Erkenntnis (2025) 90:3613–3633
https://doi.org/10.1007/s10670-024-00861-0

ORIGINAL RESEARCH

Some Worries About the Probability Account of the Feelings of (Un)Certainty

Sergiu Spatan[1]

Received: 31 January 2024 / Accepted: 21 August 2024 / Published online: 25 September 2024

Abstract
In recent papers, Peter Carruthers and others have argued that the feeling of uncertainty is not metacognitive (i.e., it is not elicited by second-order cognitive appraisals) but is elicited solely by first-order likelihood estimates—a *probability account* of the feeling of uncertainty. In this paper, I make a case for why a probability account is sufficient to explain neither the feeling of uncertainty nor the feeling of certainty in self-reflecting humans. I argue first that humans' feelings of (un)certainty vary in ways that their probability estimates on the matter do not, and second that probability accounts elide the essentially epistemic nature of epistemic feelings.

1 Introduction

Here is an experience you have surely had. You close the door of your apartment, ready to set off on vacation, only to freeze: you might not have turned off the stove. A feeling of uncertainty has crept into your mind and consulting your memory on the doorstep cannot dissipate it. You feel compelled to go back in and check with your own eyes that the stove really is turned off. Once you do so, your worries vanish entirely; the unwelcome feeling of uncertainty is replaced by calm assuredness. You leave the apartment feeling relaxed.

The feeling of uncertainty before you check the stove and the assurance that you feel afterwards are instances of *epistemic feelings* (see de Sousa, 2009; Arango-Muñoz & Michaelian, 2014; Meylan, 2014). Epistemic feelings are affective states that concern the epistemic position in which the subjects find themselves—am I justified in believing that I turned off the stove?—and which therefore pertain to believing truths and avoiding errors. Other epistemic feelings include the tip-of-the-tongue phenomenon (Hart, 1965), the feelings of knowing (Koriat, 2000), forgetting (Koriat

✉ Sergiu Spatan
sergiu.spatan@tu-dresden.de

[1] Institute of Philosophy, Technische Universität Dresden (Dresden University of Technology), Zellescher Weg 17, Room A425, 01069 Dresden, Germany

et al., 2004), familiarity (Whittlesea & Williams, 2000), curiosity (Goupil & Proust, 2022), and many others (see Arango-Muñoz & Michaelian, 2014 for an overview).

The most urgent ontological question in the literature about the nature of epistemic feelings is whether or not these affective states are metacognitive (see Carruthers, 2008; Kornell, 2014; Smith, Couchman, and Beran 2014; Carruthers, 2017; Beran, 2019). Metacognition, or second-order cognition, is usually defined as the thought processes by which we evaluate, monitor, control, and generally regulate our first-order cognitions (Dunlosky & Metcalfe, 2009). Supporters of a metacognitive account of epistemic feelings claim that these feelings are metacognitive because they are elicited by second-order appraisals of first-order mental states. According to this line of reasoning, my feeling of uncertainty regarding whether the stove is turned off is the result of my appraisal of the justification that I have to believe that the stove is turned off: assessing this justification, I find that it is insufficient.

Peter Carruthers and others have recently argued against a metacognitive account of epistemic feelings (Carruthers, 2008, 2017; Carruthers & Ritchie, 2012; Carruthers & Williams, 2019; Kornell, 2014). The main argument for the claim that epistemic feelings are not metacognitive is based on the observation that nonhuman animals seem to experience epistemic feelings as well. However, the argument runs, metacognition is too complex a cognitive system to attribute to nonhuman animals. If nonhuman animals are indeed incapable of metacognition, then epistemic feelings cannot be metacognitive. These authors suggest that epistemic feelings ought to be explained in first-order cognitive terms alone. Carruthers, for instance, claims that the feeling of uncertainty about the completion of an action is elicited by first-order appraisals that the action is unlikely to succeed; a definition that avoids appealing to metacognition (Carruthers, 2017, p. 70).[1] Call this the *probability account* of the feelings of certainty and uncertainty (feelings of (un)certainty, for short).

In this paper, I shall not settle the issue of whether or not epistemic feelings in general, nor in nonhuman animals in particular, are metacognitive. My principal aim is to argue that the probability account of the feelings of (un)certainty *in self-reflecting humans* faces some serious problems—which, indeed, might end up being problematic for the anti-metacognitivist camp.

The paper is organized as follows. In Sect. 2, I sketch a metacognitive account of the feelings of (un)certainty with which to compare the probability account. In Sect. 3, I introduce three versions of a non-metacognitive, probability-based account of the feelings of (un)certainty, based on suggestions by Carruthers and collaborators: the *simple probability account*, the *weighted probability account*, and the *shifty-threshold probability account*. In the remainder of the paper, I present two arguments against these accounts: the *argument from variability* against the first two accounts (in Sect. 4), and the *argument from epistemicity* against the shifty-threshold probability account (in Sect. 5). I conclude in Sect. 6.

[1] The assumption here is that appraisals of the likelihood of success are cognitively less demanding than metacognitive appraisals. The former do not involve representations of other mental states; they merely require processing probabilistic information, a feat that, e.g., mice and rats seem to be capable of (Carruthers and Ritchie 2012; Gallistel et al., 2001; Balci, Freestone & Gallistel, 2009).

2 The Metacognitive Account

In what sense are epistemic feelings—more specifically, the feelings of (un)certainty—metacognitive? According to one line of thought, epistemic feelings are metacognitive because they are based on meta-representations of first-order cognitions (Flavell, 1979; Nelson & Narens, 1990). This view entails that there are two levels of cognition, an object level and a meta level, such that the states at the second-order meta level monitor and control the states at the first-order object level. Whereas first-order cognitions are about reality, second-order cognitions are about first-order cognitions; second-order cognitions are thus representations of representations, or meta-representations. For instance, the tip-of-the-tongue feeling I have when I try to think of Mark Twain's birth name[2] is based, according to this account, on a metacognitive evaluation of my memory, such that a meta-representation of the required token memory is formed but the memory itself is incomplete or not retrieved. It is in virtue of the meta-representation that I have the tip-of-the-tongue feeling. The object of the tip-of-the-tongue feeling is a first-order cognitive state—the incomplete or irretrievable memory—and the feeling is therefore metacognitive. Similarly, my feeling of uncertainty concerning whether or not the stove is turned off is elicited by the meta-representational appraisal that I am not really warranted to believe that it is turned off. Epistemic feelings are metacognitive, according to this line of thought, because they are based on meta-representational appraisals.

Subsequent research in comparative psychology (e.g., Beran, 2019; Morgan et al., 2014; Smith et al., 2003, 2008, 2014, 2019) has shown that nonhuman animals might also have epistemic feelings, particularly the feeling of uncertainty.[3] Because the possibility of meta-representational abilities in nonhuman animals seems far-fetched to some authors,[4] this research leads some to deny that metacognition is

[2] If you are like me, when asked about the birth name of Mark Twain, you have the distinct feeling that you know the answer to this question but cannot recollect it right now. The answer is *on the tip of your tongue*. (The answer is Samuel Clemens, by the way).

[3] In one influential experiment (Smith, Shields & Washburn, 2003, 321), humans and rhesus monkeys were presented with sparsely and densely pixelated stimuli and three responses: response D for a densely pixelated stimulus, response S for a sparsely pixelated stimulus, and the Uncertain response (a star) for when the participants did not know what to choose. In the case of a match between the stimulus and the response, the participants received a reward, and otherwise received an unwelcome time-out. Pushing the star, nevertheless, simply allows the participants to decline the trial and move immediately to the next one. As experimenters increased the difficulty of the task and introduced more stimuli on the dense/sparse boundary, both humans and monkeys utilized the Uncertainty button more often, especially when the identification got very hard. Post-experimentally, the humans reported that they pushed the star whenever they felt uncertain. By analogy, and noticing the similar behaviour, it seems that the monkeys also experienced some kind of feeling of uncertainty, in virtue of which they chose the safe option. The assumption that the animals experience epistemic *feelings* is based on an analogy with us humans when we are posed in similar situations (Smith, Jackson, and Church 2019). Similar results were obtained with dolphins (Smith, 2010), pigeons (Iwasaki, Watanabe & Fujita 2018), or rats (Foote & Crystal, 2012).

[4] See Carruthers, 2008; Carruthers, Fletcher & Ritchie, 2012; Carruthers & Ritchie, 2012; Kornell, 2014; Perner, 2012. This skepticism is motivated by the fact that meta-representational metacognition requires representing concepts of mental states (e.g., belief), a capacity that human infants and nonhuman animals do not seem to possess (see, for instance, the incapacity of young children to pass the false belief test; Wimmer & Perner, 1983; Rakoczy, 2017).

essentially meta-representational (Proust, 2007, 2013, 2019; Dokic, 2012, 2014; Arango-Muñoz, 2014; see also Hart, 1965 for a precursor of this view). Proust, for instance, defines metacognition as "the set of capacities through which an operating cognitive subsystem is evaluated *or* represented by another subsystem in a context-sensitive way" (2013, p. 4, my emphasis). According to this view, an evaluation of one cognitive subsystem by another is still metacognitive even if the latter does not represent the former. Proust calls the left disjunct of her definition, where one subsystem evaluates another without representing it, "procedural metacognition." In procedural metacognition, judgments are inferred from various cues and are not based on explicit meta-representations.[5] Some of the most important cues are one's familiarity with a certain input, the ease of processing some information, or the general fluency that occurs when, e.g., a subject attempts to recollect a particular memory item (see Alter & Oppenheimer, 2009).[6] For instance, I might have a tip-of-the-tongue experience when thinking about Mark Twain's birth name not because I have a partial representation of the answer, but merely because the name "Mark Twain" sounds familiar to me. Similarly, the feeling of uncertainty that a monkey or a human experience in Smith and colleagues' experiments may not be elicited by a meta-representation of their epistemic situation but rather may be inferred unconsciously from these sorts of cues. Metacognition is involved in both cases, as long as one operating cognitive subsystem evaluates or represents another such subsystem.

In what follows, I sketch a hybrid metacognitive account of the feelings of (un) certainty (which I discuss at greater length in Spatan, 2022) that incorporates both meta-representational capacities (in organisms capable of self-reflection) and procedural metacognition (also in organisms that lack meta-representational capacities). This toy metacognitive account will then be used as a point of reference with which to contrast the probability account.

I take it that one of the functions of metacognition is to manage the organism's cognitive resources (Proust, 2013, pp. 4–5). Call this *gate-keeping metacognition*, as it amounts to monitoring the organism's internal states in order to regulate the deployment of cognitive effort. Now, the feelings of (un)certainty can be conceived as a part of this gate-keeping mechanism (in those organisms capable of epistemic

[5] Note that supporters of the non-metacognitive view of epistemic feelings also accept that our epistemic feelings are typically based on cues such as familiarity and fluency. They do not think, though, that metacognition is involved, because they take meta-representations to be an essential part of metacognition (see Carruthers, 2017).

[6] Alter and Oppenheimer (2009) review a vast amount of experimental work showing that processing fluency influences people's judgments on an extensive range of social domains. For instance, people's truth judgments seem to be influenced by how the evaluated statements are presented (font size and type, color contrast, background contrast etc.). In a study by Reber and Schwarz (1999; apud Alter and Oppenheimer 2009, 227), participants judge statements like "Lima is in Peru" against varying text contrasts on a computer screen. The participants are more likely to believe that such statements are true when presented in easy-to-read colors on white background (e.g., dark blue) compared to harder-to-read colors on white background (e.g., light blue). Similar cues are used in preference, confidence, or frequency judgments (see Reber et al., 2004; Alter et al., 2007; Koriat, 1993; Reber & Zupanek, 2002). The upshot of this literature is that our epistemic judgments are often based on cues such as fluency or ease of processing, without us explicitly processing them as such, or even without us realizing that this is the case.

feelings), where the feeling of certainty signals that there is no need for further inquiry, given the available evidence, while the feeling of uncertainty signals that further inquiry is needed. This epistemic evaluation will then inform the organism whether it must deploy more cognitive effort to settle the issue at hand, or not. One way to cash out this intuition (using Jesse Prinz's theory of emotions; see Prinz, 2004; see also Dokic, 2012, pp. 307–309) is to think of the feelings of (un)certainty as affective states that monitor the organism's internal mental states and processes, paying attention to cues like fluency or ease of processing, in order to represent *core relational themes or concerns*, in this case, pertaining to whether the organism needs to continue inquiry or stop it. The role of the feelings of (un)certainty, therefore, is to check whether everything is alright, from an epistemic point of view, with the belief-forming mechanisms by which the subject arrives at the target belief, say the belief that *p*. If the subject feels uncertain that *p*, this signals that more inquiry is needed—and so further cognitive effort needs to be deployed. If the subject feels certain that *p*, this signals that it is epistemically fitting to stop inquiry, given the available evidence.

This model presupposes procedural metacognition (the subject monitoring cues like ease of processing or fluency) but does not always require meta-representations. The organism merely checks its internal states to determine whether further inquiry is needed or not. This aligns with the non-meta-representational account. Granted, in organisms capable of meta-representations (e.g., non-infant humans), the mechanism sketched above might also involve the representation of the belief that *p* as either epistemically fitting or as epistemically unfitting, pending on the available evidence, which would amount to a meta-representation (a representation of the representation of *p*). This capacity of self-reflection might account for important differences between humans and less developed organisms, but feelings of (un)certainty can be experienced even without presupposing the existence of meta-representations. In both cases, the appraisals eliciting the feelings of (un)certainty are almost always automatic and happen at a sub-personal level (Dokic, 2012; Koriat, 2012; Proust, 2013).

Having sketched a toy metacognitive account of the feelings of (un)certainty, it is now time to contrast this account with a non-metacognitive, probability-based account inspired by Peter Carruthers and colleagues.[7]

[7] I focus on Carruthers because his explicit argument against the metacognitive view of epistemic feelings (epistemic emotions, in his terms) involves the most developed account on the issue (see especially Carruthers, 2017). That being said, I do not aim here to target Carruthers' particular views on the issue (his views appear to have evolved over time), but to use his remarks as an inspiration for further probability-based accounts that I then discuss for their own merits.

3 The Probability Account

Carruthers and his collaborators purport to explain the epistemic feelings of both human and nonhuman animals without appealing to metacognitive concepts (Carruthers, 2008). The feelings of (un)certainty, in their framework, are to be explained solely by appeal to probability estimates. In what follows, I mention some of the remarks made by Carruthers on this matter, and then differentiate between three possible probability-based accounts of the feelings of (un)certainty that are inspired by these remarks.

Carruthers distinguishes between two forms of the feeling of uncertainty: feeling uncertain of some fact and feeling uncertain of one's ability to complete an action (Carruthers, 2017, p. 70).[8] He focuses on the second notion, assuming that this is the more basic of the two (Carruthers, 2008; Carruthers & Ritchie, 2012; Carruthers & Williams, 2019, 2022).[9] In this paper, however, I am interested in the first notion: what it means to feel certain or uncertain that a proposition p is true. I shall therefore extrapolate from what Carruthers says about the feeling of uncertainty in one's ability to complete an action to what that might suggest about feeling (un)certain about the truth of a proposition.

In an earlier paper, Carruthers and Ritchie (2012) differentiate between uncertainty, which they take to be a fundamentally cognitive state, and the *feeling* of uncertainty, which is an affective state. On this account, to be uncertain that p is to have a low degree of belief that p, while to be certain that p is to have a high degree of belief that p (Carruthers and Ritchie 2012, p. 81). The feeling of uncertainty regarding a particular action is the affective state that accompanies a low degree of belief in the truth of the proposition on which that action depends and is elicited by the appraisal that the target action is unlikely to succeed. As Carruthers and Ritchie put it (2012, p. 85):

> when an animal has a low degree of belief in something (that the pattern on the screen is dense rather than sparse, or that it has just touched the longest of the lines on the screen, for example), then actions that depend upon the truth of that belief will be appraised as unlikely to succeed.

[8] I take it that the former is an affective state about the truth of a proposition (e.g., whether it is true that it rains outside or not), while the latter is an affective state about the success of an action (e.g., whether I am capable of jumping from this cliff to the other one).

[9] His argument for its basicness is that not every affective state that is elicited by a failure to access knowledge is a feeling of uncertainty. Curiosity or surprise are such states. The awareness of failure to access knowledge constitutes a feeling of uncertainty only if "there is some task whose successful execution depends on the fact in question" (Carruthers, 2017, p. 70). I have two comments. First, it is unclear how Carruthers would account for feeling uncertain about things that happened in the past. Second, it seems that we could simply translate talk of completion of action into terms of propositions describing the completion of action. For instance, to feel uncertain of one's ability to make a jump simply means that one feels uncertain *that* one will be able to take the jump. In any case, I do not wish to argue here for which feeling of uncertainty is more basic. I only take Carruthers' remarks as an inspiration for a possible account of the propositional feeling of uncertainty.

In a more recent paper (2017, p. 70), Carruthers defines the feeling of uncertainty as:

> the affective state that arises when one appraises an action as unlikely to succeed. (Or at any rate, as not likely *enough* to succeed, given what is at stake and one's own tolerance for risk.)

Importantly, the appraisal that elicits the feeling of uncertainty concerns the world directly (how likely it is that the target action will succeed) and not a lower-order cognitive state. That is, it does not involve a meta-representation or appeal to another cognitive state, but merely an estimation of evidential probability.[10] Of course, this view does not deny that human beings can be metacognitively aware of their feelings of (un)certainty (see Carruthers, 2017, p. 60). What it emphasizes is that the feelings of (un)certainty are not essentially metacognitive.

Let us now consider how these remarks might inform an account of the feelings of (un)certainty about a proposition. I distinguish between three possible versions of such a probability account: a simple probability account, a weighted probability account, and a shifty-threshold probability account.[11]

According to the simple probability account (suggested by the first quote; Carruthers & Ritchie, 2012, p. 85), *S* feels certain that *p* only if *S* has a high degree of belief in *p* and therefore appraises that *p* is likely to be true. Otherwise, if *S* feels uncertain that *p*, then *S* has a low degree of belief in *p* and thus appraises that *p* is unlikely to be true. I take it that, in this framework, the feelings of (un)certainty involve some kind of threshold of likelihood, such that the subject *S* experiences a feeling of certainty in *p* only if *S*'s probability estimate of *p* being true passes a threshold of likelihood that corresponds to the feeling of certainty. If *S*'s probability estimate is beneath this threshold (call it the *threshold of certainty*), then *S* feels uncertain that *p*.

But what about the influence of non-epistemic factors (like high stakes or intolerance of risk) mentioned in the second quote (Carruthers, 2017, p. 70)? Here, I take that we can have two interpretations, which would correspond to two versions of the probability account: a *weighted probability* account and a *shifty-threshold probability* account.

The first interpretation of Carruthers' (2017, p. 70) quote is that, when forming a feeling of (un)certainty, the subject has a *weighted* probability estimation of the likelihood of the target proposition, such that the appraisal of the likelihood of the target

[10] I assume that the concept of probability that Carruthers has in mind here is evidential probability. Evidential probability is the measurement of how likely it is that *p* is true given the subject's evidence for and against *p*. This contrasts with objective understandings of probability in terms of frequency or propensity, and is usually cashed out in terms of an ideal agent's level of belief, credence, or confidence, given the available evidence on the issue (see Ramsey, 1931; Eder, 2023; see also Williamson, 2000, Ch. 10, for a slightly different understanding of evidential probability).

[11] Ultimately, it is not clear which one Carruthers would prefer, given that his remarks seem to allude to all of them. My intention here is not to reconstruct what Carruthers himself would have to say about the feelings of (un)certainty in a proposition, but rather to sketch probability accounts that I shall analyze for their own merits.

proposition is weighted by non-epistemic factors from the subject's context (such as the stakes for getting *p* right, *S*'s tolerance for risk, time pressure etc.; See Kahneman et al., 1982, on the effects of non-epistemic factors—e.g., information availability—on probability estimates). For instance, if a lot is at stake with regard to *p*, then the subject might have a lower probability estimate of *p*, based on the available evidence, than if nothing was at stake with regard to *p* (see Kruglanski, 1989, on the effects of the need to avoid closure, elicited by higher stakes, on confidence levels). In such a situation, the subject might feel uncertain about whether *p*, because the subject's decreased estimation (due to the influence of high stakes) does not meet the threshold of certainty. Note that, under this reading, the threshold of certainty can remain fixed in all contexts. It is just that, in some contexts, non-epistemic factors might bias the purely epistemic evaluation of whether the subject's evidence meet the fixed threshold of likelihood corresponding to the feeling of certainty. We can call this the *weighted probability* account.[12]

The second interpretation of Carruthers' (2017, p. 70) quote is that the threshold of certainty is not fixed in all contexts, but changes with non-epistemic factors. This account claims that a feeling of uncertainty in *p* need not be associated with a low probability estimate of *p*—*S* could feel uncertain that *p* even if *S* thinks that *p* is very likely to be true. *S* will feel uncertain that *p* only if *S* appraises that *p* is not likely enough, where the threshold for "likely enough" is determined by non-epistemic factors from *S*'s context. In this framework, the appraisals that produce the feelings of (un)certainty do not concern merely the evidential probability of *p*, but also whether *p* is probable *enough*, where this threshold is determined by a complex formula that takes into consideration contextual factors such as one's stakes regarding *p*, the vividness with which one thinks of possibilities of error, one's own character traits, time pressure, and so on. We can call this the *shifty-threshold* probability account.

In all these probability accounts, the appraisals that elicit the feelings of (un)certainty in a proposition *p* concern the evidential probability of *p*'s truth given the evidence available to the subject. This contrasts with the metacognitive account, in which the appraisals that elicit the feelings of (un)certainty concern the evaluation of one's internal states in order to determine whether the subject should continue inquiry or not. One way to understand this distinction is by contrasting two scales that can be used for evaluating one's epistemic position with regard to a proposition *p*. Consider the metacognitive account of the feelings of (un)certainty sketched in Sect. 2. According to that account, the appraisals that elicit the feelings of (un)certainty do not concern the evidential probability of the target proposition, but how fitting it is to stop inquiry on the matter. We can think of the scales of epistemic fittingness and evidential probability as two ways to measure a subject's epistemic position with regard to a target proposition.[13] While the scale of epistemic

[12] I thank an anonymous reviewer of this journal for pushing me to spell out this interpretation.

[13] A useful analogy might be to think of how electricity can be measured either by looking at voltage, in which case it is measured in volts, or by looking at resistance, in which case it is measured in ohms. The two measurements measure different aspects of the same phenomenon. I develop at greater length the comparison between the scales of epistemic fittingness and that of evidential probability in Spatan, 2022. See Deigan & Piñeros Glasscock, 2023; Howard, 2018, for clarifying discussions about the concept of fittingness.

fittingness measures how appropriate it is to stop inquiry and believe that *p* given the evidence, the scale of evidential probability measures how likely it is that *p* is given the evidence. Both are about the epistemic position one finds oneself in, but they concern two critically different aspects of that construct: epistemic fittingness concerns the status of one's internal states and is therefore oriented toward the self, while evidential probability concerns the status of *p* itself and is therefore oriented toward the world. While the two types of evaluation should correlate, given that they measure the same construct, in practice they often do not, as we shall discuss below.

In what follows, I advance arguments against each version of the probability account sketched here. I must emphasize that the ensuing arguments do not show that a non-metacognitive account of the feelings of (un)certainty would be indefensible, but merely that the above versions of the probability account are unsatisfactory, at least when it comes to humans capable of self-reflection. The argument from variability (Sect. 4) poses a problem for the simple probability and the weighted probability accounts, while the argument from epistemicity (Sect. 5) poses a problem for the shifty-threshold probability account.

4 The Argument from Variability

Put simply, the problem with the simple probability account is that human feelings of (un)certainty[14] seem to vary in ways that their probability estimates on the matter do not. Let us investigate two cases that illustrate the point. Consider first REINFECTION, wherein two people have opposing epistemic feelings regarding *p* even though they hold *p* to be equally likely to be true:

> REINFECTION. Sarah and Hannah are both medical doctors who believe, based on very good scientific evidence, that the chance of getting disease D twice is 0.001%. Both Sarah and Hannah have had D before, so they both come to believe that it is 99.999% likely that they will not be reinfected. For Sarah, getting reinfected would be inconsequential, but Hannah has medical preconditions that mean that getting D again would be fatal. Thus, while Sarah feels certain that she is safe from reinfection, given the extremely low chances of this happening, Hannah does not feel at all certain that she is safe from reinfection, despite her awareness of the odds. The stakes are too high for her (the risk of getting reinfected is all too vivid in Hannah's mind), so, unlike Sarah, Hannah takes all available measures to protect herself against reinfection. Doing otherwise would seem irrational to her.

REINFECTION is problematic for the simple probability account for two reasons. First, it demonstrates that some people will feel distinctly uncertain of a proposition

[14] I do not discuss here non-human's or infant's feelings of (un)certainty, because it is not clear if such organisms are capable of self-reflection. The point developed here is that, in the case of those beings capable of self-reflection, it is implausible that the probability accounts explain differences in feelings of (un)certainty.

p even if they have a very high probability estimation of its truth. Second, it shows that one cannot predict a feeling of certainty versus uncertainty using only probability estimates, given that two people can have opposing epistemic feelings regarding a proposition *p* despite holding the same probability estimate regarding *p*.

The problem generalizes, because one and the same person can hold opposing epistemic feelings toward two propositions even if that person evaluates them as equally likely to be true. Consider SIGNATURE:

> SIGNATURE. Clotilde reads an encyclopedia entry on statistics and learns that the probability of writing your signature exactly the same twice is 0.001%. Amused by this information, she feels certain that she will never write her signature the same way twice. Later that day, Clotilde buys a lottery ticket in a fair lottery with 100,000 participants. She does the calculations and realizes that the chance of winning the lottery is 0.001%. Understanding the odds, she nevertheless buys the ticket. At the end of the day, she reasons, somebody has to win the lottery, and it might as well be her. She cannot be sure that she will not win, so she takes her chances.

Let *p* be the proposition that "Clotilde will not write her signature perfectly the same twice" and *q* be the proposition that "Clotilde will not win the lottery." Clotilde knows that *p* and *q* are equally probable. Nevertheless, she feels certain that *p* but does not feel certain that *q*. Whatever the explanation for the distinction between her surety in *p* versus in *q* is,[15] the mere possibility that one might have opposing epistemic feelings concerning two propositions that one estimates as equally likely to be true is sufficient to raise a significant problem for the simple probability account.

Moreover, this type of variability poses a problem also for the weighted probability account. While the weighted probability account *does* take into consideration the influence of non-epistemic factors on our appraises of epistemic likelihood (e.g., it sometimes happens that high stakes lower our probability estimations below the threshold of certainty), the way REINFECTION and SIGNATURE are construed is such that the subject's probability estimates remain *the same* in parallel cases in which the subject experiences opposing epistemic feelings. Therefore, whatever weighting of non-epistemic factors might have occurred, the opposing epistemic feelings cannot be explained by this weighting alone, given that the subjects' likelihood estimates are the same.[16]

[15] The explanation is presumably connected to the nature of lotteries. When you buy a ticket at a fair lottery, you inadvertently think very vividly about the remote chance of winning—you make plans, you imagine yourself having all that money etc. On the other hand, writing your signature perfectly the same twice has no serious stakes attached to it, nor is it usually taken seriously by anyone. The possibility of writing your signature the same twice is so remote and so unimportant that it makes no sense to spend any time on the idea.

[16] At this point, the supporter of the weighted probability account might respond that the description of the cases is misconstrued. For instance, in REINFECTION, while Hannah overtly accepts the same objective probabilities (about the likelihood of getting reinfected) as Sarah does, Hannah's evidential probability estimates are nevertheless (irrationally) lower than Sarah's—below the threshold of certainty. The problem is that we can still imagine perfectly rational agents holding extremely high probability estimations (even higher than Sarah or Hannah) and still feel uncertain about whether *p* (e.g., because of extremely high stakes). This does not square well with a fixed-threshold construal of the probability account (as

Note furthermore that the argument from variability poses no problem for the metacognitive account, because the metacognitive account does not claim that feelings of (un)certainty are elicited by first-order appraisals of evidential probability. Sarah and Hannah can experience opposing feelings of (un)certainty despite sharing the same probability estimates simply because their metacognitive appraisals differ (Sarah thinks that it is fitting to stop inquiry with regard to the target proposition, while Hannah does not). Similarly, Clotilde feels certain that p but uncertain that q, despite appraising p and q as equally probable, because her metacognitive evaluations diverge.

Finally, although the argument from variability poses problems for the simple probability and the weighted probability accounts, it does not pose a problem for the shifty-threshold probability account. According to the latter, for S to feel uncertain that p it must be that S appraises that p is not sufficiently likely given some non-epistemic factors from S's context, such as the stakes for getting p right or S's tolerance for risk. According to this account, Sarah feels certain that she will not get reinfected with the disease D because she appraises that a probability of 99.999% meets her threshold for the feeling of certainty, while Hannah appraises that the same probability does not meet her threshold for the feeling of certainty, given Hannah's certain death if she is reinfected. Similarly, Clotilde appraises that a probability of 99.999% meets her threshold of certainty that she will not write her signature perfectly the same twice, but it does not meet her threshold of certainty that she will not win the lottery. She therefore feels certain about the former proposition and uncertain about the latter. Nonetheless, the shifty-threshold probability account faces issues of its own.

5 The Argument from Epistemicity

In this section, I show that the way in which the shifty-threshold probability account overcomes the challenge from variability creates its own problems. If feelings of (un)certainty are determined not only by appraisals of evidential probability, but also by considerations about stakes, vividness of errors, time pressure, and so on, then these feelings are no longer purely epistemic, but also practical. However, this elides an essential feature of epistemic feelings—namely, their epistemicity.

First, let us distinguish between *purely epistemic evaluations* and *mixed epistemic evaluations*. This distinction concerns whether the construct that is to be evaluated contains only epistemic elements (evidence, truth-conducing factors, etc.), or whether it contains both epistemic elements and non-epistemic elements (stakes, vividness of error possibilities, desires, time pressure, etc.). In mixed epistemic evaluations, the non-epistemic elements are part of the construct that is being evaluated,

Footnote 16 (continued)
assumed in the weighted probability account), given that the lack of the feeling of certainty would not be explained by diminished probability estimates—the probability estimates with such an agent would remain *very* high.

which makes these non-epistemic elements *constitutive* of their evaluation. In purely epistemic evaluations, on the other hand, any non-epistemic elements are not constitutive of the evaluation. While these elements might exert an external influence on the evaluation—biasing the evaluator, as it were—they are not themselves part of the construct that is being evaluated.

An analogy might be helpful here. There have been many studies showing that the physical attractiveness of defendants influences criminal jurors' and judges' assessments of guilt in court (DeSantis & Kayson, 1997; Downs & Lyons, 1991; Kramer et al., 2023; Stewart, 1980). Does this mean that the jurors make their evaluations of guilt based, at least partly, on explicit considerations regarding the defendant's attractiveness? Presumably not. Attractiveness is not an essential part of what constitutes guilt or lack thereof. Thus, considerations of a defendant's attractiveness are *not* constitutive of the jurors' evaluations, but merely an influencing factor on those evaluations (a sort of 'beauty bias'). On the other hand, the attractiveness of contestants in a beauty pageant *is* a constitutive element of the pageant jury's evaluation, because the construct that this jury evaluates is at least partly constituted by attractiveness (alongside, presumably, self-confidence, social skills etc.).

We can draw the analogy to the epistemic case. Just as the juror's judgment is not explicitly about non-judicial factors (e.g., attractiveness), the purely epistemic judgment is not explicitly about non-epistemic factors (e.g., high stakes). Whatever influence non-epistemic factors might have on purely epistemic evaluations, this is not constitutive of the purely epistemic evaluation per se, just as any influence non-judicial factors might have on jurors, this is not constitutive of the jurors' judgments of guilt per se. At best, these are influencing factors, but they are not constitutive of the respective judgments. On the other hand, beauty pageant judgments are *also* about non-judicial considerations like attractiveness, just as mixed epistemic evaluations are *also* about non-epistemic considerations like high stakes. In mixed epistemic evaluations (e.g., assessments in which beliefs and preferences are explicitly weighted to arrive at a decision; see classical decision theory, Savage, 1954; Jeffrey, 1965), the non-epistemic considerations are constitutive of the respective evaluations and are not merely an influencing factor on them.

Another way of understanding the distinction is by looking at whether the evaluator is *rational* in making the non-epistemic factors *explicit*. If they are, then their evaluation is mixed epistemic. If they are not, then their evaluation is purely epistemic. Think again of the attractiveness analogy. In the criminal case, the jurors would be considered irrational, were they to make the non-judicial factors explicit. Attractiveness is not the object of their evaluation, and hence should not be taken into consideration. In the beauty pageant case, on the other hand, the judges would be perfectly rational to explicitly take attractiveness into consideration when making their decisions. Similarly, in a purely epistemic evaluation, making explicit the non-epistemic factors would be considered irrational, given that those factors are not the object of the evaluation, and should thus not be taken into consideration. In a mixed epistemic evaluation, on the other hand, it is perfectly rational to explicitly take non-epistemic factors into consideration.

In what follows, I argue that the metacognitive account is better suited than the shifty-threshold probability account in accommodating the fact that the feelings

of (un)certainty are elicited by purely epistemic evaluations. My argument is the following:

(P1) The threshold probability account presupposes that the feelings of (un)certainty are elicited by mixed epistemic evaluations.
(P2) It is implausible that the feelings of (un)certainty are elicited by mixed epistemic evaluations.
(P3) The metacognitive account does not make this presupposition.
(C) Therefore, the metacognitive account is more plausible than the shifty-threshold probability account in explaining the feelings of (un)certainty.

Let us begin with claim (P1). According to the shifty-threshold probability account, *S* feels certain that *p* only if *S* appraises that *p* is likely enough to be true, given *S*'s evidence with regard to *p* and *S*'s non-epistemic context. For *S* to feel uncertain that *p*, *S* must appraise that *p* is not likely enough to be true, given *S*'s evidence with regard to *p* and *S*'s non-epistemic context. It is apparent from these definitions that the appraisals that elicit the feelings of (un)certainty contain both epistemic elements (the likelihood of *p* given the evidence available to *S*) and non-epistemic elements (the non-epistemic context that determines *S*'s threshold of certainty). The threshold of certainty changes with the subject's non-epistemic context, such that every time the subject evaluates her own epistemic position, they evaluate it in relation to their non-epistemic context. The subject's evaluation is not of their epistemic position alone, but of ther epistemic position relative to their non-epistemic context. The non-epistemic context is therefore part of the construct that is assessed; that is, it is constitutive of the evaluation that elicits the feeling of (un) certainty. The shifty-threshold probability account thus presupposes that the evaluations eliciting the feelings of (un)certainty are mixed epistemic evaluations.

Note that the simple probability or the weighted probability accounts need not presuppose mixed epistemic evaluations in explaining the emergence of the feelings of (un)certainty. According to both accounts, the threshold of certainty is always fixed at the same level (corresponding to a very high evidential probability). Granted, according to the weighted probability account, the epistemic evaluation of evidential probability is sometimes weighted by non-epistemic factors, but this weighting need not be constitutive of the respective epistemic evaluation. Just like in the case of a juror's judgment of guilt being influenced by considerations of attractiveness (weighting thus on the final assessment of guilt), without the judgment becoming explicitly about attractiveness, a subject's evaluation of evidential probability can be influenced by practical considerations of stakes (weighting thus on the final assessment of evidential probability), without the evaluation becoming explicitly about non-epistemic considerations.

This move is not available to the shifty-threshold probability account, given that, according to this account, setting the threshold of certainty is an integral part of the epistemic evaluation leading to feelings of (un)certainty, and this threshold is directly dependent on the subject's assessment of non-epistemic factors. According to this account, when I evaluate whether it is appropriate to feel certain that *p*, I inadvertently adjust the threshold of certainty based on the non-epistemic factors

from my context. Thinking about these factors is a constitutive part of my evaluation. In the fixed-threshold model, on the other hand, there is no adjustment in the threshold of certainty. If non-epistemic factors affect my evaluations, they do so from the outside, so to say, pushing me to arrive at a desired judgment, but without being a constitutive part of that judgment.

Let us move to claim (P2), that it is implausible that the feelings of (un)certainty are elicited by mixed epistemic evaluations. I argue that the evaluations that elicit the feelings of (un)certainty are purely epistemic, and therefore that the non-epistemic factors that sometimes influence these evaluations are not also constitutive of them. To see this, consider first FLAT EARTH:

> FLAT EARTH. A flat earth terrorist holds you at gunpoint and asks you to feel certain that the earth is flat; otherwise, he will shoot you (let us suppose that he has the technical resources to check whether you actually feel certain or not that the earth is flat).

If the feelings of (un)certainty were actually elicited by mixed epistemic evaluations, it would appear natural to fix the threshold of certainty at an extremely low level. Thus, even a probability of, say, 1% should meet the threshold for feeling certain that the earth is flat.

But this is unlikely to be the case. It seems flatly wrong (sic!) to allow that an epistemic probability this low could constitute an epistemically sufficient condition for feeling certain of something. Furthermore, even if you thought that the flat earth hypothesis is more than 90% likely to be true, it would be unlikely that you would simply accept that non-epistemic considerations are very important and thus find a less than ideal probability estimate sufficient to feel certain on the matter. Instead, these non-epistemic considerations would prompt you to engage in a process of further evidence-gathering by which to convince yourself that it is actually *epistemically* fitting to feel certain that the earth is flat. While your non-epistemic situation would surely trigger this evidence-gathering process, it is the epistemic closure that is ultimately necessary for surety. For you to feel certain that the earth is flat, it seems that you must actively try to convince yourself on purely epistemic grounds that the earth is actually flat, and that there are essentially no doubts on the issue. In the absence of a positive and purely epistemic appraisal, you will not experience the feeling of certainty per se.

A supporter of the probability account might amend the account to say that non-epistemic considerations can only raise, never lower, the threshold of certainty. This would make it virtually impossible for you to feel certain that the earth is flat as a consequence of a terrorist's threat; you could feel certain only if that threat made you believe that it is (close to) 100% likely that the earth is flat—in other words, only if the threat changed your epistemic condition. While I find this response problematic on separate grounds,[17] it still fails to accommodate the epistemicity of the

[17] First, it rules out instances of people feeling very confident or even certain that p despite their awareness of the low odds of p, such as optimistic lottery players. Second, its defense requires theoretically compelling and non-ad hoc reasons why non-epistemic considerations can only increase one's threshold for certainty. There is also counterevidence; for instance, increasing time pressure has been shown to

feelings of (un)certainty. To see this, consider another case, adapted from Worsnip (2021, p. 2), TWO HANDS, in which the non-epistemic factors are supposed to raise the threshold for certainty rather than lower it:

> TWO HANDS. I have two hands and extremely good evidence to believe that I have two hands (I see them, I feel them, I even taste them and talk to others about them). But you, an eccentric millionaire, offer me a bribe: if I feel uncertain that I have two hands, you will give me $1,000,000 (suppose that you have the technical resources to check whether I feel uncertain or not that I have two hands).

If the feelings of (un)certainty were actually elicited by mixed epistemic evaluations, then even on the modified probability account it would be natural for me to increase the threshold for certainty to an extremely high level, such that the evidence that I currently have is insufficient for feeling certain.

But that seems unlikely to happen here. When presented with the bribe of $1,000,000, I am not immediately able to appraise that it is insufficiently likely that I have hands. To actually feel uncertain about having hands, I must convince myself that there are epistemic grounds for doubt. Perhaps I think about *The Matrix*, Putnam's brains in vats, or Descartes' evil demon. I shall attempt to find some epistemic grounds to find my current evidence insufficiently strong to guarantee that I have hands. But until I have a real, active doubt on the matter, I will not feel uncertain that I have hands.[18] This epistemic searching is initiated by the bribe, but the bribe itself is not a constitutive part of my evaluation. If I do end up feeling uncertain that I have hands, it will not be because I think that my evidence is insufficient from a practical point of view, 'given what is at stake'; it must be insufficient from a purely epistemic point of view.

Thus, the evaluations that elicit the feelings of (un)certainty are purely epistemic. Non-epistemic factors can influence these appraisals; in all the cases discussed so far, non-epistemic factors put pressure on the subject to appraise their epistemic position one way or another. However, these non-epistemic factors are never constitutive of the evaluations that elicit the feelings of (un)certainty. By defining the feelings of (un)certainty as mixed epistemic-practical attitudes, the shifty-threshold probability account fails to account for their *epistemicity*.[19]

Footnote 17 (continued)
lower the certainty threshold, whereas increasing stakes raises it (see Kruglanski, 1989; Kruglanski & Webster, 1996).

[18] As Charles Peirce famously phrased it, "the mere putting of a proposition into the interrogative form does not stimulate the mind to any struggle after belief. There must be a real and living doubt, and without this all discussion is idle" (1877, p. 6).

[19] An alternative probability account that would not require mixed epistemic evaluations is one that posits that the threshold of certainty is always probability 1 (I thank a reviewer of this journal for pushing me to discuss this option). Given that the underlying assessment (of whether the evidential probability of p is 1) is a purely epistemic evaluation, this account would avoid the challenge from epistemicity. The downside of this view is the entailment that we could rarely feel certain about anything (given that we are rarely in an infallible epistemic position, and we acknowledge that). To meet this challenge, we might assume that the feeling of certainty comes in degrees, such that, when it comes to propositions that we ascribe a very high probability to, we feel *almost* certain that they are true. One worry with this option

Finally, let us move to claim (P3), according to which the metacognitive account does not presuppose mixed epistemic evaluations. Consider the toy metacognitive account sketched in Sect. 2. According to that account, the role of the feelings of (un)certainty is to determine whether or not the subject needs more evidence on the matter, or the current evidence is sufficient to settle the matter. This is a purely epistemic evaluation: it concerns solely the epistemic fittingness of stopping or not one's inquiry, given the evidence that is currently available to one. Granted, non-epistemic considerations can influence these metacognitive evaluations, but they are *not constitutive* of those evaluations.[20]

Returning to FLAT EARTH, if a flat earth terrorist successfully compels you to feel certain that the earth is flat (if that can ever happen), it will not be because you lower your probability threshold for certainty, but because your fear of the threat pushes you to judge that it is fully epistemically fitting to stop inquiry and believe that the earth is flat.[21] This happens only when you lack any real, epistemic doubts that the earth is flat. The external non-epistemic context pushes you to feel certain that the earth is flat, but it is not a constitutive part of your appraisal.

Similarly, in TWO HANDS, I would not be able to feel uncertain about the existence of my hands by increasing my probability threshold for certainty, but only because the thought of winning $1,000,000 would push me to find epistemic grounds for doubt, and so to judge that it is not fully epistemically fitting to believe that I have hands. In both cases, the appraisals that elicit the feelings of (un)certainty are purely epistemic, in that they are solely concerned with the epistemic fittingness of stopping inquiry and settling on the target belief. The non-epistemic features of one's context are not constitutive of one's epistemic fittingness evaluations; they are only influencing factors, biasing one in one direction or another.

Note that the probability account does not have the resources for an analogous move. The crucial point of departure between the metacognitive account and the probability account is that the former makes use of the fact that there are two scales, evidential probability of propositions, and epistemic fittingness of beliefs, by which the epistemic position of a subject can be evaluated. It is the latter scale that is used in the formation of the feelings of (un)certainty. Of course, for a perfectly rational agent, the two scales will align. In practice, though, non-epistemic factors can influence the two measurements differently, as seen in all the examples presented above. The probability account uses a single scale for measuring an agent's epistemic

Footnote 19 (continued)
is that it does not deal well with the challenge from variability. Suppose, for instance, that I estimate that p (the proposition that I cannot write my signature the same twice) has a probability of 99.999, and, because nothing hinges with regard to p, I feel distinctively certain that p. At the same time, despite estimating that q (the proposition that I will lose the lottery) has a much higher probability, say, 99.999999, I might feel much less certain of q than of p.

[20] Just like attractiveness can sometimes be an influencing factor in judges' or jurors' decisions of guilt (Kramer et al., 2023), implicitly biasing them, without being explicitly part of the rationale that is given for those decisions (i.e., without being constitutive of the judgments leading to those decisions).

[21] The existence of overconfident optimist lottery players demonstrates that people can hold extremely low probability estimates but nevertheless feel confident—sometimes even certain—that they will win.

position, and therefore must postulate a moving probability threshold for certainty that is determined by non-epistemic features. The metacognitive account does not need to postulate such a thing. The appraisals responsible for eliciting the feelings of (un)certainty are purely epistemic (concerning solely the epistemic fittingness of inquiry), but they can be associated with different assessments of evidential probability, therefore accommodating both the variability cases (REINFECTION and SIGNATURE), and the epistemicity of the feelings of (un)certainty.

At this point, the proponents of the shifty-threshold view might push back and argue for a version of the shifty-threshold probability account that delineates between two phases of assessment, in the hope of meeting the challenge from epistemicity.[22] At the first phase, the story would go, the subject ascertains the threshold of certainty, considering factors such as stakes or tolerance of risk, and arrives at a numerical value, let us say 0.95. Subsequently, during the second phase, the subject evaluates the evidential probability of proposition p given the evidence for p, in order to determine whether that probability meets the threshold established in the first phase. It appears that during this second phase, only purely epistemic factors are in play, such that the appraisals eliciting the feelings of (un)certainty are purely epistemic.

Nevertheless, this version faces some challenges, too. Firstly, it assumes a fairly complex picture of the way people form feelings of (un)certainty; a picture in which subjects first determine a threshold of certainty necessary for the feeling of certainty, and then ascertain whether the current evidence meets that threshold or not. One might wonder how well this perspective accommodates the phenomenology of the feelings of (un)certainty (e.g., the fact that these feelings seem automatic and are based on cues like fluency or ease of processing; see Dokic, 2012; Arango-Muñoz, 2014). Secondly, it is not clear if such a view would not ultimately turn out to be metacognitive, given that the aim of the first phase is to determine the threshold of evidential probability required for the feeling of certainty, which presumably involves representing the feeling of certainty and the requirements needed to reach it. Thirdly, even if the second phase concerns only epistemic considerations, the subject could not have arrived at their final judgment without going through the first phase, which involves non-epistemic considerations. Non-epistemic considerations are thus constitutive of the process by which the subject arrives at their feeling of certainty, even if that process is split into two successive phases. This is notably different for the metacognitive account. Non-epistemic considerations are never part of the content that is assessed during the formation of the feelings of (un)certainty. And that is because, in the metacognitive account, the threshold of certainty is never changed relative to non-epistemic factors. The metacognitive account is thus better equipped to accommodate the epistemicity of the feelings of (un)certainty.

[22] I thank an anonymous reviewer of this journal for pushing me to discuss this view.

6 Conclusion

In this paper, I have presented a case against defining the feelings of (un)certainty in probabilistic terms, at least when it comes to self-reflecting humans. Based on work by Carruthers and collaborators, I have identified three probability accounts of the feelings of (un)certainty. I then argued that those accounts that posit a fixed threshold of evidential probability required for certainty have troubles dealing with the challenge from variability, while those accounts that posit a shifty threshold have troubles dealing with the challenge from epistemicity.

The metacognitive account sketched in Sect. 2 allows us to avoid both challenges. On the one hand, the metacognitive account posits two ways of evaluating one's epistemic position with regard to a proposition *p*: the evidential probability of *p* and the epistemic fittingness of believing that *p*. The feelings of (un)certainty are elicited by appraisals concerning the latter measurement, and not the former. This accommodates cases (see REINFECTION and SIGNATURE) in which people have opposing epistemic feelings despite holding the same probability estimates. On the other hand, the metacognitive account posits a fixed threshold of certainty set at the highest level, such that one feels certain that *p* only if one appraises that it is fully fitting to believe that *p*. This avoids the challenge from epistemicity (illustrated by FLAT EARTH and TWO HANDS), given that people's metacognitive appraisals need not take into account non-epistemic considerations (as in the shifty-threshold probability account), but can nevertheless be influenced by them.

Still, the foregoing discussion does not offer a definitive case in favor of the metacognitivist camp regarding a more general discussion of the feelings of (un) certainty, as well as other epistemic feelings, like the tip-of-the-tongue phenomenon, surprise, curiosity, and so on. It is possible that there are other non-metacognitive alternatives for explaining the feelings of (un)certainty in human adults that avoid the issues faced by the two probability accounts. Moreover, I have analyzed only the feelings of (un)certainty as experienced by non-infant humans; that is, in organisms capable of self-reflection. Much more needs to be said about how these feelings relate to those experienced by human infants and nonhuman animals. That said, my arguments do offer good reasons to doubt that these feelings can be defined in probabilistic terms, given that a crucial subclass of these feelings—those experienced by non-infant human—are inexplicable by the most salient probability accounts.

Acknowledgements I thank my PhD supervisors, Thomas Krödel and Moritz Schulz, for their constant support and helpful comments on the margins of this paper, as well as to the participants in the colloquia in theoretical philosophy at the University of Hamburg and Dresden University of Technology, where I presented various versions of the ideas contained in this paper.

Funding Open Access funding enabled and organized by Projekt DEAL. I have benefited in writing this manuscript by being employed in two projects based at the Dresden University of Technology. (1) At the center "Disruption and Societal Change" (TUDiSC), in the project "The Disruptivity of the Others in Urban Transformations". (2) In the project „Gefangen im Netz? Chancen und Perspektiven von epistemischer Netzwerkanalyse", financed by Sächsische Aufbaubank, and based at Technische Universität Dresden.

Conflict of interest The authors declare that they have no conflict of interest.

Ethical Approval This article does not contain any studies with human participants or animals performed by the author.

References

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13*(3), 219–235. https://doi.org/10.1177/1088868309341564

Alter, A. L., Oppenheimer, D. M., Epley, N., & Eyre, R. N. (2007). Overcoming intuition: Metacognitive difficulty activates analytic reasoning. *Journal of Experimental Psychology. General, 136*(4), 569–576. https://doi.org/10.1037/0096-3445.136.4.569

Arango-Muñoz, S. (2014). The nature of epistemic feelings. *Philosophical Psychology, 27*(2), 193–211. https://doi.org/10.1080/09515089.2012.732002

Arango-Muñoz, S., & Michaelian, K. (2014). Epistemic feelings, epistemic emotions: Review and introduction to the focus section. *Philosophical Inquiries, 2*, 97–122.

Balci, F., Freestone, D., & Gallistel, C. R. (2009). Risk assessment in man and mouse. *PNAS Proceedings of the National Academy of Sciences of the United States of America, 106*(7), 2459–2463. https://doi.org/10.1073/pnas.0812709106

Beran, M. (2019). Animal metacognition: A decade of progress, problems, and the development of new prospects. *Animal Behavior and Cognition, 6*, 223–229.

Carruthers, P. (2008). Meta-cognition in animals: A skeptical look. *Mind and Language, 23*(1), 58–89. https://doi.org/10.1111/j.1468-0017.2007.00329.x

Carruthers, P. (2009). How we know our own minds: The relationship between mindreading and metacognition. *The Behavioral and Brain Sciences, 32*(2), 121–138. https://doi.org/10.1017/S0140525X09000545

Carruthers, P. (2017). Are epistemic emotions metacognitive? *Philosophical Psychology, 30*(1–2), 58–78. https://doi.org/10.1080/09515089.2016.1262536

Carruthers, P. (2021). Explicit nonconceptual metacognition. *Philosophical Studies, 178*(7), 2337–2356. https://doi.org/10.1007/s11098-020-01557-1

Carruthers, P., Fletcher, L., & Brendan Ritchie, J. (2012). The evolution of self-knowledge. *Philosophical Topics, 40*(2), 13–37.

Carruthers, P., & Ritchie, J. B. (2012). The emergence of metacognition: Affect and uncertainty in animals. IN EDITED BOOK? IF SO, NAMES OF EDITORS? *Foundations of Metacognition* (pp. 76–93). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199646739.003.0006

Carruthers, P., & Williams, D. M. (2019). Comparative metacognition. *Animal Behavior and Cognition, 6*(4), 278–288. https://doi.org/10.26451/abc.06.04.08.2019

Carruthers, P., & Williams, D. M. (2022). Model-free metacognition. *Cognition, 225*, 105117. https://doi.org/10.1016/j.cognition.2022.105117

de Sousa, R. (2009). Epistemic feelings. *Mind and Matter, 7*, 139–161.

Deigan, M., & Piñeros Glasscock, J. S. (2023). A fitting definition of epistemic emotions. *The Philosophical Quarterly*. VOLUME, PAGE NUMBERS? https://doi.org/10.1093/pq/pqad096

DeSantis, A., & Kayson, W. A. (1997). Defendants' characteristics of attractiveness, race, and sex and sentencing decisions. *Psychological Reports, 81*, 679–683. https://doi.org/10.2466/pr0.1997.81.2.679

Dokic, J. (2012). Seeds of self-knowledge: Noetic feelings and metacognition. EDITORS OF BOOK? *Foundations of metacognition* (pp. 302–321). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199646739.003.0020

Dokic, J. (2014). Feelings of (un) certainty and margins for error. *Philosophical Inquiries, 2*(1), 123–144. https://doi.org/10.4454/philinq.v2i1.80

Downs, A. C., & Lyons, P. M. (1991). Natural observations of the links between attractiveness and initial legal judgments. *Personality and Social Psychology Bulletin, 17*, 541–547. https://doi.org/10.1177/0146167291175009

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition*. Sage Publications Inc.

Eder, A.-M.A. (2023). Evidential probabilities and credences. *The British Journal for the Philosophy of Science, 74*(1), 1–23. https://doi.org/10.1093/bjps/axz043

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911. https://doi.org/10.1037/0003-066X.34.10.906

Foote, A. L., & Crystal, J. D. (2012). 'Play it again': A new method for testing metacognition in animals. *Animal Cognition, 15*(2), 187–199. https://doi.org/10.1007/s10071-011-0445-y

Gallistel, C. R., Mark, T. A., King, A. P., & Latham, P. E. (2001). The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *Journal of Experimental Psychology: Animal Behavior Processes, 27*(4), 354.

Goupil, L., & Proust, J. (2023). Curiosity as a metacognitive feeling. *Cognition, 231*, 105325. https://doi.org/10.2139/ssrn.4141925

Hart, J. T. (1965). Memory and the feeling-of-knowing experience. *Journal of Educational Psychology, 56*, 208–216. https://doi.org/10.1037/h0022263

Howard, C. (2018). Fittingness. *Philosophy Compass, 13*(11), e12542. https://doi.org/10.1111/phc3.12542

Iwasaki, S., Watanabe, S., & Fujita, K. (2018). Pigeons (*Columba livia*) know when they will need hints: Prospective metacognition for reference memory? *Animal Cognition, 21*(2), 207–217. https://doi.org/10.1007/s10071-017-1153-z

Jeffrey, R. C. (1965). *The logic of decision*. University of Chicago Press.

Kahneman, D., Slovic, P., & Tversky, A. (1982). *Uncertainity: Heuristics and biases*. Cambridge University Press.

Koriat, A. (1993). How do we know that we know? The accessibility model of the feeling of knowing. *Psychological Review, 100*(4), 609–639.

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition, 9*(2), 149–171. https://doi.org/10.1006/ccog.2000.0433

Koriat, A. (2012). The self-consistency model of subjective confidence. *Psychological Review, 119*(1), 80–113. https://doi.org/10.1037/a0025648

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology. General, 133*(4), 643–656. https://doi.org/10.1037/0096-3445.133.4.643

Kornell, N. (2014). Where is the 'meta' in animal metacognition? *Journal of Comparative Psychology, 128*(2), 143–149. https://doi.org/10.1037/a0033444

Kramer, R. S. S., Jarvis, J.-L., Green, M., & Jones, A. L. (2023). The relationship between facial attractiveness and perceived guilt across types of crime. *Quarterly Journal of Experimental Psychology*. VOLUME AND PAGE NUMBERS? https://doi.org/10.1177/17470218231218651

Kruglanski, A. W. (1989). *Lay epistemics and human knowledge: cognitive and motivational bases*. Springer US. https://doi.org/10.1007/978-1-4899-0924-4

Kruglanski, A. W., & Webster, D. M. (1996). Motivated closing of the mind: 'Seizing' and 'freezing.' *Psychological Review, 103*(2), 263–283. https://doi.org/10.1037/0033-295x.103.2.263

Meylan, A. (2014). Epistemic emotions: A natural kind? *Philosophical Inquiries, 2*(1), 173–190.

Morgan, G., Kornell, N., Kornblum, T., & Terrace, H. S. (2014). Retrospective and prospective metacognitive judgments in rhesus macaques (Macaca Mulatta). *Animal Cognition, 17*(2), 249–257. https://doi.org/10.1007/s10071-013-0657-4

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *Psychology of learning and motivation* (Vol. 26, pp. 125–73). Academic Press.

Peirce, C. S. (1877). The fixation of belief. *Popular Science Monthly, 12*(1), 1–15.

Perner, J. (2012). MiniMeta: In search of minimal criteria for metacognition. In M. J. Beran, J. Brandl, J. Perner, & J. Proust (Eds.), *Foundations of metacognition. PAGE NUMBERS* Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199646739.003.0007

Prinz, J. J. (2004). *Gut reactions: A perceptual theory of emotion*. Oxford University Press.

Proust, J. (2007). Metacognition and metarepresentation: Is a self-directed theory of mind a precondition for metacognition? *Synthese, 159*(2), 271–295.

Proust, J. (2013). *The philosophy of metacognition: Mental agency and self-awareness*. Oxford University Press.

Proust, J. (2019). From comparative studies to interdisciplinary research on metacognition. *Animal Behavior and Cognition, 6*, 309–328. https://doi.org/10.26451/abc.06.04.10.2019

Rakoczy, H. (2017). In defense of a developmental dogma: Children acquire propositional attitude folk psychology around age 4. *Synthese, 194*(3), 689–707. https://doi.org/10.1007/s11229-015-0860-8

Ramsey, F. P. (1931). Truth and probability. In R. B. Braithwaite (Ed.), *Foundations of mathematics and other essays.* Routledge.

Reber, R., & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition: An International Journal, 8*(3), 338–342. https://doi.org/10.1006/ccog.1999.0386

Reber, R., Schwarz, N., & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience? *Personality and Social Psychology Review, 8*(4), 364–382. https://doi.org/10.1207/s15327957pspr0804_3

Reber, R., & Zupanek, N. (2002). Effects of processing fluency on estimates of probability and frequency. *ETC. Frequency processing and cognition* (pp. 175–188). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198508632.003.0011

Savage, L. J. (1954). *The foundations of statistics*. Wiley Publications in Statistics.

Smith, J. D. (2010). Inaugurating the study of animal metacognition. *International Journal of Comparative Psychology, 23*(3), 401–413.

Smith, J. D., Beran, M. J., Couchman, J. J., & Coutinho, M. V. C. (2008). The comparative study of metacognition: Sharper paradigms, safer inferences. *Psychonomic Bulletin and Review, 15*(4), 679–691. https://doi.org/10.3758/PBR.15.4.679

Smith, J. D., Couchman, J. J., & Beran, M. J. (2014). Animal metacognition: A tale of two comparative psychologies. *Journal of Comparative Psychology, 128*(2), 115–131. https://doi.org/10.1037/a0033105

Smith, J. D., Jackson, B. N., & Church, B. A. (2019). The cognitive architecture of uncertainty. *Animal Behavior and Cognition, 6*(4), 236–246.

Smith, J. D., Shields, W. E., & Washburn, D. A. (2003). The comparative psychology of uncertainty monitoring and metacognition. *Behavioral and Brain Sciences, 26*(3), 317–339. https://doi.org/10.1017/S0140525X03000086

Stewart, J. E. (1980). Defendant's attractiveness as a factor in the outcome of criminal trials: An observational study. *Journal of Applied Social Psychology, 10*, 348–361. https://doi.org/10.1111/j.1559-1816.1980.tb00715.x

Whittlesea, B. W., & Williams, L. D. (2000). The source of feelings of familiarity: The discrepancy-attribution hypothesis. *Journal of Experimental Psychology, Learning, Memory, and Cognition, 26*(3), 547–565.

Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Worsnip, A. (2021). Can pragmatists be moderate? *Philosophy and Phenomenological Research, 102*(3), 531–558. https://doi.org/10.1111/phpr.12673