

Higher-order Probabilism

May 15, 2024

Abstract. Rational agents are often uncertain about the truth of many propositions. To represent this uncertainty, two options are typically on the table, precise and imprecise probabilism, but both fall short in some respect. Precise probabilism is not expressive enough, while imprecise probabilism suffers from belief inertia and the impossibility of proper scoring rules. We put forward a novel version of probabilism, higher-order probabilism, and we argue that it outperforms existing alternatives.

Keywords: Probabilism; Imprecise probabilities; Evidence; Probability; Belief inertia; Bayesian networks; Proper scores.

1 Introduction

As rational agents, we are uncertain about the truth of many propositions since our evidence is often fallible. To represent this uncertainty, two options are typically on the table: precise and imprecise probabilism. Precise probabilism models an agent's state of uncertainty (or credal state) with a single probability measure: each proposition is assigned a probability between 0 and 1 (a sharp credence). The problem is that a single probability measure is not expressive enough to distinguish between intuitively different states of uncertainty (§2). Alternatively, a *set* of probability measures can model an agent's uncertainty. This approach, known as imprecise probabilism, does better than precise probabilism in some respects but also runs into problems such as belief inertia and the impossibility of defining proper scoring rules (§3).

To make progress, this paper argues that a rational agent's uncertainty should be represented by a set of (higher-order) probability densities, one for each (first-order) proposition of interest. The theory we propose is not mathematically novel, nor are philosophers unfamiliar with higher-order probabilities **CITE SOME: Skyrms, Gaifmann, Domotor, Uchii, Pearl, Peijnenburg and Atkinson**. Here we address specifically how our approach to higher-order probabilism can overcome the problems of precise and imprecise probabilism (§4 and §??). We also argue that higher-order probabilism fares better when assessing the probability of multiple propositions, dependent or independent (§?? and ??). Many of the examples in this paper are about coin tosses, but in the final two sections, we discuss legal examples and illustrate the broader applicability of higher-order probabilism.

2 Precise probabilism

Precise probabilism holds that a rational agent's uncertainty about a proposition is to be represented as a single, precise probability measure. Bayesian updating regulates how the prior probability measure should change in light of new evidence that the agent learns. This is an elegant and simple theory, but representing our uncertainty about a proposition with a single, precise probability measure fails to capture an important dimension of how our fallible beliefs reflect the evidence we have (or have not) obtained. A couple of stylized examples featuring coin tosses should make the point clear.

No evidence v. fair coin You are about to toss a coin but have no evidence about its bias. You are completely ignorant. Compare this to the situation in which you know, based on overwhelming evidence, that the coin is fair.

On precise probabilism, both scenarios are represented by assigning a probability of .5 to the proposition *the coin will land heads on the next toss* (or *heads* for short). If you are completely ignorant, the principle of insufficient evidence suggests that you assign .5 to this proposition. Similarly, if you are sure the coin is fair, assigning again .5 seems the best way to quantify your uncertainty about the outcome. The agent's evidence in the two scenarios is quite different, but the precise probability of *heads* cannot capture this difference.

Learning from ignorance You toss a coin with unknown bias. You toss it 10 times and observe 5 heads. You toss it further and observe 50 heads in 100 tosses.

Since the coin initially had an unknown bias, you should presumably assign a probability of .5 to the proposition *heads*, if you stick with precise probabilism. After the 10 tosses, you again assess the probability to be .5. You must have learned something, but whatever that is, it is not reflected in the precise probability of *heads*. When you toss the coin 100 times and observe 50 heads, you also learn something new, but your precise probability does not change.¹

These examples show that the precise probability of *heads* is not appropriately responsive to evidence. But instead of focusing on this probability, the precise probabilist might respond that we should extend the algebra and include propositions about the bias of the coin. As evidence accumulates that the coin is fair, the (precise) probability of the proposition *the coin has a .5 bias* should go up, even though the (precise) probability of *heads* should remain .5. We think this response is on the right track, but how does it generalize beyond cases of coin tosses?

Suppose that, given a certain stock of evidence, *A* is more likely than *B*. Further, suppose that the acquisition of new evidence does not change the probabilities. Admittedly, something has changed in the agent's state of uncertainty: the quantity of evidence on which the agent can assess whether *A* is more probable than *B* has become larger. And yet, this change is

¹Another problem for precise probabilism is known as *sweetening* (Hare, 2010). Imagine a rational agent who does not know the bias of the coin. For precise probabilism, this state of uncertainty is represented by a .5 probability assignment to heads. Next, the agent learns that the bias towards heads, whatever the bias is, has been slightly increased, say by .001. Intuitively, the new information should leave the agent equally undecided about betting on heads or tails. After sweetening, the agent still does not know much about the actual bias of the coin. But, according to precise probabilism, sweetening should make the agent bet on heads: if the probability of heads was initially .5, it should now be slightly above .5.

not reflected in the precise probabilities assigned to A and B .² The precise probabilist might recommend we extend the algebra and include propositions of the form *the probability of event X is such and such*. Since such propositions contain a probability, assigning a probability to them effectively amounts to using higher-order probabilities.³ Our approach to higher-order probabilism will not be quite that, for reasons we shall explain. Before going higher-order, however, we should explore another view in the literature.

3 Imprecise probabilism

Imprecise probabilism holds that a rational agent's credal stance towards a hypothesis is to be represented by a set of probability measures, typically called a representor \mathbb{P} , rather than a single measure P . The representor should include all and only those probability measures which are compatible with the evidence (more on this point later).⁴ Modeling an agent's credal state by sets of probability measures can easily accommodate the scenarios in the previous section without adding propositions about coin biases. For instance, if an agent is sure the coin is fair, their credal state would be represented by the singleton set $\{P\}$, where P is a probability measure that assigns .5 to *heads*. If, on the other hand, the agent knows nothing about the coin's bias, their credal state would be represented by the set of all probabilistic measures, since none of them is excluded by the available evidence.

So far so good. But now consider this scenario:

Even v. uneven bias: You have two coins and you are sure the probability of *heads* is .4, if you toss one coin, and .6, if you toss the other. But you cannot tell which is which. You pick one coin at random and toss it. Contrast this with an uneven case. You have four coins and you are sure three of them have bias .4 and one of them bias .6. You pick a coin at random and toss it. You should be three times more confident that the probability of getting heads is .4. rather than .6.

The first scenario can be easily modeled by imprecise probabilism. The representor would contain two probability measures; one assigns .4. and the other assigns .6 to the proposition *the coin will land heads*. Yet imprecise probabilism cannot model the second scenario. Since the probability measures in the set are all compatible with the agent's evidence, no probability measure can be assigned a greater (higher-order) probability than any other.⁵

These scenarios show that imprecise probabilism is not always expressive enough. Worst

²The distinction here is sometimes formulated in terms of the *balance* of the evidence (that is, whether the evidence available tips in favor a hypothesis or another) as opposed to its *weight* (that is, the overall quantity of evidence regardless of its balance); see Keynes (1921) and Joyce (2005).

³Interestingly, **PERAL ADD REFERENCE** argues that the semantics of higher-order probability statements can still be expressed in the language of first-order probability.

⁴For the development of imprecise probabilism, see Keynes (1921); Levi (1974); Gärdenfors & Sahlin (1982); Kaplan (1968); Joyce (2005); Fraassen (2006); Sturgeon (2008); Walley (1991). Bradley (2019) is a good source of further references. Imprecise probabilism is closely related to what we might call interval probabilism (Kyburg, 1961; Kyburg Jr & Teng, 2001). In interval probabilism, precise probabilities are replaced by intervals of probabilities. Imprecise probabilism is more general since the representor set need not be an interval.

⁵Other scenarios can be constructed in which imprecise probabilism fails to capture distinctive intuitions about evidence and uncertainty; see, for example, (Rinard, 2013).

still, imprecise probabilism suffers from three shortcomings that do not affect precise probabilism: first, the idea of compatibility between the available evidence and the probability measures in the representator set is not clearly defined; second, updating imprecise probabilities can run into belief inertia; and third, no proper scoring rules exist for imprecise probabilities.

The first shortcoming has not received extensive discussion in the literature. For imprecise probabilism, an agent's state of uncertainty is represented by those probability measures that are *compatible* with the agent's evidence, but the notion of compatibility lacks a clear definition in the literature. Roughly, think of it as the fact that the agent's evidence does not outright rule out the probability measure in question. The problem is that observations, evidence and data will hardly rule out any probability measure. Admittedly, there will be clear-cut cases: if you see the outcome of a coin toss is heads, you reject the measure with $P(\text{heads}) = 0$, and similarly for tails. Another class of cases might arise while randomly drawing objects from a finite set where the objective chances are known.⁶ But clear-cut cases aside, what else?

A second, related problem for imprecise probabilism is belief inertia (Levi, 1980). Precise probabilism offers an elegant model of learning from evidence: Bayesian updating. Imprecise probabilism, at least *prima facie*, offers an equally elegant model of learning from evidence, richer and more nuanced. When faced with new evidence E between time t_0 and t_1 , the representor set should be updated point-wise, running the standard Bayesian updating on each probability measure in the representor:

$$\mathbb{P}_{t_1} = \{P_{t_1} | \exists P_{t_0} \in \mathbb{P}_{t_0} \forall H [P_{t_1}(H) = P_{t_0}(H|E)]\}.$$

Belief inertia arises in situations in which no amount of evidence can lead the agent to change their initial belief state, according to a given modeling strategy. Consider a situation in which you start tossing a coin knowing nothing about its bias, so the range of possibilities is $[0, 1]$. After a few tosses, if you observed at least one tail and one head, you can exclude the measures assigning 0 or 1 to *heads*. But else what can you learn? For any sequence of outcomes that you can obtain and any probability value in $(0, 1)$, there will exist a probability measure (conditional on the outcomes) that assigns that probability to *heads*. Consequently, the edges of your resulting interval will remain the same.

Some downplay the problem of belief inertia and blame the uniform prior $[0, 1]$.⁷ Just as contingent propositions should not be assigned (precise) probabilities of 0 or 1 since these extreme values are unrevisable, by the same token the uniform prior $[0, 1]$ should not be used whenever it is impervious to revision (Moss, 2020). More generally, it is unsurprising you will not learn anything when you start from a place of ignorance (Joyce, 2010). But, as we will see in the next section, uniform priors are not necessarily unrevisable and can be a starting point

⁶Probability measures can be inconsistent with evidential constraints that agents believe to be true (Bradley, 2012), for example, *structural constraints* such as “ X and Y are independent”. But, unless they come from an oracle, there will usually be some degree of uncertainty about the acceptability of these constraints.

⁷Another strategy is to say that, in a state of ignorance, a special updating rule should be deployed. Elkin (2017) suggests the rule of *credal set replacement* that recommends that upon receiving evidence the agent should drop measures rendered implausible, and add all non-extreme plausible probability measures. This, however, is tricky. One needs a separate account of what makes a distribution plausible from a principled account of why one should use a separate special update rule when starting with complete ignorance.

for learning. The problem lies with imprecise probabilities, not with uniform priors as such.

The third problem for imprecise probabilism is to find suitable proper scoring rules. In the precise case, scoring rules measure the distance between a rational agent's probability measure and the actual value. The Brier score is the most common scoring rule for precise probabilities.⁸ A requirement of scoring rules is *propriety*: any rational agent will expect their probability measure to be more accurate than any other. A bit more formally, the expected inaccuracy of p from the perspective of p should always be smaller than the expected inaccuracy of p from the perspective of another distribution q . After all, if an agent thought a different probability measure was more accurate, they should switch to it. Well-known results demonstrate the strict propriety of the Brier score for precise probabilities.⁹ Can similar results be established for imprecise probabilities? The answer is likely to be negative.

In the precise case, let $I(p, w)$ be an inaccuracy score of a probability distribution p relative to the true state w . Its expected inaccuracy, from the perspective of p equals

$$\sum_{w \in W} p(w) I(p, w).$$

In the imprecise case, let $I([p_-, p_+], w)$ be the inaccuracy score for the interval $[p_-, p_+]$. What is its expected inaccuracy from the perspective of, say, the interval $[p_-, p_+]$ itself? There is no straightforward answer to this question because $I([p_-, p_+], w)$ cannot be multiplied by $[p_-, p_+]$, in the way in which $I(p, w)$ can be multiplied by $p(w)$. The expected inaccuracy of $[p_-, p_+]$ can be evaluated from the perspective of the precise probabilities at its edges.¹⁰ The problem is that, in this case, finding a proper inaccuracy score that is also continuous is mathematically impossible (Seidenfeld, Schervish, & Kadane, 2012).

Proper scoring rules are often used to formulate accuracy-based arguments for precise probabilism. These arguments show (roughly) that, if your precise measure follows the axioms of probability theory, no other non-probabilistic measure is going to be more accurate than yours whatever the facts are. But, without proper scoring rules for imprecise probabilities, the prospects for an accuracy-based argument for imprecise probabilism look dim (Campbell-Moore, 2020; Mayo-Wilson & Wheeler, 2016).¹¹

⁸The Brier score is defined as the squared distance between the true state and the probability forecast, or formally, $(p(x) - V(x, w))^2$, where $p(x)$ is the probability forecast and $V(x, w)$ determines if a proposition obtains at w ($V(x, w) = 1$) or not ($V(x, w) = 0$). If, for example, the proposition 'rain' obtains at w , the forecast 'rain with .9 probability' would be more accurate at w than the forecast 'rain with .8 probability'. If, on the other hand, the proposition 'not rain' obtained at w , the latter forecast would be more accurate.

⁹Besides propriety, other common requirements (which the Brier score also satisfies) are: the score $I(p, w)$ should be a function of the probability distribution p and the true state w (extensionality); and the score should be a continuous function around p (continuity).

¹⁰So the expected inaccuracy would equal $\sum_{w \in W} p_-(w) I([p_-, p_+], w)$ or $\sum_{w \in W} p_+(w) I([p_-, p_+], w)$.

¹¹Moreover, as shown by Schoenfield (2017), if an accuracy measure satisfies certain plausible formal constraints, it will never strictly recommend an imprecise stance, as for any imprecise stance there will be a precise one with at least the same accuracy.

4 Higher-order probabilism

Let us take stock. Imprecise probabilism is expressive enough to model the difference between a state in which there is no evidence about a proposition (or its negation) and a state in which the evidence for and against a proposition is in equipoise. However, imprecise probabilism has its own expressive limitations, for example, it cannot model the case of uneven bias. In addition, imprecise probabilism faces difficulties that do not affect precise probabilism: the notion of compatibility between a probability measure and the evidence is too permissive; belief inertia makes it impossible for a rational agent to learn via Bayesian updating; and no proper scoring rules exist for imprecise probabilism. In this section and the next, we show that higher-order probabilism overcomes the expressive limitations of imprecise probabilism without falling prey to any such difficulties.

Proponents of imprecise probabilism already hinted at the need to rely on higher-order probabilities. For instance, Bradley (2012) compares the measures in a representor to committee members, each voting on a particular issue, say the true chance or bias of a coin. As they acquire more evidence, the committee members will often converge on a chance hypothesis. But such convergence up cannot be modeled by imprecise probabilism alone: a probability distribution over chance hypotheses is needed.¹² That one should use higher-order probabilities has also been suggested by critics of imprecise probabilism. Carr (2020) argues that sometimes evidence requires uncertainty about what credences to have. Carr, however, does not develop this suggestion formally, and does not explain how her approach would fare against the difficulties affecting imprecise probabilism. We now set out to do precisely that.

The central idea of our approach to higher-order probabilism is this: instead of a probability between 0 and 1, rational agents assign a probability density to each proposition of interest. For example, consider the scenario in which the agent is about to toss a coin whose known bias is either .4 or .6, but the former is three times more likely. This scenario—which imprecise probabilism cannot model—can be easily modeled within high-order probabilism. The proposition *the coin will land heads on the next toss* is assigned a probability density whose mass is 3/4 on the .4 value and the remaining 1/4 on the .6 value, as illustrated in Figure 1. To be sure, an agent’s uncertainty about a proposition could—perhaps, should—sometimes be represented by a single probability value. Higher-order probabilism does not prohibit that. There may well be cases in which an agent’s uncertainty is aptly represented by the expectation, usually defined as $\int_0^1 pf(p) dp$, where p is the first-order probability of a given proposition and f is the density representing the agent’s uncertainty about p . But this need not always be the case. If the probability distribution is not sufficiently concentrated around a single value, a one-point summary will fail to do justice to the nuances of the agent’s credal state.¹³

So, an agent’s state of uncertainty (or credal stance) towards a proposition x is not represented by a probability value $p(x)$ between 0 and 1, but by a probability density $f_x(p(x))$,

¹²In a similar vein, Joyce (2005), in a paper defending imprecise probabilism, explicates the notion of weight of evidence using a probability distribution over chance hypotheses.

¹³In the case of uneven biases, for example, representing the agent’s credal state with the expectation $.75 \times .4 + .25 \times .6 = .45$ would fail to capture the agent’s different epistemic attitudes towards the two biases. The agent believes the two biases have different probabilities, and is also certain the bias is *not* .45.

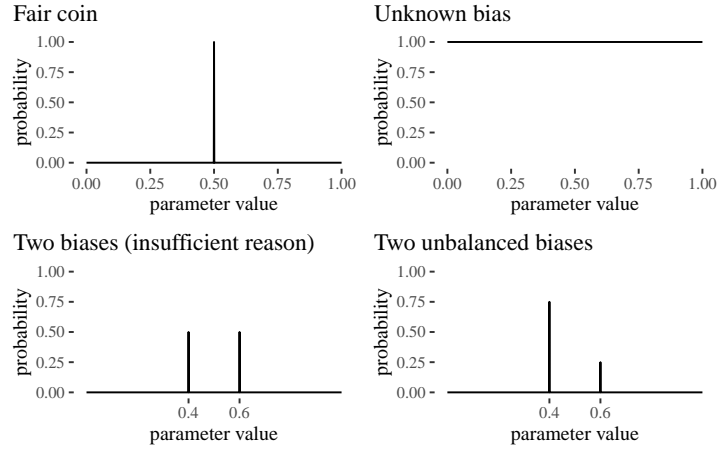


Figure 1: Examples of higher-order distributions for a few scenarios problematic for both precise and imprecise probabilism.

where the first-order probability $p(\mathbf{x})$ is the parameter in question and is treated as a random variable. (For computational ease, we will use a higher-order discretized density ranging over 1000 first-order probabilities.) More generally, let \mathbf{X} be the set of first-order propositions an agent is interested in (more on this restriction soon). For each proposition $\mathbf{x} \in \mathbf{X}$, the rational agent will assign a probability density $f_{\mathbf{x}}(p(\mathbf{x}))$, and not simply a probability value $p(\mathbf{x})$. The mapping from propositions to densities models the agent’s overall state of uncertainty.¹⁴

Besides its greater expressive power in modeling uncertainty, higher-order probabilism does not fall prey to the technical difficulties of imprecise probabilism. It does not fall prey to belief inertia. Consider a situation in which you have no idea about the bias of a coin. You start with a uniform distribution over $[0, 1]$ as your prior. Observing any non-zero number of heads will exclude 0 and observing any non-zero number of tails will exclude 1 from the basis of the posterior. The posterior distribution will become more centered as the observations come in. This result is a straightforward application of Bayesian updating. Instead of plugging sharp probability values into the formula for Bayes’s theorem, the factors to be multiplied in the theorem will be probability densities (or ratios of densities as needed). Figure 2 illustrates—starting with a uniform prior distribution—how the posterior (beta) distribution changes after successive observations.¹⁵

The impossibility of defining proper scoring rules was another weakness of imprecise proba-

¹⁴This approach lines up with common practice in Bayesian statistics, where the primary role of uncertainty representation is assigned to the whole distribution. Summaries such as the mean, mode standard deviation, mean absolute deviation or highest posterior density intervals are only succinct ways of representing uncertainty.

¹⁵Assuming independence and constant probability for all the observations, learning is modeled the Bayesian way. You start with some prior density p over the parameter values. If you start with a complete lack of information, p should be uniform. Then, you observe the data D which is the number of successes s in a certain number of observations n . For each particular possible value θ of the parameter, the probability of D conditional on θ follows the binomial distribution. The probability of D is obtained by integration. That is:

$$\begin{aligned} p(\theta|D) &= \frac{p(D|\theta)p(\theta)}{p(D)} \\ &= \frac{\theta^s(1-\theta)^{(n-s)}p(\theta)}{\int (\theta')^s(1-\theta')^{(n-s)}p(\theta') d\theta'}. \end{aligned}$$

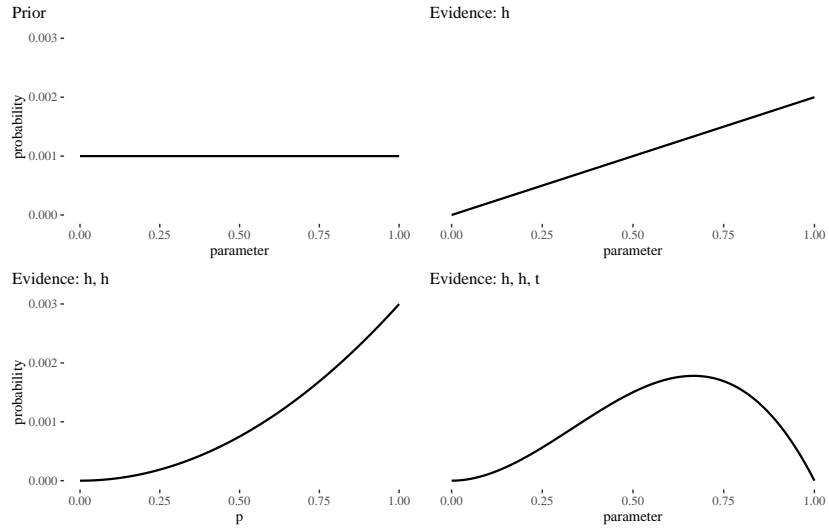


Figure 2: As observations of heads, heads and tails come in, extreme parameter values drop out of the picture and the posterior is shaped by the evidence.

bilism. This is a significant shortcoming, especially because proper scores do exist for precise probabilism. Fortunately, one can show that there exist proper scoring rules for higher-order probabilism. We defend this claim in greater detail in the next section. But before turning to this topic, we will discuss an alternative approach to higher-order probabilism.

As noted earlier, the set \mathbf{X} of propositions of interest is restricted to first-order ones. What does that mean? We are here relying on the intuitive distinction between propositions more directly about events in the world, such as *the coin will land heads on the next toss* and more abstract constructs such as the probabilities of these propositions. Only the former are included in the set \mathbf{X} . Now, the precise probabilist might agree with this distinction but insist that going higher-order in the way we are suggesting is unnecessary. After all, the set of propositions of interest need not be limited to first-order ones. It can be extended to include second-order propositions, those about the probabilities of first-order propositions. Then, an agent's state of uncertainty can be modeled by a single, precise probability measure defined over all members of this extended set denoted by \mathbf{X}^2 . The scenario of uneven coins would not pose a problem, since the (second-order) propositions about the coin biases can be assigned different probabilities, that is, $p(\text{bias} = .4) = 3/4$ and $p(\text{bias} = .6) = 1/4$. Call this approach \mathbf{X}^2 -precise probabilism to distinguish it from the simpler version of precise probabilism.

Higher-order probabilism and \mathbf{X}^2 -precise probabilism are quite similar, but a few differences are worth pointing out. Instead of representing one's uncertainty about the coin bias by, say, a Beta distribution, the agent is expected to have attitudes towards uncountably many propositions of the form *bias* = x . This latter route might be viable, but it is in no way simpler. For consider the (first-order) proposition *the coin will land heads on the next toss* (in short, *heads*) and the (second-order) *the coin has a bias of .5*. We expect them to be related in precise ways, for example, it is inconsistent to assign a probability of one (or an extremely high or low probability) to both of them. Evidently, the two propositions are not probabilistically independent, but their dependence is conceptual, not empirical: that a coin is fair just means that

its probability of landing heads cannot be extremely high. Higher-order probabilism can easily model this conceptual dependence by positing that the expectation of the (higher-order) density $f_{heads}(p(heads))$ should equal the (first-order) probability $p(heads)$. Instead, X^2 -precise probabilism will have to hard code (infinitely many!) dependencies between first- and second-order propositions as *ad hoc* conceptual postulates. Besides its greater simplicity, another advantage of our higher-order approach is that, at the technical level, it is by no means novel. Bayesian probabilistic programming languages embrace the well-known idea that parameters can be stacked and depend on each other (Bingham et al., 2021). But, while the technical machinery has been around for a while, it has not been deployed by philosophers to model a rational agent’s uncertainty or credal state.¹⁶

References

- Bingham, E., Koppel, J., Lew, A., Ness, R., Tavares, Z., Witty, S., & Zucker, J. (2021). Causal probabilistic programming without tears. *Proceedings of the Third Conference on Probabilistic Programming*.
- Bradley, S. (2012). *Scientific uncertainty and decision making* (PhD thesis). London School of Economics; Political Science (University of London).
- Bradley, S. (2019). Imprecise Probabilities. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2019). <https://plato.stanford.edu/archives/spr2019/entries/imprecise-probabilities/>; Metaphysics Research Lab, Stanford University.
- Campbell-Moore, C. (2020). *Accuracy and imprecise probabilities*.
- Carr, J. R. (2020). Imprecise evidence without imprecise credences. *Philosophical Studies*, 177(9), 2735–2758. <https://doi.org/10.1007/s11098-019-01336-7>
- Elkin, L. (2017). *Imprecise probability in epistemology* (PhD thesis). Ludwig-Maximilians-Universität; Ludwig-Maximilians-Universität München.
- Fraassen, B. C. V. (2006). Vague expectation value loss. *Philosophical Studies*, 127(3), 483–491. <https://doi.org/10.1007/s11098-004-7821-2>
- Gärdenfors, P., & Sahlin, N.-E. (1982). Unreliable probabilities, risk taking, and decision making. *Synthese*, 53(3), 361–386. <https://doi.org/10.1007/bf00486156>
- Hare, C. (2010). Take the sugar. *Analysis*, 70(2), 237–247.
- Joyce, J. (2005). How probabilities reflect evidence. *Philosophical Perspectives*, 19(1), 153–178.
- Joyce, J. (2010). A defense of imprecise credences in inference and decision Making1. *Philosophical Perspectives*, 24(1), 281–323. <https://doi.org/10.1111/j.1520-8583.2010.00194.x>
- Kaplan, J. (1968). Decision theory and the fact-finding process. *Stanford Law Review*, 20(6), 1065–1092.

¹⁶Our approach to higher-order probabilism should also be distinguished from what we might call higher-order (imprecise) probabilism. On this account, an agent’s state of uncertainty is modeled by a set of probability measures defined over a fixed algebra of first-order propositions (just like imprecise probabilism does), and in addition, a probability distribution over those first-order probability measures. This is a generalization of imprecise probabilism.

- Keynes, J. M. (1921). *A treatise on probability, 1921*. London: Macmillan.
- Kyburg, H. E. (1961). *Probability and the logic of rational belief*. Wesleyan University Press.
- Kyburg Jr, H. E., & Teng, C. M. (2001). *Uncertain inference*. Cambridge University Press.
- Levi, I. (1974). On indeterminate probabilities. *The Journal of Philosophy*, 71(13), 391. <https://doi.org/10.2307/2025161>
- Levi, I. (1980). *The enterprise of knowledge: An essay on knowledge, credal probability, and chance*. MIT Press.
- Mayo-Wilson, C., & Wheeler, G. (2016). Scoring imprecise credences: A mildly immodest proposal. *Philosophy and Phenomenological Research*, 92(1), 55–78. <https://doi.org/10.1111/phpr.12256>
- Moss, S. (2020). Global constraints on imprecise credences: Solving reflection violations, belief inertia, and other puzzles. *Philosophy and Phenomenological Research*, 103(3), 620–638. <https://doi.org/10.1111/phpr.12703>
- Rinard, S. (2013). Against radical credal imprecision. *Thought: A Journal of Philosophy*, 2(1), 157–165. <https://doi.org/10.1002/tht3.84>
- Schoenfield, M. (2017). The accuracy and rationality of imprecise credences. *Noûs*, 51(4), 667–685. <https://doi.org/10.1111/nous.12105>
- Seidenfeld, T., Schervish, M., & Kadane, J. (2012). Forecasting with imprecise probabilities. *International Journal of Approximate Reasoning*, 53, 1248–1261. <https://doi.org/10.1016/j.ijar.2012.06.018>
- Sturgeon, S. (2008). Reason and the grain of belief. *Noûs*, 42(1), 139–165. Retrieved from <http://www.jstor.org/stable/25177157>
- Walley, P. (1991). *Statistical reasoning with imprecise probabilities*. Chapman; Hall London.