values' of  $P$  and  $\neg P$  at the two relevant possible worlds  $w_1$  and  $w_2$ . It is standard in this context (beginning with de Finetti) to use what is called the *Brier score* (of a credence function  $b$ , at a world  $w$ ), which is the sum of squares of differences between credences and truth values. For our toy agent  $S$ , it is defined in the following way:

- The Brier score of  $b$  in  $w_1$   $(0 - b(P))^2 + (1 - b(\neg P))^2$ .
- The Brier score of  $b$  in  $w_2$   $(1 - b(P))^2 + (0 - b(\neg P))^2$ .

The idea behind all such scoring rules is that 'distance from truth' or 'inaccuracy' of a credence function  $b$  (at a world  $w$ ) is measured in terms of  $b$ 's 'distance (at  $w$ ) from the numerical truth-values' of the set of propositions in the agent's doxastic space. There is a rather vast literature on alternative scoring rules, but none of those controversies about how to measure 'accuracy' or 'verisimilitude' of a credence function will be important for present purposes. The worry we describe below will not depend on which scoring rule one adopts. So, for simplicity, we will just assume the Brier score (which is acceptable to both Joyce and de Finetti).

With these basics in mind, we can now cast Joyce's argument for probabilism in the standard 'mould' from above. For Joyce, the 'bad' property  $B$  that an incoherent agent succumbs to is the property of being *accuracy-dominated*. More precisely, the formal property  $F$  that underlies this 'accuracy-domination' is the existence of a coherent credence function  $b'$  that has a *strictly lower Brier score – in every possible world*.

Unlike the traditional arguments for probabilism, Joyce thinks of his argument as *non-pragmatic*. This is because he thinks of 'accuracy' as a *purely epistemic aim*, and he thinks of the Brier score as an adequate measure of 'inaccuracy'. Thus, if there exists a credence function  $b'$  with a lower Brier score (hence, lower 'inaccuracy') than yours ( $b$ ) – in every possible world – then this is supposed to reveal an *epistemic defect* of your credence function  $b$ .

Furthermore, Joyce's argument appears (at first) to say more – it gives an incoherent agent some *specific* coherent functions  $b'$  that should 'look *epistemically better*' to him than his current  $b$ .<sup>3</sup> Traditional arguments for probabilism do no such thing.<sup>4</sup> This seems to make Joyce's (or de Finetti's) argument 'more

<sup>3</sup> This is merely a comparative claim. Nothing about the argument suggests that an agent whose current credence function is  $b$  should specifically adopt  $b'$ . For all this argument says, some alternative  $b''$  may be best overall.

<sup>4</sup> For instance, the Dutch Book argument only implies that you'd be (pragmatically) 'better off' if you were coherent *rather than incoherent*. It doesn't single out any *specific set* of coherent functions (among the totality of such functions) that should 'look (pragmatically) better' to you. As far as the Dutch Book argument is concerned, if you're incoherent then *all* coherent functions should 'look (pragmatically) better' to you – and to precisely the same extent.

informative' than traditional arguments. As we'll see shortly, this 'additional informativeness' of Joyce's approach opens Joyce's argument up to a novel objection, not faced by traditional arguments.

### 3. *The worry – conflicts with evidential norms for credences*

Suppose an agent  $S$  is trying to figure out which candidate credence functions to rule out. In this process,  $S$  might appeal to various sorts of norms and considerations. For instance,  $S$  might want to rule out credence functions that are susceptible to a Dutch Book (DB). Moreover,  $S$  might want to avoid credence functions that violate the Principal Principle (PP) (Lewis 1980), given her current knowledge concerning objective chance.

(PP) If  $S$  knows that the objective chance of  $p$  is less than  $r$ , and  $S$  has no inadmissible evidence regarding  $p$ , then  $S$  should not assign a credence greater than  $r$  to  $p$ .

Suppose (1)  $S$  rules out those credence functions that are susceptible to Dutch Book (DB), and then (2)  $S$  rules out those credence functions that violate (PP) given her current knowledge  $K$  concerning objective chance. This two-step procedure will rule out the same set of credence functions as the procedure which performs (2) first, and then (1). Specifically, no non-probabilistic  $b$ 's will survive either two-step ruling out procedure. In this sense, *the order in which the two norms (DB) and (PP) are applied does not matter*.

Interestingly, this order-independence property is *violated* by 'accuracy-dominance' (AD) norms like de Finetti's and Joyce's (as well as other 'scoring-rule-based' dominance norms). For instance, suppose (a)  $S$  rules out those credence functions that are (Brier) accuracy-dominated (AD) by a credence function that is not yet ruled out, *and then* (b)  $S$  rules out those credence functions that violate (PP) given her current knowledge  $K$  concerning objective chance. No non-probabilistic credence functions  $b$  can survive *this* two-step procedure. But, if we *reverse the order* – that is, if we perform (b) first *and then* (a) – then some non-probabilistic credence functions *can* survive.

Here is a simple example that illustrates this *interaction/order-effect*. Suppose  $S$ 's background knowledge  $K$  contains (*exactly*) the following information about the chance of  $P$  (and *no inadmissible evidence*):

( $K$ ) The objective chance of  $P$  is at most 0.2.

We can understand the effect by looking at the diagrams in Figure 1. The square represents the set of credence functions taking values between 0 and 1, with  $x$ -axis representing the agent's credence in  $P$  and the  $y$ -axis representing the agent's credence function in  $\neg P$ . For a given credence function  $b$  (represented by the dot)

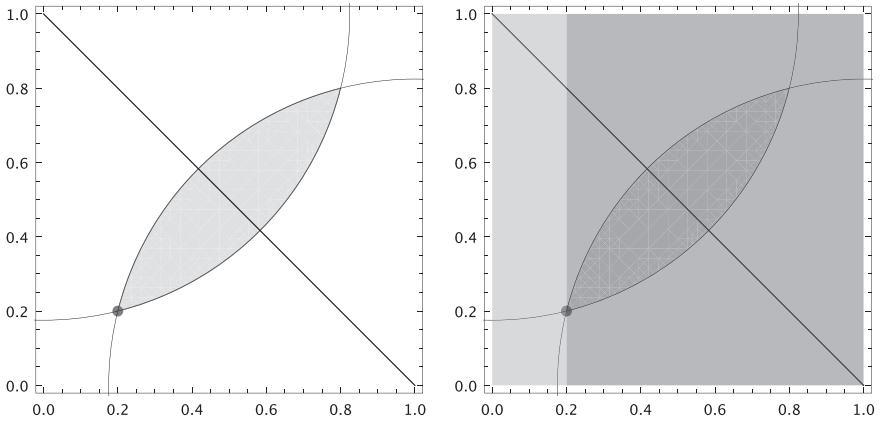


Figure 1. Visualizing the order-dependence of (AD) and (PP).

the two circular arcs delineate the regions that are at least as accurate as  $b$  in worlds  $w_2$  (the upper-left corner) and  $w_1$  (the lower-right corner). The shaded region in the left diagram then represents the set of credence functions that accuracy-dominate  $b$ . The region to the left of 0.2 ( $x$ -axis) in the right diagram represents the credence functions that satisfy (PP) given knowledge  $K$ .

Thus, if the agent first applies (a), she will rule out every credence function that is not on the diagonal line, because they have a non-empty shaded region on the left diagram. Applying (b) second, she will then rule out every remaining credence function in the region to the right of 0.2 ( $x$ -axis) on the right diagram, leaving just the upper-left part of the diagonal line.

However, if the agent first applies (b), she will rule out every credence function in the region to the right of 0.2 ( $x$ -axis) on the right diagram. When she then applies (a), what happens will be different. For a credence function like  $b$  (which survives application of (b), since it is part of the region to the left of 0.2 ( $x$ -axis)), (a) will not say anything, because the only  $b'$  that dominate it have already been ruled out by (b). As the lemma below will show, the functions that survive this order of application will be the ones on the upper-left part of the diagonal line, *and also all the ones on the border of the two shaded regions in the right diagram that are below the diagonal line.*

In the toy case where the algebra consists only of four propositions, and credences in  $\top$  and  $\perp$  are fixed at 1 and 0 respectively, the following result applies:

LEMMA. If  $b'$  dominates  $b$ , and both take values only in  $[0,1]$ , then:

*either  $b(P) > b'(P)$  and  $b(\neg P) > b'(\neg P)$  or  $b(P) < b'(P)$  and  $b(\neg P) < b'(\neg P)$ .*

PROOF. If  $b'$  dominates  $b$ , then  $b'$  must have lower inaccuracy both in  $w_1$  and in  $w_2$ . The inaccuracy of a credence function in  $w_1$  is the sum of two terms. If  $b'(P) > b(P)$ , then the first term is greater for  $b'$  than for  $b$ . Thus, if  $b'$  is less inaccurate than  $b$  in  $w_1$ , then the second term must be lower for  $b'$  than for  $b$ . But this means that  $b(\neg P) > b'(\neg P)$ . By considering the two terms summing to the inaccuracy in  $w_2$ , we can show that if  $b'(P) < b(P)$ , then  $b'(\neg P) < b(\neg P)$ . Thus, if  $b'(P) \neq b(P)$ , then  $b'(\neg P)$  must also differ from  $b(\neg P)$ , in the same direction. Similar reasoning establishes the converse.

An agent with credence function  $b$  will evaluate herself as having violated a norm if she applies (a) before (b), but not if she applies (b) before (a).<sup>5</sup>

In fact, our worry is much more general than this example involving (PP) suggests. Joyce's argument tacitly presupposes that – for any incoherent agent  $S$  with credence function  $b$  – some (coherent) functions  $b'$  that Brier-dominate  $b$  are always 'available' as 'permissible alternative credences' for  $S$ . But, there are various reasons why this may not be the case.<sup>6</sup> The agent could have good reasons for adopting (or sticking with) *some* of their credences. And, if they do, then the fact that some accuracy-dominating (coherent) functions  $b'$  'exist' (in an abstract mathematical sense) may not be epistemologically probative, from their current epistemic perspective. Thus, our use of the Principal Principle (PP) is merely one illustration of this more general phenomenon. Therefore, while one may object to our use of (PP) here for various reasons,<sup>7</sup> our worry will remain pressing, provided only that the following sorts of cases are possible:

- (★) Cases in which (a)  $S$  is incoherent, (b)  $S$  assigns  $b(p) \in \mathcal{I}$ , for some  $p$  and some interval  $\mathcal{I}$ , (c)  $S$  has good reason to believe (or even *knows*) that

<sup>5</sup> What we have here is a conflict between *evidential* norms for credences and a certain (accuracy-dominance) *coherence* norm for credences. This is analogous to conflicts that can arise between evidential and coherence norms in the case of full belief (e.g., the preface case). Niko Kolodny 2007 has argued that there are *only* evidential norms, and that 'coherence norms' do not really exist. We take it that our worry is something that Kolodny would find compelling (and welcome). While we do not wish to take a stand on this issue here, we do think that these sorts of *epistemic priority* questions are crucial in this context.

<sup>6</sup> Presently, we're concerned with *evidential* reasons why such  $b$ 's may be unavailable to an agent. There may also be *psychological* reasons why some 'alternative  $b'$ -functions' may be unavailable, but we are bracketing that possibility here.

<sup>7</sup> For instance, the (PP) was originally intended (Lewis 1980) to be applied to agents with 'reasonable' – indeed, *probabilistic* – credence functions. In our case, the agent doesn't recognize, for instance, that  $K$  entails that the objective chance of  $\neg P$  is at least 0.8. While this may be hard to imagine for such a small example, if the algebra is sufficiently large, then it becomes plausible that the agent won't recognize all the constraints on objective chance.

In addition, the (PP) was originally intended to be applied to *initial* credence functions, which would not be informed by specific bodies of empirical knowledge regarding objective chances (e.g., our  $K$  above). For these reasons, our present applications of (PP) are not (strictly speaking) kosher.

Ultimately, however, our worry will remain – so long as examples satisfying (★) are possible (see below). And, we see no reason to doubt that such examples exist.

epistemic rationality *requires*  $b(p) \in \mathcal{I}$ , but (d) all the (coherent) credence functions  $b'$  that Brier-dominate  $S$ 's credence function  $b$  are such that  $b'(p) \notin \mathcal{I}$ .

We have tried to describe a simple, toy case satisfying (★) – by making use of (PP). Even if one thinks our toy (PP) example is infelicitous (see fn. 7), this won't be enough to make our worry go away. In order to avoid our worry completely, one would need to argue that no examples satisfying (★) are possible. And, *that* is a tall order. Surely, we can imagine that an *oracle* concerning epistemic rationality has informed  $S$  that  $b(p) \in \mathcal{I}$  is *required* – despite the fact that all (coherent) Brier-dominating functions  $b'$  are such that  $b'(p) \notin \mathcal{I}$ . While such cases are fanciful, it seems to us that they are sufficient to motivate our worry.

It is interesting to note that Dutch Book arguments do *not* have this feature. As far as (DB) is concerned, it doesn't matter if you have good reasons for sticking with some of your credences. Suppose you do. Nonetheless, it remains true that if (and only if) you're incoherent, you're susceptible to Dutch Book. And, this gives you some reason (albeit a pragmatic reason) to change your *other* credences, so as to bring yourself into a coherent doxastic state. In the example(s) depicted in Figure 1, for instance, (DB) would give  $S/S'$  some reason (albeit a pragmatic one) to change their credences to a probabilistic function in the region to the left of 0.2 [ $x$ -axis] in the right diagram. So, this 'order-dependence' is a *peculiarity* of 'accuracy-dominance'-based approaches to probabilism.

#### 4. Modesty and propriety

A version of this worry applies to another argument Joyce makes on a related topic. In Joyce (2009), he wants to consider inaccuracy measures other than the Brier score. For any inaccuracy measure, he says that a credence function  $b$  is *modest* if it assigns a lower expected inaccuracy to some credence function other than itself. He says,

Modest credences, it can be argued, are epistemically defective because they undermine their own adoption and use. . . . If, relative to a person's own credences, some alternative system of beliefs has a lower expected epistemic disutility, then, by her own estimation, that system is preferable from the epistemic perspective. This puts her in an untenable doxastic situation. (Ib., 277)

He uses this argument to assert a principle,

IMMODESTY: An epistemic scoring rule  $S$  should not render any credences modest when there are epistemic circumstances under which those credences are clearly the rational ones to hold. (Ib., 278)

He then uses the Principal Principle (PP) to argue that *any* coherent set of credences  $\mathfrak{b}$  *could*, in certain circumstances (*viz.*, circumstances in which *S* knows that the salient objective chance function is identical to  $\mathfrak{b}$ ), be clearly the rational ones to hold, which then puts a constraint on what the scoring rule *S* should look like. (In particular, it should satisfy a principle he calls “Propriety”.<sup>8</sup>)

However, this argument relies on the same point about the order of application of rules for ruling out credence functions. He defines modesty in terms of the existence of some alternative set of credences that would have lower expected inaccuracy. But the mere formal existence of such an alternative is no problem, if that alternative has already been ruled out for some other reason. Compare this pragmatic parallel of Joyce’s principle:

If, relative to a person’s own credences, some alternative [action] has a [higher expected utility], then, by her own estimation, that [action] is preferable from the [pragmatic] perspective. A [utility function *U*] should not render any [action] modest when there are [pragmatic] circumstances under which [that action is] clearly the rational one to [perform].

This initially sounds plausible. But when we consider the actions we perform in everyday life, they will clearly all be ‘modest’ in this sense. There is always some formally defined alternative that would be better – rather than betting a dollar at even odds on the outcome of a coin flip, I should choose the action that pays me a million dollars regardless of how the coin comes up! But this is no criticism of my action, or my utility function, since the alternative that is better is one that is not *available* to me. If we applied the modesty constraint to utility functions, we would have to make some *ad hoc* move to say that my utility function should always assign these unavailable actions lower expected utility than the best available action. The right thing to do here is to define ‘modest’ actions as ones such that some alternative *available* action has a higher expected utility.

If we apply this thought back to the epistemic case, then we see that we can only argue for Immodesty where ‘modest’ is understood relative to the *available* credence functions. If violations of the Principal Principle are taken to render a credence function *unavailable*, then Immodesty no longer supports the constraint Joyce argues for.

This requires some consideration of what the role of the Principal Principle is in constraining belief. If it helps determine which credence functions are available, then as we have seen, both of Joyce’s arguments run into problems. One might instead think that violation of the Principal Principle doesn’t make a credence

<sup>8</sup> Alan Hájek objects to Joyce’s argument here by arguing that there are in fact coherent credence functions that couldn’t be the objective chances (Hájek 2008, 814–816). We argue here that *even if* every coherent credence function *could be* the objective chance function, (PP) may interact differently with Joyce’s other rules than he suggests.



function unavailable, but instead just represents some dimension of epistemic 'badness'. If this badness is different from the badness of inaccuracy, then it becomes clear that Joyce's arguments need to be modified – even if  $b'$  dominates  $b$  with respect to inaccuracy, if  $b$  has less *overall* epistemic badness, then  $b$  may still be perfectly acceptable as a credence function. Thus, Joyce's arguments would need to consider overall badness rather than just inaccuracy.

The only way to save Joyce's arguments here seems to be to say that somehow the badness of violating the Principal Principle is *already included* when one has evaluated the accuracy of a credence function. Perhaps there is some way to argue for this claim.<sup>9</sup> But this claim needs more support than it has been given. And nothing here turns on the use of the Principal Principle in particular – if there can be any epistemic norm whose force is separate from accuracy, then the same sort of problem will arise. Joyce's argument works only if *all* epistemic norms spring from accuracy.

#### REFERENCES

- DE FINETTI, B. 1974, *The Theory of Probability*, New York, Wiley.
- HÁJEK, A. 2008, "Arguments for – or Against – Probabilism?", *The British Journal for the Philosophy of Science*, **59** (4), pp. 793–819.
- JOYCE, J. 1998, "A Nonpragmatic Vindication of Probabilism", *Philosophy of Science*, **65** (4), pp. 575–603.
- JOYCE, J. 2009, "Accuracy and Coherence: Prospects for an Alethic Epistemology of Partial Belief", in: F. Huber and C. Schmidt-Petri, eds, *Degrees of Belief*, Synthese Library, vol. 342, part II, Springer, pp. 263–297.
- KOLODNY, N. 2007, "How Does Coherence Matter?", *Proceedings of the Aristotelian Society*, **107** (3), pp. 229–263.
- LEWIS, D. 1980, "A Subjectivist's Guide to Objective Chance", in: R.C. Jeffrey, ed., *Studies in Inductive Logic and Probability*, vol. II., Berkeley, University of California Press.

<sup>9</sup> Perhaps the arguments in Joyce (1998) should be taken to show that *all* dimensions of epistemic badness are already included in the measure of inaccuracy. His defense of "Extensionality" on Ib. (591–592) already suggests that he expects some additional epistemic constraints to arise out of accuracy once we consider realistic models rather than toy models.