Two norms of epistemic humility

Jakob Donskov, jakobdonskov@gmail.com

[Draft April 2026]

Abstract. Our fallibility calls for epistemic humility: the available evidence is often difficult to evaluate so we should be open to the possibility of error. Here I argue that Bayesian models of higher-order uncertainty can capture rational epistemic humility by imposing two norms on credences. *Total Calibration* requires lower-order credences to match higher-order expectations of the rational credence. The principle avoids epistemic akrasia without requiring higher-order certainty while explaining why the humble decrease their confidence. *Proportional Resilience* requires higher-order uncertainty to constrain how much to expect credences to change in the face of new evidence. The principle thereby links higher-order uncertainty to open-mindedness, explaining why the humble expect new evidence to be counterevidence.

Keywords: Epistemic humility · Higher-order uncertainty · Calibration · Akrasia · Doxastic resilience · Open-mindedness

1 Introduction

If you care about believing truths and avoiding falsehoods, figuring out what to believe is difficult. Your best bet is to follow the evidence: believe in accordance with what your evidence indicates. As Hume (1748) famously put it, the wise proportion their beliefs to the evidence. On the Bayesian conception of belief,

this amounts to giving propositions credence to the degree your total evidence indicates that those propositions are true.

Our evidence is often complex, ambiguous, and hard to evaluate, and there is ample reason to think that we are prone to bias and error in assessing it. As such, we have reason to be epistemically humble: we should be open to the possibility that our beliefs are not proportioned to what our evidence indicates. To account for such epistemic humility, Bayesians appeal to higher-order uncertainty.[1] Where our first-order credences reflect our uncertainty about how the world is, higher-order credences reflect our uncertainty about our own rationality.

The central claim of this paper is that higher-order uncertainty has two distinct normative consequences. First, higher-order uncertainty about your own rationality bears on your beliefs about how the world is. Second, higher-order uncertainty can require a distinctive form of open-mindedness, namely an increased expectation that further evidence will overturn one's current views. To capture this, I impose two norms on rational credence functions. The first governs how higher-order uncertainty constrains first-order credences. The second governs how such uncertainty bears on how firmly one should expect those beliefs to withstand future inquiry. Taken together, the two norms articulate what epistemic humility requires.

The first norm is a generalised reflection principle: *Total Calibration*. This requires first-order credences to match higher-order expectations of rational credence. I argue that it is irrational for our actual credences to diverge

[1] Skyrms (1980b) gives an early overview of the literature and defends the coherence of higher-order degrees of belief. See also Gaifman (1985) and more recently Egan and Elga (2005), Elga (2007, 2013), Briggs (2009), Christensen (2010a), and Dorst (2019) for influential papers.

from what we expect the rational credence to be. Absent such a constraint, higher-order uncertainty leads to epistemic akrasia: believing something while believing that one should not believe it.

Humility is not merely a matter of reduced confidence. It also involves non-dogmatism: a readiness to be corrected by further evidence. The second norm, *Proportional Resilience*, therefore links higher-order uncertainty and one's expectation that new evidence will overturn one's beliefs. This can be modelled directly in the agent's credence function. Since rational agents update their beliefs by conditionalising on what they learn, the Bayesian conception of belief encodes the dispositional profile of a rational agent's belief updating strategy directly into the doxastic states.

These dispositions can be characterised in terms of degrees of credal resilience. Highly resilient credences are expected to remain stable across a wide range of evidential updates, whereas non-resilient credences are expected to change. I argue that higher-order uncertainty should constrain not only an agent's credence in a proposition but also the expected resilience of that credence. When agents doubt that their beliefs are rational, they should expect new evidence to be counterevidence.

In §2, I clarify the connection between epistemic humility and higher-order uncertainty. In §3, I review extant calibration principles and defend Total Calibration, and in §4, I defend Proportional Resilience. In §5, I show how the resulting framework illuminates higher-order defeat and higher-order confirmation. In §6, I conclude.

2 Fallibility calls for humility

Evaluating one's evidence correctly is no trivial matter. We can get it right or wrong, we can get close to evaluating it correctly, or completely misunderstand what bearing our evidence has on what we should think about the world. Even in highly controlled environments such as scientific inquiry, it will often be unclear what the evidence rationalises. Indeed, what makes scientific inquiry difficult is not merely gathering the data. The hard part is figuring out what those data points in the spreadsheet tell us about the world.

For example, it has been observed that the vocal sexual displays or 'songs' by male humpback whales in the southwestern Pacific Ocean vary systematically between communities and that there are certain patterns in how the songs evolve across communities (Garland et al. 2011). The working hypothesis is that whale pods located off the east coast of Australia serve as cultural centres for the development of new songs. New songs originate in the coastal communities and then spread east from pod to pod for thousands of nautical miles. Consider now the following scenario:

> **The marine biologist:** Ellen is observing humpback whales during the feeding season in the Antarctic. Using standard acoustic tools and transcription methods, she records the song of a particular whale and compares its spectrogram to an existing database of regional song types. The song closely matches the dominant song currently observed in the French Polynesian population and matches less well with songs from neighbouring regions. Taking into account the degree of similarity, known rates of song change, and overlap between regional song types, Ellen judges it substantially more likely than not that the whale belongs to the French Polynesian population.

Two features of this case are worth separating. First, doxastic attitudes admit of degree. Ellen is not certain that the whale belongs to the French Polynesian population, but she judges it to be more likely than not. On the Bayesian interpretation, this uncertainty is represented by a credence between 0 and 1. Suppose that Ellen's credence in the French Polynesia hypothesis is 0.75. This first-order credence reflects her uncertainty about how the world is. Second, evidential support also admits of degree. Suppose, for example, that Ellen's evidence favours what we might call the French Polynesia hypothesis, but that it does not conclusively rule out that the whale is an outlier. It is possible, therefore, that Ellen's 0.75 credence is exactly the rational response to the evidence she has.

Unfortunately, it need not be transparent to Ellen exactly how much the evidence supports the hypothesis. Suppose that Ellen is uncertain whether she has assessed her evidence correctly:

> **The humble marine biologist:** Ellen recognises that her confidence in the French Polynesia hypothesis rests on a complex chain of reasoning. It depends on contested choices about how songs are segmented and weighted, on assumptions about migration and cultural transmission, and on the representativeness of the database. Since she is aware that these matters are complicated, and small methodological changes can shift the resulting probabilities, Ellen is uncertain whether her judgment is itself rational.

Ellen's uncertainty is rational. She should harbour at least some doubt as to whether 0.75 is a rational response to her evidence. Notice that this is plausible even if we assume that her 0.75 credence is in fact rational given her first-order evidence. Generally, we have reason to doubt the rationality of our

credences in proportion to how difficult the total evidence is to evaluate. When our beliefs hinge on many different pieces of evidence, when the evidence is ambiguous, or difficult to interpret due to cognitive limitations or potential bias, we have reason to doubt the rationality of our beliefs.

This is easy to see on objectivist theories of evidential support. If there are facts about how much a batch of evidence supports a given proposition, we will often have reason to doubt whether we have formed our credences in proportion to these facts. For example, perhaps rational belief is constrained by evidential probabilities.[2] Or, perhaps rational belief in a hypothesis is constrained by how well that hypothesis explains the evidence.[3] In either case, we should often be higher-order uncertain whether our credences rationally reflect our evidence.

But even Bayesian subjectivists, who reject that there are objective support relations between propositions, may need higher-order uncertainty on other grounds. Agents may, for example, be uncertain what they actually believe (Skyrms 1980b, 1984) or what they will believe (van Fraassen 1984). Worse yet, agents will often have reason to doubt whether their beliefs are coherent and whether they have updated their beliefs according to the correct updating rule. Subjectivists may well judge that Ellen should harbour some doubt as to whether her 0.75 credence is rational, even if this doubt is not grounded in facts about evidential support.

Objectivists and subjectivists alike, therefore, should say that we often have reason to doubt that our evidential evaluations are correct. This is the sense in which fallibility calls for humility. We are epistemically limited creatures, and rational belief formation should reflect that fact. We should take

[2] See Williamson (2000), Hawthorne (2005), and Maher (2006) for defences.

[3] See Lipton (2004), Weisberg (2009), Henderson (2014), and Hedden (2015) for defences.

seriously the possibility that our current beliefs are epistemically suboptimal. What does this amount to? In the rest of the paper, I argue that we can make sense of epistemic humility in terms of higher-order uncertainty. But, as I argue in the following, we need to steer clear of two pitfalls when introducing higher-order uncertainty.

2.1 Akrasia and dogmatism

There is something dogmatic about someone who fails to recognise their fallibility. An adequate account of epistemic humility should explain what is wrong with such dogmatism, and it should do so in a way that connects humility to rational reasoning and belief revision. It will not be enough to appeal to the intrinsic vice of dogmatism.

Several prominent Bayesians explicitly link higher-order uncertainty to epistemic humility (Christensen 2010a, 2020; Elga 2013; Dorst 2019; Hedden and Dorst 2022).[4] The idea is that evidence of irrationality should induce doubt in one's rationality. The more reason you have to be higher-order uncertain, the more reason you have to doubt the rationality of your beliefs. The slogan is that epistemically humble people sometimes doubt whether their credences are rational.

The literature on peer disagreement, which led to a more general literature on higher-order evidence, has emphasised this point. Learning that a peer disagrees with you, that a prankster has drugged your morning coffee, or that

[4] Outside the Bayesian literature, appeal to higher-order beliefs has been used to state requirements on justified belief. For example, the principle that justified belief requires justified higher-order belief that one's first-order beliefs are reliably formed. For an overview, see Graf (2025).

you currently suffer from hypoxia, all give rise to higher-order evidence.[5] While such evidence bears not on the truth of the content of your first-order beliefs, but rather on the rational standing of the beliefs themselves, it can nonetheless still influence what you should believe.

Indeed, it seems plausible that our credences about what is rational should somehow influence what we should believe. Without a way of linking higher-order uncertainty to first-order credences we face a dilemma. Either we permit agents to hold beliefs they themselves believe to be irrational—thereby permitting epistemic akrasia—or else we require agents to be higher-order certain—thereby implausibly requiring agents to be dogmatic.

To see why unconstrained higher-order uncertainty leads to akrasia, consider Hazlett's (2012) theory of epistemic humility. He argues that epistemic humility should be understood as a disposition to form appropriate attitudes toward the rationality of one's own beliefs. On Hazlett's view, what matters for humility is not what you believe about the world, but what you believe about your own rationality. Cashed out in a categorical framework, Hazlett requires that the humble suspends judgment on whether their belief that p is rational. Or, if the higher-order evidence against the rationality of belief that p is sufficiently strong, they ought to outright disbelieve that their belief is rational.

On Hazlett's account, this is all higher-order evidence does. What matters is whether one appropriately doubts the rationality of one's beliefs, not whether such doubt has any downstream effect on one's first-order reasoning. Indeed, Hazlett uses this to argue for a steadfast position on peer disagreement: the rational first-order credence remains unchanged in the face of peer disagreement.

[5] For a spectrum of views on how to deal with disagreement, see Christensen (2007), Kelly (2010), and Tal (2020).

On this way of thinking, epistemic humility is exhausted by higher-order attitudes. However, understood in this way, Hazlett's account threatens to render higher-order uncertainty epiphenomenal. If higher-order attitudes toward one's own rationality place no constraints on what one should believe, then they make no difference to deliberation or decision-making.

Suppose that the humble correctly assess the rationality of their beliefs, but are not required to change those attitudes which they recognise as irrational. Then there will be agents who rationally judge themselves to be irrational while rationally continuing to hold their beliefs. Such a set of attitudes seems to be incoherent. It would be better if you were to resolve the tension by changing some of your beliefs, and, importantly, you yourself will also judge this to be the case. That is just what it means to believe that one's own beliefs are irrational. On a categorical conception of belief, the incoherence is difficult to deny: you believe that p while you disbelieve that, or suspend on whether, it is rational for you to believe that p.

The upshot is that Hazlett's view permits *epistemic akrasia*: believing something you believe you should not believe. As several writers have noted, such akrasia seems paradigmatically irrational (Horowitz 2014; Greco 2014; Smithies 2019).[6] It seems irrational not to heed the advice of what you yourself take your evidence to suggest. What characterises the akratic is not that they hold suboptimal beliefs; anyone with a false belief does that. Rather, the akratic hold beliefs which are, by their own lights, suboptimal.

Akrasia is standardly defined in terms of categorical belief, but it also

[6] While some authors have argued that epistemic akrasia can be rationally permissible, their arguments typically appeal to the implausibility of putative bridge principles between lower-order and higher-order attitudes (Williamson 2014; Lasonen-Aarnio 2015, 2020). However, I will argue that there *is* a plausible constraint, so I will assume that akrasia is irrational.

seems to apply to certain combinations of first-order and higher-order credences. I will make the definition of credal akrasia precise below. For now, call an agent akratic if, by the agent's own lights, there is some alternative credence she expects to be epistemically superior. Akrasia, so understood, involves a mismatch between one's first-order credences and one's higher-order evaluation of those credences.

Clearly, something has gone wrong if the marine biologist with her 0.75 credence is also higher-order *certain* that her 0.75 credence is irrational (Skipper 2021; Tal 2020). An ideal version of Ellen would somehow have managed to revise her beliefs to resolve the tension.[7] If she is certain her credence in the French Polynesia hypothesis is irrational, it seems she should somehow change her credence to something she is more confident is rational (White 2009; Sliwa and Horowitz 2015).

Less than higher-order certainty that you are irrational will also lead to incoherence if the uncertainty plays no role in what first-order credences you should have. Suppose instead the biologist is uncertain that 0.75 is the rational credence to have, and so gives at least some credence to some other credence, say 0.6, being rational. It seems that her uncertainty whether 0.6 is the rational response should play some role in her deliberation and decision-making. Suppose she is 0.9 confident that 0.75 is rational and 0.1 confident that 0.6 is rational. If she disregards her higher-order uncertainty and solely relies on her 0.75 credence, it seems that she is, by her own lights, simply disregarding

[7] There may be some fringe cases where this is arguably not so. Christensen (2024) argues that akrasia may be permissible in cases where the evidence suggests a divergence between rationality and accuracy. Such cases will not be typical, however, so we will still need a bridge principle which explains why akrasia is typically irrational. But if Christensen is right, we can tacitly add the proviso to any bridge principle that it applies only if the total evidence does not suggest that rational beliefs are inaccurate.

part of her evidence. After all, whatever made her give 0.1 credence to 0.6 being rational plays no part in guiding her deliberations.

One immediate way to rule akrasia irrational is to deny that higher-order uncertainty is ever rational (Titelbaum 2015; Smithies 2019; Skipper 2021). If agents are required to be certain of their own rationality, akrasia can never be rational.

The challenge to such views is obvious: it often seems unduly dogmatic to have such extreme confidence that one has evaluated the evidence correctly. Higher-order certainty fixes your view about whether you are rational. If you are higher-order certain that your current credence function is the rational one, you thereby close off the possibility that further reflection on your epistemic performance could reveal a mistake. Once higher-order certain, you eliminate the possibility that you have misevaluated the evidence by overlooking relevant considerations, misjudged the weight of some piece of evidence, and so on. More generally, you should reject that the common experience of bias or cognitive limitations applies to yourself.

This will typically not be appropriate. For example, what would justify the biologist to be *certain* that her 0.75 credence is rational?[8] In light of the fact that her evidence is complex and therefore difficult to evaluate, in addition to the fact that other biologists and presumably her past self has made mistakes before, it seems that she has strong reason to doubt whether she has, on this occasion, judged her evidence exactly right. As such, it seems that higher-order uncertainty is exactly warranted in her epistemic situation.

[8] Hedden (2015) argues that how much a piece of evidence supports a proposition is an *a priori* question. Even so, agents should sometimes be uncertain whether they have evaluated their evidence correctly. Marine biologists, in this regard, are in the same boat as mathematicians. Both should, when their evidence is complex enough, be higher-order uncertain. See Christensen (2004) for an argument.

We thus face a dilemma. Higher-order certainty threatens dogmatism, while unconstrained higher-order uncertainty threatens akrasia. What is needed is an intermediate position that allows for rational higher-order uncertainty while preventing self-undermining epistemic states. In the next section, I will argue for a calibration principle which does just that.

3 Humility requires Total Calibration

Ellen should find a way to let the first-order credences she thinks might be rational play an appropriate role in determining what she actually believes. If this is correct, we need a principle which tells us how the biologist should calibrate her first-order uncertainty in light of her higher-order uncertainty. Such a calibration principle must be strong enough to rule out akratic attitudes, yet weak enough to allow genuine epistemic humility.

I will argue for two overall claims in this section. First, that epistemic akrasia can be understood in terms of credence only, without appeal to categorical beliefs. I suggest that akratic attitudes are misalignments between one's credence and one's expectation of the rational credence. My second claim is that a generalised reflection principle, Total Calibration, allows us to rule akrasia irrational without opting for dogmatism. To get there, I will start by reviewing two instructive accounts found in the literature.

3.1 Reliability calibration

One natural way to implement calibration is by appealing to beliefs about one's own reliability. If I take myself to be fallible in a systematic way, it is tempting to think that my credences should reflect this self-assessment. If I recognise myself to be, say, 90% reliable on propositions of a certain type, then

my credences should arguably be adjusted to reflect this.

White (2009) and Sliwa and Horowitz (2015) suggest that the rational credence is the weighted average of one's initial assessment and one's estimated frequency of error. On this picture, if an agent is aware of her own reliability, then her first-order credences ought to be adjusted accordingly.[9] Here is White's (2009: 234) formulation: 'Conditional on the reliability of a process producing a certain outcome being x, our credence that that outcome obtains should be x, absent more direct evidence bearing on the outcome.' We can formalise the idea as follows:

> **Reliability Calibration:** Where R_p is the reliability of judging that the proposition p is true, for any credence function $\mathcal{C}$, $\mathcal{C}$ is rational only if
>
> $$\mathcal{C}(p \mid R_p = x) = x$$

White generalises the proposal to cases where you are uncertain how reliable you are. Perhaps, you are uncertain whether you are 90% reliable or whether you are only 80% reliable. Take the average of these reliability hypotheses, weighted by your credence in the respective reliability hypotheses and you get your *expected* reliability.

$$\mathbb{E}[p] = \sum_i i \cdot \mathcal{C}(R_M = i) \tag{1.1}$$

According to this proposal, conditional on your expected reliability being x,

[9] Of course, your beliefs about your own reliability can be irrational. For example, they might be based on wishful thinking or forgetfulness of past mistakes. In such cases, your reliability beliefs will not provide normative guidance. What you *should* do is to change your beliefs in accordance with your evidence. Having done that, you will have rational beliefs about your own reliability. And these, according to reliability calibrationists, do provide normative guidance.

the rational credence to have is x.[10]

Suppose our biologist has a good idea of her own reliability in identifying baleen whales when they breach for air, but her evidence does not conclusively settle the question. Thus, she is rationally uncertain whether she is 90% reliable (she gives this hypothesis 0.6 credence) or whether she is 80% reliable (she gives this hypothesis 0.4 credence). Taking the weighted average, we get her expected reliability: $\mathbb{E}[p] = 0.9(0.6) + 0.8(0.4) = 0.86$. Suppose the biologist observes a baleen whale breach for air. Reliability Calibration tells us what first-order credence the biologist should have: $\mathcal{C}(p \mid R_p = 0.86) = 0.86$.

Despite its intuitive appeal, the reliability-based approach faces a difficulty. As Isaacs (2019) argues, relying on expected reliability in this way leads to a form of the base-rate fallacy. Place the biologist in coastal waters in the North Atlantic. She observes a whale and judges it to be a Sei. Realising that her expected reliability is 86%, should she therefore form a 0.86 credence in the whale being a Sei? No. Sei whales are rare in these waters. Conditional on observing a whale, the prior probability of the whale being Sei is rather small. It is antecedently much more likely she is observing, say, a Finn whale which looks similar and much more widespread. Consequently, the posterior probability of the whale being a Sei will be significantly lower than 0.86 even though the biologist is quite reliable in identifying Sei whales when they do in fact surface.

I have little doubt that the reliability approach can be made to work by tinkering with the technical details. A solution would be welcome since

[10] Summing to determine expectation only works on the assumption that the state space is finite. If an agent considers an infinite range of hypotheses, the expectation is determined by the integral of the hypotheses. This need not concern us here, except that it calls for a simplification of the notation: use the $\mathbb{E}$ function as a shorthand for the appropriate expectation function.

it is indeed intuitive that our rational judgments about our own reliability should inform what first-order credences are rational. For now, I will bypass the question of how higher-order judgments about *reliability* constrain rational beliefs, and ask instead how higher-order judgments about *rationality* constrain first-order judgments. This raises the level of abstraction, so to speak, since higher-order beliefs about reliability (and accuracy, for that matter[11]) arguably constrain higher-order beliefs about rationality.

3.2 Rational reflection

We can formulate calibration in terms of *expected rational credence* rather than *expected reliability*. Christensen's (2010b) Rational Reflection does this.[12] The idea is to require agents to match their credences to what they expect to be rational. We do this by introducing a probability function $\mathcal{P}$ which specifies the rational credence for an agent given that agent's evidence. When Christensen introduces $\mathcal{P}$, he says it denotes 'the credence that would be maximally rational for someone in that agent's epistemic situation'.[13] What determines the maximally rational credence is the *total evidence* available to

[11] Lewis (1971) argues that rational credences must be *immodest* in the sense that any rational credence function must judge itself to be at least as *accurate* as any other. And it seems plausible that rational higher-order credences about one's accuracy also inform which first-order credences are rational.

[12] The principle is so named given its analogy to van Fraassen's (1984) diachronic reflection principle. Christensen is the first that I know of to formulate the principle in terms of one's current beliefs about what is rational. A version of the principle is sometimes attributed to Miller (1966) who assumes it in a *reductio* against the coherence of higher-order probabilities. See Skyrms (1980b) for a response.

[13] For subjectivists and for at least some epistemic expressivists, $\mathcal{P}$ may ultimately be reducible to $\mathcal{C}$. See Greco (2014). Nothing in my argument requires an objectivist conception of rationality. I assume only that agents can represent alternative candidate rational credences and be uncertain which of these is correct.

the agent. The following reflection principle therefore requires that you defer to the credences that you believe an idealised version of yourself to hold:

> **Rational Reflection:** Where $\mathcal{P}$ is the rational credence function given the available evidence, for any credence function $\mathcal{C}$,$\mathcal{C}$ is rational only if
>
> $$\mathcal{C}(p \mid \mathcal{P}(p) = x) = x$$

This principle tells you to adopt the credence which you believe that your total evidence rationalises. If you learn—and thereby become certain—that the rational credence in p is x, then your credence in p should be x.

Rational Reflection is formulated as a conditional constraint. But just as we saw for Reliability Calibration, once this constraint is imposed on every possible value of the rational credence, standard probabilistic coherence fixes the unconditional credence as well. In this way, Rational Reflection entails a calibration requirement: an agent's credence must equal her *expectation* of the rational credence. To see this, we need only appeal to the Law of Total Expectation.[14] This tells us that:

$$\mathcal{C}(p) = \sum_x \mathcal{C}(p \mid \mathcal{P}(p) = x) \cdot \mathcal{C}(\mathcal{P}(p) = x) \tag{1.2}$$

Rational Reflection tells us that the first term in the scope of the sum is equal to x. Substituting identicals, we get the following requirement:

$$\mathcal{C}(p) = \sum_x x \cdot \mathcal{C}(\mathcal{P}(p) = x) \tag{1.3}$$

The right-hand side is the expectation of the rational credence. Rational

[14] The Law of Total Expectation says that the expected value of some random variable X equals the expected value of the conditional expected value of X given Y: $\mathbb{E}[X] = \mathbb{E}[\mathbb{E}[X \mid Y]]$. When the state space is finite and the random variable a proposition, this means that $\mathcal{C}(X) = \sum_i \mathcal{C}[X \mid Y_i]\mathcal{C}(Y_i)$. So, let the random variable X be $\mathcal{C}(p)$. And let Y_i be $\mathcal{P}(p) = x_i$.

Reflection therefore entails:

$$\mathcal{C}(p) = \mathbb{E}_{\mathcal{C}}[\mathcal{P}(p)] \tag{1.4}$$

In words: an agent is rational only if their first-order credence matches their higher-order expectation of the rational credence.

Apply this to the biologist. The biologist should set her first-order credence at the value she expects a rational evaluator of her evidence to adopt. If she expects that a rational assessment of her total evidence supports the French Polynesia hypothesis to degree x then she should adopt x herself. This also holds in the whale sighting case. Since the base-rate of Sei whales is low, the rational credence is also low.[15] She distributes her uncertainty over a number of hypotheses, for example $\mathcal{C}(\mathcal{P}(\text{Sei}) = 0.3) = 0.8$ and $\mathcal{C}(\mathcal{P}(\text{Sei}) = 0.35) = 0.2$. Taking the weighted average, the rational credence for the biologist according to equation 1.4 is $\mathcal{C}(\text{Sei}) = 0.31$. This allows us to respect base-rates since the rational credence respects base-rates. Even if the biologist is highly reliable with respect to identifying baleen whales, her rational credence in observing a Sei whale will still be low.

I think the spirit of Rational Reflection is correct. However, there is a hiccup with epistemic akrasia which will require us to generalise the principle. I turn to that in the following.

[15] This is arguably where considerations of reliability enter the picture. Information about one's reliability can rationally affect how one distributes credence over the competing rationality hypotheses. I do not attempt a general account of how reliability information determines these higher-order credences, but nothing in what follows depends on any particular view.

3.3 Total calibration

I said above that credal akrasia is a mismatch between one's credence and one's evaluation of that credence. We can now make that idea precise. I propose to understand epistemic akrasia as a mismatch between an agent's credence and her expected rational credence. Since Rational Reflection requires that agents align their credences with that expectation, Rational Reflection deems mismatches between first-order credence and higher-order expectation to be irrational.

However, before we can assess whether Rational Reflection succeeds in ruling out akrasia, we need to clarify the domain of quantification the principle ranges over. Suppose the range of propositions Rational Reflection ranges over is the set of all propositions. Then we get the result that an agent should be higher-order certain. Elga (2013) shows why.

Suppose that the proposition $\mathcal{C} = \mathcal{P}$ is part of the domain of propositions which Rational Reflection applies to. For example, you may consider whether your credence function is rational. Then Rational Reflection tells us that $\mathcal{C}(p \mid \mathcal{C} = \mathcal{P}) = \mathcal{P}(p)$. Now let p be the proposition that your credence function is rational, i.e. $\mathcal{C} = \mathcal{P}$. Then $\mathcal{C}(\mathcal{C} = \mathcal{P} \mid \mathcal{C} = \mathcal{P}) = \mathcal{P}(\mathcal{C} = \mathcal{P})$. The left-hand side is of course equal to 1, so the right-hand side must be as well. This means that if you give any positive credence to some credence function being rational then you must assign credence 1 to that function being rational.

As a result, an unrestricted Rational Reflection principle yields a requirement of higher-order certainty. Akrasia will thereby also be deemed irrational: the rational credence matches the higher-order expectation of the rational credence and rational agents are always certain what they expect the rational credence to be.

This does not mean that Rational Reflection entails that we should be insensitive to higher-order evidence. Evidence or impairment or disagreement may still call for revision, for example. But once you have revised, you should return to certainty that your current credence is perfectly rational. As such, Rational Reflection is extremely demanding. Indeed, it entails the kind of dogmatism which motivated us to introduce higher-order uncertainty in the first place. Said differently, unrestricted Rational Reflection does not permit epistemic humility.[16]

One natural response is to restrict the range of propositions to which Rational Reflection applies, in order to make it compatible with epistemic humility. While Christensen does not specify which propositions Rational Reflection is supposed to quantify over, we can stipulate that it applies to a wide class of what we might call descriptive propositions. Let S_1 be the set of descriptive propositions, which we can assume to be closed under the usual Boolean operations so that it forms an algebra. Descriptive propositions are ordinary propositions about the world such as propositions about whale behaviours and their relative frequencies in certain locations. They also include propositions that can constitute higher-order evidence, such as that one is impaired, that one's evidence is complex, or that an epistemic peer disagrees.

As such, the class of descriptive propositions is extremely broad. What the class excludes are normative propositions that ascribe epistemic status to credences or credence functions, for example that a credence function is rational, justified, or ideally rational. We can then let S_{n+1} be the set of propositions

[16] Elga's (2013) own solution is the much discussed New Rational Reflection principle. However, as Dorst (2019: §3.2) draws attention to, New Rational Reflection also allows for intuitively akratic attitudes. I therefore set Elga's principle to the side here and pursue my own attempt to salvage Rational Reflection.

of the form *The rational credence function over the propositions in S_n is $\mathcal{P}$*; we write this proposition as $\text{Rat}_n = \mathcal{P}$. To keep things simple, I assume that there are only finitely many candidate rational credence functions at each level. And I assume that S_{n+1} forms a partition, meaning the rationality-hypotheses are mutually exclusive and jointly exhaustive.

Under the restriction to descriptive propositions, Elga's proof no longer goes through. However, this restriction introduces a second problem. If calibration applies only to descriptive propositions, then it guarantees alignment between first-order credence and higher-order expectation only relative to whatever higher-order distribution you currently have. But agents may themselves have higher-order doubt about the rationality of that very distribution over rationality-hypotheses.

As Skipper (2021) points out, even a restricted reflection principle can leave room for akrasia once we allow uncertainty about one's higher-order expectations. Rational Reflection constrains your first-order credence only by appealing to your current higher-order distribution over hypotheses about the rational first-order credence. When you are uncertain about the rationality of your higher-order distribution, you are correspondingly uncertain about what you should expect the rational first-order credence to be.

In such cases, Rational Reflection no longer secures the intended enkratic coherence. Although your first-order credence can match the expectation determined by your current higher-order state, you may nonetheless expect, at the next level up, that this expectation is mistaken. In the hierarchy of your doxastic attitudes, there is some mismatch between a level of your doxastic state and your own best estimate of what would be rational at that level.

We are thus faced with our dilemma once again. If Rational Reflection

ranges over all propositions, it entails implausible higher-order certainty. If it ranges only over descriptive propositions it fails to secure the enkratic coherence. Skipper (2021) uses this observation to motivate access internalism as a requirement of rationality: rationality requires higher-order certainty of what you expect to be the rational credence.[17] This way, Rational Reflection is guaranteed to rule out akratic attitudes.

Let me instead propose a less demanding solution. Notice that we are now appealing to three layers of uncertainty. First-order uncertainty about whether p is true, second-order uncertainty which of, say, $\mathcal{C}(p) = 0.2$, $\mathcal{C}(p) = 0.3$, or $\mathcal{C}(p) = 0.4$ is rational, and third-order uncertainty whether the second-order distribution is rational. If Rational Reflection only specified how second-order credences should constrain first-order credences, the point is that third-order uncertainty can reintroduce akratic attitudes. But we need not appeal to third-order *certainty* to rescue the insight of Rational Reflection.

We can generalise the calibration requirement across the hierarchy of higher-order attitudes. Whenever an agent has higher-order uncertainty about what is rational at some level, that higher-order attitude should itself be aligned with the agent's expectation of what is rational at the next level. In this way, the agent is not required to be certain of her expectations. She is required only to defer, at each level, to her best estimate of what rationality requires at that level.

By the reasons which motivated Rational Reflection in the first place, we can require second-order credences to be calibrated in light of the third-order

[17] Skipper calls the requirement Top-Down Fixation. In slightly altered notation, his requirement is: $(\mathbb{E}_{\mathcal{C}}[\mathcal{C}(p)] = x) \rightarrow (\mathcal{C}(\mathbb{E}_{\mathcal{C}}[\mathcal{C}(p)] = x) = 1)$. If you have an expectation of the rational credence, you should be certain of that expectation. As a result, second-order uncertainty is permissible but third-order uncertainty is not. You are then guaranteed to have a second-order expectation of the rational first-order credence.

expectation. If, at the third level you expect that your second-order expectation is irrational, then, by your own lights, you should change the second-order expectation. Once the second level is calibrated, we can calibrate first-order credences in light of the calibrated second level. And so on, for all levels of uncertainty the agent has.

This gives us a Rational Reflection principle which applies to all levels of uncertainty:

> **Total Calibration:** Let $\mathcal{C}_n$ be the credence function of an agent over the propositions in $\mathcal{S}_n$. Let Δ_n be the set of all possible candidate rational probability functions over the propositions in $\mathcal{S}_n$. Then, for each $n \geq 1$ and for each $p \in \mathcal{S}_n$,
>
> $$\mathcal{C}_n(p) = \sum_{\mathcal{P} \in \Delta_n} \mathcal{C}_{n+1}(\mathrm{Rat}_n = \mathcal{P}) \cdot \mathcal{P}(p).$$

Given Total Calibration, every level of uncertainty should match the higher-level expectation of the rational credence. We need not require higher-order certainty, but we can require that lower-order credences match higher-order expectations. The principle therefore accommodates cases of uncertainty about expectations. Thus, we need not assume any strong access internalist requirement. We can leave it open whether agents can have higher-order doubt to an arbitrary order. Insofar as their doxastic lives do have such complex structures, each level ought to be calibrated. This rules out any mismatches where any credences are different from your best estimate of the rational credence.

As I mentioned, mismatches between your expectation of the rational credence and your actual credence is a natural precisification of epistemic akrasia. And this is exactly the kind of mismatches Total Calibration rules out. If so, higher-order uncertainty by itself does not permit akrasia. As long

as every layer of uncertainty is calibrated according to your own expectations of the rational credence, you have done as good a job as you can. By your own lights, there is no alternative credal state you expect to be epistemically superior. Your evidence might leave you higher-order uncertain, which is unfortunate, just as our evidence may leave us first-order uncertain whether whales drink water, and so on. But there is nothing epistemically irrational about it.

I have not attempted a full defence of Total Calibration but only sketched the motivations behind accepting any such principle. Yet, insofar as we have good reason to accept Rational Reflection, it seems to me that we have good reason to accept Total Calibration: both are motivated by the idea that one should defer to the expected rational credence. The virtue of Total Calibration is that by requiring deference all the way up, we rule out akrasia without requiring higher-order certainty. Still, this is not enough to capture epistemic humility in its entirety. As I will argue in the next section, agents can satisfy Total Calibration yet differ significantly in how humble they are.

4 Humility requires Proportional Resilience

We started with the idea that our fallibility requires us to be open to the possibility of error. Total Calibration explains how higher-order uncertainty about the rationality of our credences in p should make us change our credence in p. But this does not exhaust the normative influence of higher-order uncertainty. A big part of the intuitive notion of epistemic humility is linked with open-mindedness.[18] In this section I will argue that higher-order uncertainty can

[18] Open-mindedness is usually discussed outside the Bayesian literature. See Hare (1985) for an early influential account and see Fantl (2018: Ch. 1) and Whitcomb et al. (2017) for a good overview of the literature.

be linked with such open-mindedness through a principle I call Proportional Resilience.

Let us start with an intuitive thought: if you should worry about whether you are rational, you should be open to changing your mind. Consider the following case:

> **Ellen and Simon:** Independently of each other, Ellen and Simon investigate the geographical origin of a particular humpback whale observed in the Antarctic. Each considers how their evidence bears on the hypothesis p that the whale belongs to a population originating near French Polynesia. Ellen has tested her method extensively. She has often made predictions about the origins of humpback whales based on their song, and has a good sense of how to do it. For this reason, she is quite confident that her judgment is rational. She rationally assigns a 0.75 credence which she is rationally confident is rational. By contrast, Simon has been trained but has not had occasion to test his methodology much. As it happens, Simon also rationally assigns a 0.75 credence, but he harbours significant doubt as to whether it is rational.

The salient feature of the case is that Ellen and Simon both rationally give the French Polynesia hypothesis the same credence. Yet, because of the significant difference in the quality and size of their respective batches of evidence which bear on the hypothesis, Ellen and Simon rationally have quite different higher-order credences about the rationality of their 0.75 credence.

Since they are both rational, both satisfy Total Calibration. This means that the normative force of their higher-order uncertainty has already appropriately influenced their first-order credences. Yet, because Simon is more uncertain about the rationality of his calibrated credences, Simon should be

more epistemically humble than Ellen should be.

I want to propose that this difference in humility should manifest itself as a difference of much Simon and Ellen should expect to change their minds in the face of new evidence. Suppose that both Ellen and Simon consider a possible future piece of evidence e: the whale is later observed migrating eastward, behaviour more characteristic of populations originating off the coast of Chile in their return for the breeding season.

What should they think about this possible piece of evidence? In particular, how should they regard the potential impact of e given their present epistemic situations? Simon should expect that learning e would significantly lower his credence in p. Given how limited his experience is, he should anticipate that new information of this kind would likely require substantial revision. Ellen, by contrast, might expect the same evidence to have a more modest effect. Because her belief is based on greater experience, she can reasonably regard it as more stable in the face of additional information. In other words, whereas Simon should regard e as a piece of counterevidence, Ellen should incline towards treating e as misleading evidence. She should not disregard the evidence, of course, but it will sway her very little.

We might say that Simon's credence in p should be less *resilient* than Ellen's. He should be disposed to changing his mind if relevant counterevidence arises. And, since he already doubts his own rationality, he should expect that such changes are likely given new evidence. By contrast, Ellen's greater evidential position should make her expect her beliefs to be relatively more stable.

The intuitive cases motivating epistemic humility suggest that rational agents should often recognise that their present beliefs are provisional. They

may be rational given the evidence currently available, but that evidence may be limited, noisy, or difficult to assess. In such circumstances, higher-order uncertainty should make the rational expect their credence to be sensitive to further information. As such, epistemic humility consists in open-mindedly recognising that one's beliefs are fragile.

We can spell this out in the Bayesian framework. A credence function does not merely assign unconditional probabilities. It also encodes a dispositional profile of belief revision, given by the agent's conditional credences. These determine how the agent expects her beliefs to change upon learning various possible pieces of evidence. What is missing is a normative link between higher-order uncertainty and this dispositional profile. Once that link is made explicit, we can give a Bayesian account of open-mindedness without introducing any new doxastic attitudes.

I propose to model open-mindedness using Skyrms' notion of resilience. Following Skyrms (1977, 1980a), the resilience of a credence measures how much it would change upon conditionalising on a given piece of evidence. Highly resilient credences are those that would remain largely unchanged across a wide range of evidential updates. Fragile or unresilient credences are those that would shift substantially in light of new information.

Skyrms' original notion defines resilience in terms of the maximum change a credence can undergo given any proposition one might learn within some set of salient propositions. For present purposes, however, what matters is not merely how much a credence could *possibly* change conditional on some evidence, but, given an agent's current epistemic situation, how much she expects to change.

Once again, we get the expectation by weighting the resilience measure

with respect to some evidence proposition e with the credence that e will be learnt. In what follows, I treat the relevant evidence outcomes $E = \{e_i\}$ as a partition of possible experiences the agent might have. The weight $\mathcal{C}(e_i)$ is therefore the agent's credence of undergoing e_i. On this interpretation, expected resilience measures how stable the agent expects $\mathcal{C}(p)$ to be across the experiences she expects to have.

This yields a measure of expected resilience, understood as the agent's expectation, prior to having any new experiences, of how stable her credence will be:

> **Expected Resilience:** Let $\mathcal{C}$ be the credence function of an agent and let $E = \{e_i\}$ be a partition of epistemically possible experiences for that agent. Then, where $|x - y|$ is the absolute difference between x and y, for any proposition p, the expected resilience of $\mathcal{C}(p)$ is given by:
>
> $$1 - \sum_{e \in E} \mathcal{C}(e)|\mathcal{C}(p) - \mathcal{C}(p \mid e)|.$$

Expected resilience is one minus the weighted average of how much the agent expects her credence in p to change upon learning epistemically possible evidence, where the weights are given by her current credences in those evidential propositions. The principle thus measures the expected degree of doxastic stability.

This allows us to characterise open-mindedness within the Bayesian framework. To be close-minded is for one's expected resilience to be high. To be open-minded is for one's expected resilience to be low: to expect one's beliefs to be sensitive to further information. On this account, open-mindedness is a structural feature of a credence function, reflecting how seriously the agent takes the possibility of future revision.

At this point, it should be noted that higher-order uncertainty, even when combined with Total Calibration, does not by itself determine expected resilience. Two agents may have identical first-order credences in p and identical higher-order expectations of the rational credence in p, yet differ arbitrarily in how much they expect their credence in p to respond to new evidence. Total Calibration constrains the level of credence an agent may hold, but it does not constrain the sensitivity of that credence to further information. As a result, an agent may be fully calibrated and highly confident in her rationality, while still regarding her beliefs as fragile. Another agent may have the same calibrated credence but treat it as highly resilient. Calibration and resilience are therefore logically independent.

But plausibly they are normatively linked. It seems plausible that higher-order uncertainty normatively constrains how much you should expect your credence to withstand the addition of new evidence. The more reason you have to doubt your own rationality, the more reason you have to expect there to be counterevidence to what you believe. We can spell this out with the following principle:

> **Proportional Resilience:** Let $\mathcal{C}$ be the credence function of an agent and hold fixed the agent's first-order evidential situation. Let E be a partition of the agent's epistemically possible experiences. Then, for any descriptive proposition p, $\mathcal{C}$ is rational only if the expected resilience of $\mathcal{C}(p)$ relative to E is a weakly increasing function of the agent's higher-order confidence that $\mathcal{C}(p)$ is rational.

Proportional Resilience requires that increases in higher-order confidence be accompanied, all else equal, by increases in expected resilience, and that increases in higher-order *un*certainty be accompanied by *de*creases in expected

resilience.[19]

While I have defined open-mindedness in terms of one's *expected* resilience, Proportional Resilience places constraints on first-order credences directly. When higher-order confidence increases, Proportional Resilience requires an increase in expected resilience. That forces a change in the first-order credences themselves since expected resilience is defined directly in terms of first-order credences. And since first-order credences guide reasoning and action, this change has such downstream effects. I will return to this point in the next section.

It is important to emphasise that Proportional Resilience does not rest on any simple association between the amount of first-order evidence and the expected stability of a credence. In particular, it is not the case that having little evidence automatically calls for low expected resilience, or that having a large body of evidence automatically calls for high expected resilience. In many cases, a rich evidential base is also a complex one, involving noisy data, methodological assumptions, or difficult judgments of reliability. Such complexity can itself be a source of higher-order uncertainty, and in those circumstances an agent may rationally expect her credences to be relatively fragile despite being supported by a great deal of evidence.

Conversely, sparse evidence can sometimes rationalise a great deal of higher-order confidence. For example, an agent who has almost no information about a particular coin may rationally be just short of certainty that the

[19] I formulate the principle in terms of a weakly increasing function because there is not a unique function linking the two, such as requiring that expected resilience always be less than or equal to higher-order uncertainty. Rather, the principle merely imposes a monotonic constraint: greater confidence in one's rationality should be accompanied by equal or greater confidence in the stability of one's beliefs, while greater self-doubt requires equal or greater openness to revision.

rational credence in its landing heads is 0.5. In that sense, she may be highly confident that her credence is rational despite her limited evidence. Yet that credence should nevertheless be highly unresilient, since she should expect it to change substantially once she begins to gather information about the coin's behaviour. For example, she should expect to change her 0.5 credence significantly if she were to observe 10 heads in a row (Jeffrey 1983: Ch. 12.5). Such evidence will make her suspect the coin is biased, and so drastically change her 0.5 credence that the next throw will come up heads.

This is not a counterexample to Proportional Resilience. The principle does not require high higher-order confidence to be accompanied by high expected resilience in absolute terms. Rather, it imposes a comparative constraint: holding the first-order evidential situation fixed, decreases in higher-order confidence should be accompanied by decreases in expected resilience, and increases in higher-order confidence should be accompanied by increases in expected resilience.

5 The theory at work: peer agreement decreases humility

We can see how the two proposed norms, Total Calibration and Proportional Resilience, interact by focusing on cases of peer disagreement and peer agreement since they serve as straightforward cases of higher-order defeat and higher-order confirmation.

Let us assume that the lesson of the peer disagreement literature is that higher-order evidence can call for revision to one's first-order credences (Christensen 2007; Feldman 2010; Elga 2007; Kelly 2010). Total Calibration explains this conciliationist intuition. Learning that a peer disagrees with you should make you lower your higher-order confidence. This shifts your

expectation of the rational credence: you divide your higher-order confidence between your initial judgment and the judgment of your peer. You should then shift your first-order credence according to Total Calibration. On pain of irrationality, your higher-order uncertainty calibrates your judgment according to your expectation of the rational credence.

Conciliationists disagree, of course, over what you should expect the rational credence to be given peer disagreement. If you should give equal credence to the rationality of your own initial credence and the credence of your peer, we get Elga's (2007) Equal Weight View. If we impose stronger objectivist requirements on what is rational in light of what the truth is, we get something closer to Kelly (2010). In fact, steadfasters who think peer disagreement does not call for any belief revision at all can still accept Total Calibration. If disagreement does not influence what you should expect the rational credence to be, Total Calibration will require no belief revision.

The point is that we can view these debates as debates over how disagreement bears on your rational higher-order credences about rationality. That is, when you learn a peer disagrees with you, what *should* you believe it is rational to believe? Total Calibration will be compatible with these different theories: after you have updated your higher-order credences in light of the disagreement, Total Calibration requires that you update your lower-order credences to match your expectation of the rational credence.

Notice however, that peer disagreement should oftentimes make you increasingly humble even after having calibrated.[20] For example, the disagreement might have taught you that the topic at hand was more difficult to assess than you had previously thought. This is easy to see in Christensen's

[20] Compare Steglich-Petersen (2019) who links symmetrical higher-order defeat to resilience via higher-order uncertainty.

(2007) restaurant case where you discover that your peer disagrees with you on a simple mental maths question. Even after you have consciliated the disagreement through calibration, it seems that your higher-order confidence in your updated credence should be lower relative to your initial higher-order confidence in your initial credence. After all, you have learnt that the mental maths question was one of the difficult ones. Perhaps both you and your peer made a mistake, and so on.

It thus seems plausible that in cases like these, peer disagreement rationalises an increase in humility which is not captured by focusing on changes to the credence you assign to the disputed proposition. Sometimes, disagreement should not only make us change our minds about the issue at hand, but it should also make us more sensitive to new evidence which bears on that issue.

To bring this out, focus on cases of peer *agreement*. Suppose you learn that a trusted peer, who has accessed and reviewed the same evidence as you have, gives p the same credence as you do. This should be reassuring news. It should allow you to become more confident in the rationality of your credence in p. Thus, agreement leads to an increase—however much will depend on the details—in higher-order confidence. However, this shift in higher-order confidence will not entail a shift in your expectation of the rational credence. Otherwise Total Calibration kicks in and shifts your first-order credence. Then peer agreement leads to disagreement!

Clearly the increased higher-order confidence in $\mathcal{C}(p)$ should not make you shift $\mathcal{C}(p)$. But it should somehow influence your first-order credences. Our second norm, Proportional Resilience, delivers the desired result. Peer agreement leads to increased higher-order confidence without shifting what the calibrated first-order credence is. But the increased confidence gained by

peer agreement influences your expected resilience. Learning that a peer agrees with you should make you at least a little more resistant to changing your mind in light of future counterevidence. Either because the peer agreement makes you more confident that there will be no counterevidence (that is, $\mathcal{C}(e)$ falls), or because the peer agreement makes you think that counterevidence is misleading (that is, $\mathcal{C}(p \mid e)$ moves closer to $\mathcal{C}(p)$).

For example, suppose Simon learns that Ellen agrees with him. Somehow, Simon comes to learn that Ellen has recorded the song of the same whale, done her analysis, and formed a 0.75 credence in the French Polynesia hypothesis. This is good news for Simon. As a result, he should become more confident that his own credence is rational. Arguably, Ellen is Simon's epistemic superior, not peer. This just means that Simon's higher-order confidence should increase to a greater extent when learning of the agreement. Conversely for Ellen: Simon's agreement will confirm her own rationality to a lesser extent.

Simon's increased higher-order confidence that his credence is rational should entail an increased expectation that his credence will remain unchanged as he gains new evidence. For example, Simon should expect to be more resilient with respect to the potential counterevidence e—observing the whale migrating eastward. Depending on the details of Simon's prior probabilities and evidential situation, Ellen's agreement should either make him assign a lower credence to e or else assign a lower evidential impact on the French Polynesia hypothesis conditional on learning that e. Either way, Simon should become at least a little less open to changing his mind.

The natural upshot is that peer agreement should make you less epistemically humble. It leads to greater higher-order confidence which, in turn, should lower your expectation that new evidence will call for belief revision.

6 Conclusion

Epistemic humility is a rational response to evidence of our own fallibility. Conversely, evidence that we have judged the evidence correctly should make us less humble. Not arrogant or dogmatic, of course, but appropriately confident of our own judgments.

I have argued that Bayesians models of rational belief can make sense of such epistemic humility in terms of higher-order uncertainty by imposing two norms of credences. *Total Calibration* allows higher-order uncertainty to bear on first-order credences by influencing which credence you should expect to be rational. The principle allows this without permitting akratic attitudes or requiring implausibly dogmatic higher-order certainty. *Proportional Resilience* links higher-order uncertainty to open-mindedness. The degree to which you expect new evidence to be counterevidence, should be proportional to the degree to which you remain higher-order uncertain after having calibrated your credences.

References

Briggs, R. A. (2009) 'Distorted Reflection'. *Philosophical Review*, 118.1, pp. 59–85. DOI: 10.1215/00318108-2008-029.

Christensen, David (2004) *Putting Logic in its Place: Formal Constraints on Rational Belief*. Oxford: Oxford University Press.

— (2007) 'Epistemology of Disagreement: The Good News'. *Philosophical Review*, 116.2, pp. 187–217.

— (2010a) 'Higher-Order Evidence'. *Philosophy and Phenomenological Research*, 81.1, pp. 185–215.

— (2010b) 'Rational Reflection'. *Philosophical Perspectives*, 24.1, pp. 121–140. DOI: 10.1111/j.1520-8583.2010.00187.x.

Christensen, David (2020) 'Akratic (Epistemic) Modesty'. *Philosophical Studies*, 178.7, pp. 2191–2214. DOI: 10.1007/s11098-020-01536-6.

— (2024) 'Epistemic Akrasia: No Apology Required'. *Noûs*, 58.1, pp. 54–76. DOI: 10.1111/nous.12441.

Dorst, Kevin (2019) 'Higher-Order Uncertainty'. In: *Higher-Order Evidence: New Essays*. Ed. by M. Skipper and A. Steglich-Petersen. Oxford University Press.

Egan, Andy and Adam Elga (2005) 'I Can't Believe I'm Stupid'. *Philosophical Perspectives*, 19.1, pp. 77–93. DOI: 10.1111/j.1520-8583.2005.00054.x.

Elga, Adam (2007) 'Reflection and Disagreement'. *Noûs*, 41.3, pp. 478–502. DOI: 10.1111/j.1468-0068.2007.00656.x.

— (2013) 'The Puzzle of the Unmarked Clock and the New Rational Reflection Principle'. *Philosophical Studies*, 164.1, pp. 127–139. DOI: 10.1007/s11098-013-0091-0.

Fantl, Jeremy (2018) *The Limitations of the Open Mind*. Oxford University Press.

Feldman, Richard (2010) 'Reasonable Religious Disagreements'. In: *Philosophers without Gods: Meditations on Atheism and the Secular Life*. Ed. by Louise Antony. OUP Usa, pp. 194–214.

van Fraassen, Bas (1984) 'Belief and the Will'. *Journal of Philosophy*, 81.5, pp. 235–256. DOI: 10.2307/2026388.

Gaifman, Haim (1985) 'A Theory of Higher Order Probabilities'. In: *Causation, Chance, and Credence*. Ed. by Brian Skyrms and William Harper. Kluwer Academic.

Garland, Ellen C. et al. (2011) 'Dynamic Horizontal Cultural Transmission of Humpback Whale Song at the Ocean Basin Scale'. *Current Biology*, 21.8, pp. 687–691. DOI: https://doi.org/10.1016/j.cub.2011.03.019.

Graf, Simon (2025) 'Epistemic Justification and Higher-Order Requirements'. *Episteme*, pp. 1–23.

Greco, Daniel (2014) 'A Puzzle About Epistemic Akrasia'. *Philosophical Studies*, 167.2, pp. 201–219. DOI: 10.1007/s11098-012-0085-3.

Hare, William (1985) *In Defence of Open-Mindedness*. Kingston, [Ont.] : McGill-Queen's University Press.

Hawthorne, James (2005) 'Degree-of-Belief and Degree-of-Support: Why Bayesians Need Both Notions'. *Mind*, 114.454, pp. 277–320. DOI: 10.1093/mind/fzi277.

Hazlett, Allan (2012) 'Higher-Order Epistemic Attitudes and Intellectual Humility'. *Episteme*, 9.3, pp. 205–223. DOI: 10.1017/epi.2012.11.

Hedden, Brian (2015) 'A Defense of Objectivism About Evidential Support'. *Canadian Journal of Philosophy*, 45.5, pp. 716–743. DOI: 10.1080/00455091.2015.1123038.

Hedden, Brian and Kevin Dorst (2022) '(Almost) All Evidence is Higher-Order Evidence'. *Analysis*, 82.3, pp. 417–425. DOI: 10.1093/analys/anab081.

Henderson, Leah (2014) 'Bayesianism and Inference to the Best Explanation'. *British Journal for the Philosophy of Science*, 65.4, pp. 687–715. DOI: 10.1093/bjps/axt020.

Horowitz, Sophie (2014) 'Epistemic Akrasia'. *Noûs*, 48.4, pp. 718–744. DOI: 10.1111/nous.12026.

Hume, David (1748) *An Enquiry Concerning Human Understanding*. Oxford University Press.

Isaacs, Yoaav (2019) 'The Fallacy of Calibrationism'. *Philosophy and Phenomenological Research*, 102.2, pp. 247–260. DOI: 10.1111/phpr.12640.

Jeffrey, Richard (1983) *The Logic of Decision*. 2nd ed. University of Chicago Press.

Kelly, Thomas (2010) 'Peer Disagreement and Higher Order Evidence'. In: *Disagreement*. Ed. by Richard Feldman and Ted A. Warfield. Oxford University Press, pp. 111–174.

Lasonen-Aarnio, Maria (2015) 'New Rational Reflection and Internalism About Rationality'. *Oxford Studies in Epistemology*, 5. DOI: 10.1093/acprof:oso/9780198722762.003.0005.

— (2020) 'Enkrasia or Evidentialism? Learning to Love Mismatch'. *Philosophical Studies*, 177.3, pp. 597–632. DOI: 10.1007/s11098-018-1196-2.

Lewis, David (1971) 'Immodest Inductive Methods'. *Philosophy of Science*, 38.1, pp. 54–63. DOI: 10.1086/288339.

Lipton, Peter (2004) *Inference to the Best Explanation*. New York: Routledge.

Maher, Patrick (2006) 'The Concept of Inductive Probability'. *Erkenntnis*, 65.2, pp. 185–206. DOI: 10.1007/s10670-005-5087-5.

Miller, David (1966) 'A Paradox of Information'. *British Journal for the Philosophy of Science*, 17.1, pp. 59–61.

Skipper, Mattias (2021) 'Does Rationality Demand Higher-Order Certainty?' *Synthese*, 198.12, pp. 11561–11585. DOI: 10.1007/s11229-020-02814-w.

Skyrms, Brian (1977) 'Resiliency, Propensities, and Causal Necessity'. *Journal of Philosophy*, 74.11, pp. 704–713. DOI: 10.2307/2025774.

— (1980a) *Causal Necessity: A Pragmatic Investigation of the Necessity of Laws*. Yale University Press.

— (1980b) 'Higher Order Degrees of Belief'. In: *Prospects for Pragmatism*. Ed. by D. H. Mellor. Cambridge University Press, pp. 109–137.

— (1984) *Pragmatics and Empiricism*. Yale University Press.

Sliwa, Paulina and Sophie Horowitz (2015) 'Respecting All the Evidence'. *Philosophical Studies*, 172.11, pp. 2835–2858. DOI: 10.1007/s11098-015-0446-9.

Smithies, Declan (2019) *The Epistemic Role of Consciousness*. New York, USA: Oxford University Press.

Steglich-Petersen, Asbjørn (2019) 'Higher-Order Defeat and Doxastic Resilience'. In: *Higher-Order Evidence: New Essays*. Ed. by Mattias Skipper and Asbjørn Steglich-Petersen. Oxford University Press, pp. 209–225.

Tal, Eyal (2020) 'Is Higher-Order Evidence Evidence?' *Philosophical Studies*, 178.10, pp. 3157–3175. DOI: 10.1007/s11098-020-01574-0.

Titelbaum, Michael G. (2015) 'Rationality's Fixed Point'. *Oxford Studies in Epistemology*, 5. DOI: 10.1093/acprof:oso/9780198722762.003.0009.

Weisberg, Jonathan (2009) 'Locating IBE in the Bayesian Framework'. *Synthese*, 167.1, pp. 125–143. DOI: 10.1007/s11229-008-9305-y.

Whitcomb, Dennis et al. (2017) 'Intellectual Humility: Owning Our Limitations'. *Philosophy and Phenomenological Research*, 94.3, pp. 509–539. DOI: 10.1111/phpr.12228.

White, Roger (2009) 'On Treating Oneself and Others as Thermometers'. *Episteme*, 6.3, pp. 233–250. DOI: 10.3366/E1742360009000689.

Williamson, Timothy (2000) *Knowledge and its Limits*. Oxford University Press.

— (2014) 'Very Improbable Knowing'. *Erkenntnis*, 79.5, pp. 971–999. DOI: 10.1007/s10670-013-9590-9.