*Episteme* (2026), 1–19
doi:10.1017/epi.2026.10126

episteme

ARTICLE

# Trust Issues: Reliability, Reference Classes, and the Underdetermination of Trust

William Conner

Department of Philosophy, The University of North Carolina at Chapel Hill
Email: wconner@unc.edu

(Received 11 March 2026 UTC; revised 11 March 2026 UTC; accepted 29 May 2026 UTC)

**Abstract**

This paper concerns the problem of determining whom one should trust and how much in complex testimonial exchanges featuring conflicting reports, unclear communication, and higher-order evidence bearing on the reliability of a speaker. Drawing on a Bayesian model of source reliability, I argue for testimonial underdetermination: the claim that testimonial exchanges can underdetermine whom one ought to trust and how much. The argument proceeds by showing that assessments of a source's reliability are made relative to reference classes – sets of testimonial exchanges sharing features relevant for predicting the source's reliability. In complex cases, an agent's evidence may fail to determine which among several plausible reference classes is most appropriate, where different reference classes underwrite different degrees of trust. I then contend that testimonial underdetermination supports a form of synchronic intrapersonal permissivism about trust: an agent's evidence may permit her to adopt any of several incompatible degrees of trust in a speaker. I conclude that when purely epistemic considerations underdetermine trust, practical considerations may guide one's choice between different epistemically permissible options, putting pressure on the idea that we can have a *pure* epistemology of trust that applies to real-life, complex cases involving non-ideal agents.

**Keywords:** Testimony; trust; underdetermination; permissivism; Bayesian epistemology; formal social epistemology; reference class problem

## 1. Introduction

Imagine: two colleagues in your new department provide clashing testimony about how best to navigate a tricky interpersonal situation; a close friend disputes the trustworthiness of another, perhaps motivated by jealousy; a romantic partner provides testimony you have some antecedent reason to doubt – but then again, you're no expert and feel pulled to give them the benefit of the doubt. Complex testimonial exchanges like these abound. In them, we find ourselves facing the task of interpreting, weighing, and integrating multiple pieces of evidence – sometimes including higher-order evidence –

© The Author(s), 2026. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

some of them conflicting, sometimes with high stakes, and guided (perhaps licitly, but not always) by our preexisting relationships of trust.

One might expect epistemologists of testimony to have developed theories that help us think through these complex cases. But traditional accounts of testimony and testimonial justification have been developed, defended, and debunked on the basis of simple, highly idealized cases in which speakers offer clear, unambiguous testimony to listeners who understand it perfectly and either clearly possess or clearly lack defeaters for believing it. In such cases, little interpretive work is required of hearers, who need only decide whether to believe what they've been told.

While this methodological approach is genuinely valuable in many ways, helping us isolate and test theoretical principles, the results do not always generalize to real-world testimonial exchanges, which are rarely so clean. By abstracting away from the complexity and dynamics of real-world trust and communication, traditional theories leave open important questions about responsibly navigating the complexity of testimonial exchanges in non-ideal settings. As Lackey (2021: 224) puts it, "Rather than theorizing only about the giving and receiving of testimony in ideal situations, more work needs to be done on how recipients of testimony should navigate the world as it actually is." This paper takes up that project, using the tools of Bayesian epistemology to think through decisions about whom to trust and how much.

After laying down conceptual foundations and raising what I call the *complexity challenge* (Section 2), I argue for **testimonial underdetermination**: testimonial exchanges can underdetermine whom a recipient of testimony ought to trust and how much (Section 4.1). To do so, I employ a formal model of source reliability (Section 3) from Olsson (2011; 2013), which helps illustrate the difficulty of making determinations of speakers' reliability in complex cases where the reference class problem looms large. I then contend that testimonial underdetermination supports a kind of synchronic intrapersonal permissivism about trust: in the face of underdetermination, an agent's evidence at a single time may permit her to adopt any of several incompatible degrees of trust in another (Section 4.2). This might suggest that trust decisions are hard to make. Thankfully, I argue, practical considerations may help out, guiding one's choice among one's epistemically permissible options (Section 5). By way of conclusion, I highlight how the resulting approach to testimonial trust opens up new avenues for inquiry about the diachronic aspects of trust and the causes of persistent disagreement and polarization.

## 2. The complexity challenge

As is standard in the literature, I consider trust a three-place relation between a truster, a trustee, and a domain about which the trustee is trusted.[1] In this paper, I focus specifically on *testimonial trust*, which we exhibit when we believe something on another's say-so. The following is paradigmatic: Steve tells Bill that the Seahawks won last night, and Bill takes Steve's word for it.

Typically, we trust speakers across a range of thematically or conceptually related propositions. I trust my doctor, for instance, about a wide range of medical propositions – "Tylenol will help you feel better," "I'm sorry, but we can't save your toe," and

[1]This view is orthodox. See Baier (1986), Jones (1996, 2012, 2019), and Hawley (2014a, 2019) for discussion. For opposing arguments that trust is fundamentally a *two-place* relation ('A trusts B') akin to love and friendship, see Faulkner (2015, 2017) and Domenicucci and Holton (2017).

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

"that drinking's really going to catch up to you someday" – but not about the tree species in my backyard, quantum mechanics, or monetary policy. We'll call a set of propositions that hang together in these ways a *domain*. Sometimes, as with our loved ones, we trust others across a range of domains, not just across the propositions within one domain.

Putting these points together, here is how I understand testimonial trust.

> **Trust**: *A* trusts *B* within domain *d* if and only if *A* is disposed to give high credence to or form outright beliefs in propositions in *d* on *B*'s say-so.[2]

While trust is great, trusting *rationally* is even better.

> **Rational trust**: *A* rationally trusts *B* within domain *d* if and only if (i) *A* trusts *B* within *d* and (ii) *A* has sufficient epistemic reason to regard *B* as sincere and competent within *d*.

Speaker reliability is standardly glossed as domain-relative competence and sincerity, so we can abbreviate **Rational Trust** as follows: *A* rationally trusts *B* as to *d* just in case she has sufficient epistemic reason to regard *B* as reliable about *d*.

With these conceptual foundations in place, let's consider the following case.

> **Colleagues**: Mark and Sarah are colleagues who have developed a friendship. Over several months, they've discussed work regularly. Mark has consistently expressed satisfaction with his job, good rapport with his team, and confidence in his abilities. He has never struck Sarah as insincere or incompetent – he's always forthcoming with his thoughts and reasonable in his self-assessments. So, she trusts Mark about this. A couple of weeks later, however, another colleague, Anna, complains to Sarah about Mark's work performance. "His work is sloppy," Anna says. "And honestly, I'm not sure how much he cares about this job." Sarah is caught off guard: Anna's report doesn't match Mark's own reports at all, but as far as she knows, Anna has always been honest with her. However, Sarah *also* knows that Mark recently received an award for which Anna was in contention, too, so she wonders whether Anna's report stemmed from jealousy toward Mark. That said, Anna doesn't *really* seem like the jealous type . . . does she?

This is a tough case. Sarah's preexisting trust in Mark seems to provide reason to believe what he says, but her preexisting trust in Anna seems to provide reason to believe Anna's countervailing testimony. The weight of Anna's testimony *might be* undercut by higher-order evidence suggesting that she has a reason to be dishonest – but Sarah is unsure how weighty that higher-order evidence is. So, Sarah may reasonably be *stumped* about whether to continue trusting Mark. In this respect, she faces a *trust decision*: a choice

[2]This serves only to pick out the phenomenon I am theorizing about – testimonial trust. I do not claim that it captures the essence of trust *tout court*, nor do I offer it as a general account of trust. Indeed, I am doubtful that there is "one true essence" of trust. For defense of a similar position, see Simpson (2012).

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

between continuing to trust Mark *or* reducing or withdrawing her trust. This is, in essence, a choice between different interpretations of the available evidence. On the one hand, Sarah could view Anna's testimony as an expression of jealousy, allowing her to preserve trust in Mark. Alternatively, she could treat Anna's testimony as a valuable perspective that highlights possible blind spots in Mark's self-assessment, prompting her to reduce her trust in Mark.

Cases with these features are ubiquitous. Consider how often we receive conflicting testimony that requires us to determine which party is correct and whether one party is biased, deceptive, or incompetent: evaluating conflicting news coverage from different outlets, assessing competing reviews of products online, and assessing contradictory accounts of an event from different friends. Think of how often we have higher-order evidence about a source's reliability and must determine its weight – for example, a family member warns that a particular news outlet is biased, but we're uncertain about the soundness of *their* assessment. Take the case of vaccine hesitancy among marginalized communities with reasonable distrust of medical institutions, who must interpret physicians' advice against that historical background (Goldenberg 2016). Finally, consider cases of expert disagreement where competing authorities offer different interpretations of the same evidence (Goldman 2001). In all of these situations, agents face difficult trust decisions like Sarah's.

What these cases have in common is *testimonial complexity*. Roughly, a testimonial exchange is complex to the extent that it involves conflicting reports, ambiguous evidence, unclear communication (think: slurred speech, poor hearing, and crowded rooms), higher-order evidence that casts doubt on the trustworthiness of a speaker, or some combination thereof. Testimonial complexity, then, comes in degrees: some testimonial exchanges are quite straightforward – for example, receiving the morning forecast from a reliable meteorologist – while others, like those mentioned above, are far more complex. While developing a full *account* of testimonial complexity would require saying much more, we have enough detail here for our purposes – namely, thinking through the *upshot* of testimonial complexity. In complex testimonial exchanges, we find ourselves needing to interpret, weigh, and balance various considerations and pieces of evidence that push and pull in different ways. And it may be unclear how to do so.

To the extent that we must regularly grapple with testimonial complexity, we regularly face the following:

> **The complexity challenge**: The challenge of determining whom to trust – and how much – in the face of testimonial complexity.

The reason this is a challenge is that testimonial complexity can induce what Carr (2020) calls *normative uncertainty*: uncertainty about what one's evidence *supports* (cf. Dorst (2019) on *higher-order uncertainty*). More importantly, I will argue below that testimonial complexity can generate full-blown *underdetermination* about whom we should trust and how much.

But complexity isn't just a challenge for *us* – it's also a challenge for traditional approaches to testimony. Epistemologists of testimony are centrally concerned with *the permissibility question*: Under what conditions is one epistemically permitted to believe another's testimony? There are two main camps in this debate. On the one hand, nonreductionists hold that one has defeasible, default epistemic permission to believe a speaker's testimony whenever one understands her testimony and lacks defeaters for believing it (Reid 1764/1997; Burge 1993; Simion 2021). Crucially, this is a *default*

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

entitlement: one need not have any evidence about a speaker's reliability at all in order to enjoy this entitlement. Reductionists, however, argue that one is epistemically permitted to believe a speaker's testimony just in case one has sufficient undefeated non-testimonial positive reasons to regard the speaker as reliable (Hume 1748/1999; Fricker 1994). Other theorists have advanced hybrid views combining elements of these theories (Lackey 2008; Faulkner 2011).

I think that neither reductionism nor nonreductionism is of much assistance in helping us determine how to make trust decisions in the face of complexity. To see this, let's think about how they would advise Sarah. Nonreductionists would hold that Sarah has a defeasible default entitlement to believe Mark *and* a defeasible default entitlement to believe Anna. What she should do is ultimately determined by whether either of these entitlements has in fact been defeated – that is, whether reasons to suspect that Anna is jealous defeat Sarah's entitlement to believe her or whether Anna's testimony defeats Sarah's entitlement to believe Mark. But *the whole problem* is that Sarah does not know which, if either, of these is the case. Reductionists, meanwhile, would argue that Sarah is justified to believe, and thus trust, whichever speaker her non-testimonial evidence on balance supports believing is more reliable. But again, this is exactly what's at issue. I say this not to malign either of these theories; they were developed for different purposes than addressing real-world complexity. Given that our purpose here *is* to address real-world complexity, though, we need a different approach entirely, one that has more promise of offering guidance in situations like Sarah's. Let's turn to this now.

## 3. The model

I'll argue that Sarah's evidence underdetermines what she ought to believe about Mark and thus the degree of trust she should place in him in this domain. If this argument succeeds, it will establish the following:

> **Testimonial underdetermination**: Testimonial exchanges can underdetermine whom a recipient of testimony ought to trust and how much.

Making that argument is the task of the next section. To do so, I will draw on Olsson's (2011, 2013) formal model. The reason to go formal here is that we'll need to think about degrees of confidence and how precisely determinations of a speaker's reliability should shift our opinions and our levels of trust. Formalism, while risky, provides helpful tools for thinking about inherently graded phenomena like these carefully and rigorously. That said, I want to emphasize that the formalism is intended to lay bare a problem we all navigate in testimonial exchanges, if implicitly or subconsciously, and help us theorize about it. I do not intend to suggest that the model perfectly represents explicit thought processes. Moreover, the response I'll develop to the problem does not require the formalism. The model neither creates nor solves the problem; it just makes its structure apparent and easier to theorize about.

Let's model Sarah's doxastic attitudes as degrees of belief – credences – understood as subjective probabilities satisfying the standard axioms of probability theory. Let $A$ be the set of propositions toward which Sarah has doxastic attitudes. We'll define a credence function $P(\cdot) : A \rightarrow [0, 1]$, mapping each proposition in $A$ to her credence in it.

Prior to receiving Anna's testimony, Sarah had high credence in $h$: "Mark is competent at and enthusiastic about his work." This credence is based on Mark's testimony. Sarah also knows that Anna recently lost an award to Mark. Let's encode

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

Mark's testimony and Sarah's knowledge about the award within Sarah's background information, which we'll designate *i*. Sarah later receives new testimony, *e*, from Anna that Mark's work is sloppy and he seems apathetic. Given her background information, Sarah is unsure how to revise her credence in *h* in light of this testimony and, thus, manage her trust in Mark within the relevant domain.[3]

According to Bayesian orthodoxy, agents should revise their credences by conditionalizing on new evidence. If $P(\cdot)$ is an agent's prior credence function and *e* is new evidence, then her updated credence in some proposition *p*—her posterior credence in *p*—should be $P'(p) = P(p|e) = \frac{P(e|p)\cdot P(p)}{P(e)}$. So, Sarah's updated credence in *h*, upon receiving *e* and given her background information *i*, should be:

$$P(h|e,i) = \frac{P(e|h,i)\cdot P(h|i)}{P(e|i)}. \tag{1}$$

The critical term is the likelihood, $P(e|h,i)$, the probability that Anna testifies that Mark is sloppy and apathetic, given that he's actually competent and engaged; he said as much, and Anna recently lost an award to him. Given this setup, $P(e|h,i)$ is a function of Sarah's assessment of Anna's reliability. If she judges Anna highly reliable, then $P(e|h,i)$ should be low. This is where the complexity of the case bites, because Sarah is uncertain of Anna's reliability: her recent loss to Mark makes it salient that her testimony might be biased, but Sarah is unsure what specific value in $[0, 1]$ Anna's reliability *R* takes (where 0 amounts to perfect anti-reliability and 1 amounts to perfect reliability). Lacking a precise value for this, Sarah cannot determine what specific value $P(e|h,i)$ should take.

This kind of uncertainty calls for a model that can represent both an agent's credence in a proposition and her uncertainty about the reliability of the source on whose word that credence is based. Fortunately, Olsson (2011, 2013) has developed a model that is up to this task. In his model, the recipient of a report does not assign a precise reliability to a source but instead maintains a probability distribution over possible values their reliability could take in the interval $[0, 1]$. I'll call this distribution a *reliability distribution*.

Let $\tau(x)$ denote a probability density function representing Sarah's reliability distribution for Anna. As a shorthand, I'll call any probability density function representing a reliability distribution a *reliability profile*. For any interval $[a, b] \subseteq [0, 1]$, the probability that Anna's reliability falls within that interval is given by:

$$P(a \leq R \leq b) = \int_a^b \tau(x)dx. \tag{2}$$

The mean (or expected value) of Sarah's reliability distribution for Anna, $\mathbb{M}[\tau(x)]$, is given by:

$$\mathbb{M}[\tau(x)] = \int_0^1 x\tau(x)dx, \tag{3}$$

which we may think of as Sarah's best guess as to how reliable Anna is.

With these details in view, we can now look at two key equations in Olsson's model.[4] First, when an agent receives testimony *e* in favor of some proposition *h*, her prior credence in *h* is updated according to:

[3]Of course, she is also unsure how to manage her trust in Anna. I focus solely on her trust in Mark here to demonstrate the interpersonal upshots of judgments of a source's reliability.

[4]I draw on Merdes et al.'s (2020) presentation of the model.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

$$P_{t+1}(h) = P_t(h|e) = \frac{\mathbb{M}[\tau_t(x)] \cdot P_t(h)}{[\mathbb{M}[\tau_t(x)] \cdot P_t(h)] + (1 - \mathbb{M}[\tau_t(x)])(1 - P_t(h))} \quad (4)$$

where $P_t(h)$ is her prior credence in $h$, $P_{t+1}(h)$ is her posterior credence in $h$ after receiving $e$, and $\mathbb{M}[\tau_t(x)]$ is the mean of her reliability distribution for the speaker at time $t$.[5] Simultaneously, the reliability profile itself is updated as follows:

$$\tau_{t+1}(x) = \tau_t(x|e) = \frac{x \cdot P_t(h) + (1 - x)(1 - P_t(h))}{[\mathbb{M}[\tau_t(x)] \cdot P_t(h)] + (1 - \mathbb{M}[\tau_t(x)])(1 - P_t(h))} \cdot \tau_t(x) \quad (5)$$

where $\tau_t(x)$ is the agent's reliability profile for the speaker prior to receiving evidence $e$, and $\tau_{t+1}(x)$ is the posterior profile reflecting the agent's revised judgment about the speaker's reliability after receiving $e$.

What matters most for our purposes is that the model shows precisely how uncertainty about a speaker's reliability affects the evidential force of their testimony, since it allows us to compute the likelihood $P(e|h, i)$ as an average over possible reliability values:

$$P(e|h, i) = \int_0^1 P(e|h, i, x) \cdot \tau(x|i) dx, \quad (6)$$

where $x$ represents a possible value Anna's reliability could take. In Olsson's model, though, reliability is defined as the probability of reporting truthfully, so, more carefully, $x$ represents a possible value for the probability that Anna reports truthfully.[6] Since $e$ is a report that Mark is incompetent, Anna's report is false if $h$ ("Mark is competent") is true. So, the probability of her giving that report, $P(e|h, i, x)$, is $1 - x$. Substituting into the integral above yields:[7]

$$P(e|h, i) = \int_0^1 (1 - x)\tau(x|i) dx = 1 - \mathbb{M}[\tau(x|i)]. \quad (7)$$

This result shows that Sarah's assessment of Anna's reliability – captured by the mean of her reliability profile for Anna – directly determines the evidential force of Anna's testimony. The model, then, helps clarify the mechanics of Sarah's uncertainty. That said, it also leaves something out. While it tells us how to revise our credences based on someone's testimony, given a reliability profile for them, it does not tell us how a reliability profile should be constructed – that is, how one should determine what values a speaker's reliability might take and distribute one's credence across those values. I'll argue below that there may be no single correct answer to this question. Let's turn there now.

[5]In the **colleagues** case, Anna's testimony is *disconfirming* – she testifies *against h* – so a corresponding formula for disconfirming testimony would apply, in which $\mathbb{M}[\tau_t(x)]$ and $(1 - \mathbb{M}[\tau_t(x)])$ swap roles. In the two calculations in the Appendix, this corresponding formula is applied to the colleagues case.

[6]Olsson's definition has obvious philosophical limitations, but they are immaterial for my purpose here. The core insight of the present paper holds, provided only that reliability can be modeled quantitatively, as it would be on my preferred account of reliability as the product of sincerity and competence, where both are understood probabilistically.

[7]See Proof 1 in the Appendix for details.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

## 4. Testimonial underdetermination and permissive trust

### *4.1. From reference classes to testimonial underdetermination*

To say that a source is reliable or unreliable is to say something about how often her reports are correct, among other things. Since we rarely have complete information about any speaker's reliability, such judgments require generalizing from limited evidence: observations of her past behavior, observations of others in similar circumstances, and relevant background information. Because of this, judgments of a speaker's reliability – whether precise or rough, qualitative or quantitative – are made relative to *reference classes*, sets of testimonial exchanges taken to be similar or representative.

Importantly, what counts as "similar" or "representative" can be tricky, depending on which properties of the speaker, the audience, the environment, or the speech act are treated as relevant. For example, we might assess a speaker's reliability as a coworker, as someone with a personal grievance, or as someone with a strong track record in unrelated domains. Each of these assessments invokes a different reference class. Sometimes, it may be unclear which reference class among several plausible candidates should be used to assess a speaker's reliability. When multiple reference classes are applicable, we face the well-known *reference class problem*: the challenge of determining which reference class is appropriate in some context (Reichenbach 1949, Fetzer 1977, Hájek 2007).

Before we continue, two terminological notes are needed. In what follows, I will occasionally say that reference classes are "plausible," "left open by the evidence," "ruled out by the evidence," and "supported by the evidence." Of course, strictly speaking, reference classes cannot be plausible, nor can evidence support them or rule them out: they are sets of testimonial exchanges, not propositions. These expressions are merely useful shorthands for more cumbersome expressions, such as "Sarah's evidence rules out regarding the features that define the class as causally or statistically relevant" and the like. Second, I will occasionally talk about "endorsing," "adopting," or "prioritizing" a reference class. These should also be interpreted as shorthands for unwieldy expressions like "forming a high credence that a reference class is appropriate and using it in one's deliberations."[8]

Thinking back to **colleagues**, Sarah must have some reference class in mind to construct a reliability profile for Anna. But there seem to be multiple plausible reference classes, including at least: (i) the set of exchanges involving people discussing workplace issues generally; (ii) the set of exchanges involving people complaining about a colleague; (iii) the set of exchanges involving people complaining about a colleague to whom they recently lost an award; and (iv) the set of Sarah's own past exchanges with Anna. In turn, these different reference classes generate different reliability profiles, which feed into how she should update on Anna's testimony.

[8]And of course, neither Sarah nor any real agent typically goes around explicitly reasoning about reference classes or computing reliability distributions. In practice, we likely employ a variety of "fast-and-frugal" heuristics and cognitive shortcuts that approximate such reasoning under real-world constraints (cf. Gigerenzer and Goldstein 1996). The framework I use here, in which agents select reference classes and update their credences accordingly, should be understood as an idealized model that (i) makes explicit the kinds of distinctions our cognitive systems often track implicitly and (ii) captures the messier, subconscious processes by which we partition the space of possibilities into coarse-grained categories.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

Let's consider in slightly more detail just two of these plausible reference classes:

**Competitors**: This class consists of testimonial exchanges involving negative reports from speakers who recently lost a competition to the subject of their testimony. The underlying thought is that competitive dynamics might skew such a speaker's reliability: losers may exaggerate faults or withhold credit. While there's little hard data on workplace gossip like this, familiar social dynamics make it plausible that direct competitors tend to be less reliable sources than neutral parties. If Sarah were to assess Anna's reliability with this reference class in mind, she might reasonably arrive at a reliability distribution with mean 0.3.

**Track record**: This class consists of Sarah's past testimonial exchanges with Anna across a variety of contexts. The guiding idea is that a speaker's reliability across domains in the past underwrites generalizations to new domains: a strong track record justifies a presumption of continued reliability. If Anna has previously been highly reliable in other contexts, Sarah might regard her as likely to be highly reliable in this context, too, adopting a reliability distribution with mean 0.8.

The exact numbers are not important here. What matters is that different plausible reference classes generate very different reliability estimates, with different consequences for Sarah's posterior in $h$ and for her trust in Mark and Anna.

I think Sarah's evidence does not determine which of these reference classes is most appropriate, which she should prioritize in her deliberations. Instead, I'll argue, she faces what philosophers of science call *transient underdetermination* – a situation in which one's available evidence leaves open multiple responses to it. To see this, let's get clearer on underdetermination.

**Transient underdetermination**: An agent's available evidence $e$ transiently underdetermines how she rationally ought to respond to it if and only if:

i. there exist multiple doxastic responses, $r_1, r_2, \ldots, r_n$ to $e$,
ii. each response is incompatible with any other response,
iii. each response is consistent with $e$, and
iv. $e$ does not favor one response over any other.[9]

Sarah's situation, I argue, meets all of these conditions.[10]

We have already seen that condition (i) is met. Let's turn to condition (ii), incompatibility between responses. The relevant sense of incompatibility depends on the type of response at issue. When we're considering full beliefs, incompatibility is a matter of logical inconsistency. For credences, incompatibility is a matter of probabilistic inconsistency. In Sarah's case, the incompatibility between prioritizing the **competitors** reference class and prioritizing the **track record** class is straightforward: the former would lead Sarah to assign Anna an expected reliability of 0.3, while the latter would lead

[9]"Response" includes the formation or revision of a doxastic attitude of any kind (e.g., belief, credence, acceptance). In the case of reference classes, a response would be forming a credence that a particular reference class is appropriate to prioritize in one's deliberations.

[10]This account mirrors extant accounts in the literature but generalizes beyond scientific theory choice. See, for instance, Sklar (1975), Kitcher (2001), Turnbull (2018), and Tulodziecki (2025).

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

her to assign an expected reliability of 0.8. Obviously, Sarah cannot simultaneously assess Anna as both highly reliable and unreliable.

Are these two responses both consistent with Sarah's evidence? It is clear that Sarah's evidence constrains the reference classes she can reasonably adopt: Her past experience with Anna prevents her from dropping her credence in Anna's reliability too low, but the possibility of bias compels her to refrain from assigning too high a reliability to Anna. That's good: we want Sarah's evidence to constrain how she is permitted to respond to it. But her evidence does not rule out either of the reference classes above as appropriate. The **competitors** reference class remains viable because Anna genuinely did lose an award to Mark, providing legitimate grounds for considering how competitive dynamics and interpersonal friction affect reliability. The **track record** reference class remains viable because Anna genuinely has been highly reliable in past interactions, providing legitimate grounds for continued trust. Both reference classes fall within the range of options that Sarah's evidence leaves open, whereas some other reference class consisting solely of testimonial exchanges with known inveterate liars, for example, is ruled out: Sarah has no good reason to think the features of that class bear on Anna's reliability.

Finally, does Sarah's evidence favor either response over the other? To answer this question, we must consider what it would take for evidence to favor one reference class selection over another. Sarah's evidence would favor the **competitors** reference class if it provided a clear indication that competitive dynamics are the dominant factor affecting Anna's reliability in this context – for instance, patterns in Anna's behavior suggesting she struggles with jealousy or disappointment. On the other hand, her evidence would favor the **track record** reference class if it provided a clear indication that Anna's past reliability is the best predictor of her current reliability. But these are exactly the things about which Sarah is uncertain. So, while Sarah's background information and evidence provide relevant information and constrain plausible reference classes, they fail to favor either of our two central reference classes. The upshot: She faces genuine underdetermination, demonstrating the truth of the **testimonial underdetermination** thesis.

### *4.2. Synchronic intrapersonal permissivism in trust decisions*

It is common to hold that underdetermination generates leeway in scientific theory choice (Longino 2004; Biddle 2013; Turnbull 2018). Following Jackson and LaFore (2024) and Jackson (2025), I think the same is true in cases of underdetermination in ordinary situations we confront in our daily lives. To see why, consider the following highly plausible principle:

> **No arbitrary requirements**: Epistemic rationality cannot require an agent to adopt one response to her evidence over another when her evidence does not discriminate between them.

If this principle were false, then rationality could demand that agents make epistemically arbitrary choices, requiring that agents adopt one belief over another without any evidential basis for doing so. That picture is difficult to square with the very idea of epistemic rationality, which is supposed to track reasons given by our evidence. Accordingly, epistemic rationality cannot require Sarah to prioritize one reference class over another: from the standpoint of epistemic rationality, both **competitors** and **track**

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

**record** are live options; she is permitted to regard either as appropriate and to employ it in her deliberations.[11]

Recall that the reliability distribution Sarah accepts for Anna will in turn determine the value of the likelihood $P(e|h, i)$ via the relationship established in Equation (7) above. This has knock-on effects for Sarah's posterior credence in Mark's workplace performance and for her trust in him. To see this, consider two different but rationally permissible ways Sarah might assess Anna's reliability.

> **Trust-reducing interpretation**: Suppose that Sarah has a prior credence of 0.8 in $h$, reflecting her high trust in Mark. She then prioritizes the **track record** reference class and constructs a reliability distribution $\tau_a(x|i)$ for Anna with mean $\mathbb{M}[\tau_a(x|i)] = 0.8$. This would yield $P(e|h, i) = 1 - 0.8 = 0.2$. With this value, Anna's negative testimony would reduce Sarah's credence in $h$ to 0.5, indicating reduced trust in Mark in this domain.[12]
>
> **Trust-increasing interpretation**: Suppose instead that Sarah prioritized the **competitors** reference class and constructed a reliability distribution $\tau_b(x|i)$ with mean $\mathbb{M}[\tau_b(x|i)] = 0.3$. This would yield $P(e|h, i) = 1 - 0.3 = 0.7$. With this value, Anna's negative testimony would actually increase Sarah's credence from $h$ to 0.9, indicating increased trust in Mark in this domain.[13]

Both interpretations represent equally legitimate, rationally permissible ways for Sarah to construct a reliability distribution for Anna and proceed to revise her credence in $h$ and her trust in Mark. In other words, Sarah could rationally accept either a trust-increasing or trust-reducing interpretation of Anna's testimony, establishing a species of *synchronic intrapersonal permissivism*, owing to the difficulty of reference class selection in testimonial exchanges. To the extent that **colleagues** is representative of complex, real-life cases, rational trust may often be quite permissive in this way.[14]

### 4.3. Objections and replies

A natural response to the above is that when an agent's evidence is indeterminate, rationality requires suspension of judgment. If this is right, and Sarah's evidence fails to determine which of these competing reference classes is appropriate, she should not adopt or prioritize either to construct some new assessment of Anna's reliability on its basis. Instead, she ought to remain unmoved, holding fixed her attitudes about Anna and her posterior credence in $h$ until more evidence comes in.

The suspension objection has intuitive force, but its force arises, I think, from trading on unclarity about what "indeterminacy" means. We can distinguish different kinds and

[11]The qualifier "from the standpoint of epistemic rationality" is important: not all options consistent with Sarah's evidence may be live from an all-things-considered perspective; more on this, albeit briefly, below (Section 5).

[12]See Calculation 1 in the Appendix.

[13]See Calculation 2 in the Appendix.

[14]One might object that Sarah faces mere higher-order uncertainty about what her evidence determinately supports. Even if so, however, Dorst (2020) shows that higher-order uncertainty constrains first-order credences but leaves room for latitude (Dorst 2020: 13; cf. also p. 33). Thus, her evidence may not compel her to adopt any particular response to it – even if it in fact supports one unique response – and **testimonial underdetermination** would still be true.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

sources of indeterminacy – and when we do, the suspension objection has much less force in this case. Consider the following cases:

> **No reasonable interpretation**: Miriam is a scientist trying to discern whether Hypothesis Y or rival Hypothesis X is more likely to be correct. She needs to process a batch of raw data that might bear on these hypotheses. There are two models Miriam might use to process the data, each of which is well-established and well-regarded in her field: Model A, a regression with plausible covariates, and Model B with a different specification. Miriam tries both, and it's a wash: neither model reveals that the data clearly support Hypothesis X or Hypothesis Y. The hypotheses seem to be tied across the acceptable models. No reasonable interpretation of the data, then, yields a determinate conclusion.[15]
> **Multiple reasonable interpretations**: The setup is the same as above, except Model A now determinately supports Hypothesis X over Y, while Model B determinately supports Hypothesis Y over X. Multiple reasonable interpretations of the data yield different conclusions.

In the first case, Miriam faces *evidential indeterminacy*, while in the second she faces what statisticians, scientists, and philosophers of science sometimes call *model uncertainty* – uncertainty about which model is appropriate for interpreting one's data or evidence when multiple models are available.[16] We must clearly distinguish between these. It is highly plausible that agents facing evidential indeterminacy are rationally required to suspend judgment between responses to their evidence when the evidence fails to discriminate between them on any reasonable interpretation of it. Any response besides suspension would outstrip the evidence.

It is far less clear that suspension is required in the face of model uncertainty. Proponents of suspension here would counsel Miriam either to suspend between these rival hypotheses, since there's no cross-model winner, or to suspend between rival plausible models entirely. But this would come at a steep cost. Scientists routinely face model uncertainty. If rationality required suspension in these cases, vast swaths of responsible scientific practice would be irrational. Instead, the rational response is not the suspension of judgment about the hypotheses under consideration. Rather, it is to adopt and provisionally rely on a model or average the results of multiple models, guided by various theoretical and practical considerations, proportioning one's posterior credence in the hypothesis to one's credence that the model selection or averaging technique was appropriate.[17]

[15]Talk of "evidence" becomes tricky here. Reflecting on scientific practice suggests drawing a distinction between "data" and "evidence," where evidence may be thought of as data under an interpretation or model – a distinction that traditional epistemology has largely overlooked. On the resulting view, evidence and evidential support would always be model relative. Developing these thoughts and their implications further would take us too far afield, however, so I leave this for future work and use "data" and "evidence" interchangeably here for simplicity's sake.

[16]This is also called "structural uncertainty" and "model structure uncertainty." For foundational discussions in statistics, see, inter alia, Hodges (1987), Chatfield (1995), Draper (1995), and Clyde and George (2004), although the problem of model uncertainty was recognized earlier, if not always under that name.

[17]There is immense literature on model selection, uncertainty quantification, and statistical techniques for handling model uncertainty – see Claeskens and Hjort (2008) for a book-length overview – along with concurrent discussions of the issues in the philosophy of science (see, among countless others, Forster and Sober (1994), Sober (2002), Dowe et al. (2007), and Sterner and Lidgard (2024)).

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

I think that **colleagues** is relevantly analogous to **multiple reasonable interpretations**. Both Miriam and Sarah must settle on which features of the phenomena of interest are relevant, which relationships and patterns in the data should be prioritized, and which interpretive framework – whether a formal model, as in Miriam's case, or a judgment about reference classes and reliability, in Sarah's – should be employed to interpret the data. And in both, multiple principled interpretive frameworks yield different determinate conclusions, with indeterminacy residing in the choice of framework rather than at the level of evidential support following this choice. The structural analogy between the cases suggests that they require the same normative treatment. If this is correct, it would follow that Sarah, like Miriam, need not suspend judgment between her competing interpretive frameworks and thus need not suspend judgment about Anna's reliability.

Moving on from suspension of judgment, one might instead object that Sarah should average over the reference classes her evidence leaves open, constructing a weighted combination of the corresponding reliability distributions. The objector might claim that this approach would generate an appropriately cautious assessment of Anna's reliability. But following this strategy would require Sarah to assign weights to competing reliability distributions, and determining appropriate weights is precisely what is at issue: her evidence provides no principled basis for weighting a distribution constructed from the **competitors** class more or less heavily than one constructed from the **track record** class. The averaging strategy thus pushes underdetermination up one level rather than resolving it.[18] In summary, then, Sarah's case is genuinely permissive: Owing to underdetermination, multiple incompatible responses to her evidence are epistemically permissible, each pointing to a different posterior level of trust in Mark.

## 5. The pragmatics of trust decisions and arbitrariness

If underdetermination renders incompatible responses to testimony permissible, how should agents actually make trust decisions? While developing a complete theory of the pragmatics of trust decisions would take us too far afield here, I suggest that while purely epistemic considerations determine the bounds of epistemically permissible trust, practical considerations may guide one's choice between competing epistemically permissible options when the evidence fails to determine a unique response.[19]

[18]One might reply that if Sarah is genuinely in equipoise regarding the candidate reference classes, she should assign equal weight to the corresponding reliability distributions. This approach could be motivated by appeal to the principle of indifference (PoI). Following this approach, Sarah would assign equal weights (probability 0.5) to the propositions "the **competitors** reference class is the most appropriate" and "the **track record** reference class is the most appropriate," then use these weights to average the reliability distributions. But this application of PoI faces the familiar partition-dependence problem (Keynes 1921; van Fraassen 1989; Rinard 2014; Eva 2019): Sarah could equally well partition the space of propositions about which reference class is the most appropriate in other ways by considering additional reference classes – e.g., "workplace testimony generally," "testimony from close colleagues," and "testimony from disappointed competitors specifically" – and different partitions would yield different equiprobable weightings of reliability distributions. Given these difficulties, I reject this approach and conclude that the averaging objection fails.

[19]This position resembles arguments in the philosophy of science that non-epistemic values can shape and guide theory choice in the face of underdetermination – see Longingo (2004; 2008), Intemann (2005), and Elliott (2011), along with Brown (2013) and DiMarco and Khalifa (2019) for more discussion.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

Two kinds of practical considerations are especially salient in testimonial contexts. The first is the value of existing trusting relationships. In the face of underdetermination, agents like Sarah may allow the value of their preexisting trusting relationships to shape their trust decisions.[20] One must also consider the risks of misplaced trust – personally, professionally, and emotionally. Sarah might worry, for example, that if she continues trusting Mark and he wronged his team members, she risks becoming complicit in a pattern of workplace injustice. Moreover, one must take into account one's tolerance for epistemic risk. Just as we differ in our willingness to take practical risks, we also differ in our comfort with the risk of error, striking different balances between James's (1897) edicts to believe truly and avoid believing falsely.[21] Of course, not all practical considerations may appropriately guide trust decisions. If I must choose between preserving or reducing my trust in another, and I'm guaranteed a slice of apple pie if I make the latter choice, reducing my trust for that reason would be wrong because it manifests a failure to respect the significance of trust.

This broadly pragmatic approach to trust decisions helps blunt the force of a final objection to the permissivism developed above – namely, that it licenses problematically arbitrary trust decisions.[22] Following Quanbeck and Worsnip (2025), we can distinguish between what is *epistemically* permissible or obligatory and what is *all-things-considered* permissible or obligatory, where the latter incorporates practical, in addition to purely epistemic, considerations. Even if one's evidence renders epistemically permissible two or more incompatible trust decisions, practical considerations may further constrain what trust decisions one all-things-considered ought to make – perhaps even determining a uniquely permissible response *at that level*. In such a case, one would be forbidden from the kind of arbitrary flip-flopping between epistemically permissible options that concerns White (2005) and other proponents of uniqueness. While it may be that the choice between epistemically permissible options *is*, in some sense, arbitrary at the epistemic level, this arbitrariness is constrained at the level where it matters most – the all-things-considered level.

There is, of course, much more to be said – more than can be said here – about fitting reasons for trust and distrust, about weighing various practical considerations against one another and against epistemic considerations when making trust choices, and about the relationship between the practical and the epistemic. I cannot pretend to be able to resolve these issues here. That said, these remarks, while provisional, point toward valuable new lines of inquiry into how practical and epistemic considerations jointly shape our trust in others – lines I leave for future work.

## 6. Conclusion

In a wide range of testimonial exchanges, we find ourselves confronted with complexity. When we are, we need to make difficult decisions about whom to trust and how much. These decisions will be sensitive to our assessment of the reliability of those whose testimony we've received. And assessing someone's reliability is no mean feat: these

[20]For discussion of epistemic partiality like this, see Stroud (2006), Keller (2004), Hawley (2014b), Goldberg (2019), Mason (2023), and Brinkerhoff and Arpaly (2025).

[21]For discussion, see, among others, Kelly (2013); Levinstein (2017); Pettigrew (2022); Jackson (2025).

[22]See White (2005) for an early formulation of the arbitrariness objection to permissivism, as well as Feldman (2007), Ballantyne and Coffman (2011), White (2014), Schoenfield (2014), Simpson (2017), Titelbaum and Kopec (2019), Smith (2020), and Ye (2021) for subsequent discussion.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

assessments are made relative to reference classes, and it may be genuinely difficult to discern whether any particular, seemingly applicable reference class is most appropriate. Sometimes, multiple reference classes may be appropriate. When they are, we face genuine underdetermination, our evidence leaving open a range of options for how reliable to find someone and how to apportion our trust.

Recognizing that rational trust may often be underdetermined and permissive has several important implications. First, it opens up new hypotheses for the causes of widespread persistent disagreement, polarization, and apparent resistance to evidence, which may in some cases arise due to rational leeway in the interpretation of testimonial evidence, owing to the difficulty of assessing others' reliability. Moreover, if rational trust is often underdetermined and permissive, this raises new questions about how trust should be managed responsibly over time. My trust decisions right now will shape how I interpret my future evidence and revise my trust. If there were always a uniquely rational trust decision, this diachronic dimension would be trivial. But if there is latitude in trust, then different yet rationally permissible trust decisions today may set me on very different epistemic trajectories tomorrow – something the purely synchronic focus of many traditional accounts of testimonial trust may lead us to overlook.

The most important upshot, however, is this: navigating testimony responsibly in the real world is not always a purely epistemological affair. When our testimonial evidence is ambiguous and underdetermining, as I suspect it often is, practical considerations – relational, moral, social, and political – may play a necessary role in shaping our trust decisions. Epistemologists must not bracket these considerations if we want accounts of testimony and trust to capture their full nature and significance in actual human lives.

**Acknowledgments.** For helpful comments on earlier drafts of this paper, I thank Sam Fullhart, Taylor Koles, Travis McKenna, Conner Schultz, and the audience at the Fall 2025 UNC PPE Faculty Workshop. Special thanks go to Marina DiMarco, whose generous discussions with me about underdetermination and values in science meaningfully improved the paper.

**Competing interests.** The author declares no competing interests.

## References

**Baier A.** (1986). 'Trust and Antitrust.' *Ethics* **96**(2), 231–60.

**Ballantyne N. and Coffman E.J.** (2011). 'Uniqueness, Evidence, and Rationality.' *Philosophers' Imprint* **11**(18), 1–13.

**Biddle J.** (2013). 'State of the Field: Transient Underdetermination and Values in Science.' *Studies in History and Philosophy of Science Part A* **44**(1), 124–33.

**Brinkerhoff A. and Arpaly N.** (2025). 'Epistemic Partiality in Friendship.' In J. Lackey and A. McGlynn (eds), *Oxford Handbook of Social Epistemology*, pp. 201–18. New York: Oxford University Press.

**Brown M.J.** (2013). 'Values in Science beyond Underdetermination and Inductive Risk.' *Philosophy of Science* **80**(5), 829–39.

**Burge T.** (1993). 'Content Preservation.' *The Philosophical Review* **102**, 457–88.

**Carr, J.** (2020). 'Imprecise Evidence without Imprecise Credences.' *Philosophical Studies* **177**(9), 2735–58.

**Chatfield C.** (1995). 'Model Uncertainty, Data Mining, and Statistical Inference.' *Journal of the Royal Statistical Society, Series A* **158**(3), 419–66.

**Claeskens G. and Lid Hjort N.** (2008). *Model Selection and Model Averaging*. Cambridge: Cambridge University Press.

**Clyde M. and George E.I.** (2004). 'Model Uncertainty.' *Statistical Science* **19**(1), 81–94.

**DiMarco M. and Khalifa K.** (2019). 'Inquiry Tickets: Values, Pursuit, and Underdetermination.' *Philosophy of Science* **86**(5), 1016–28.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

**Domenicucci J. and Holton R.** (2017). 'Trust as a Two-Place Relation.' In P. Faulkner and T. Simpson (eds), *The Philosophy of Trust*, pp. 149–60. New York: Oxford University Press.
**Dorst K.** (2019). 'Higher-Order Uncertainty.' In M. Skipper and A. Steglich-Petersen (eds), *Higher-Order Evidence: New Essays*, pp. 35–61. New York: Oxford University Press.
**Dorst K.** (2020). 'Evidence: A Guide for the Uncertain.' *Philosophy and Phenomenological Research* **100**(3), 586–632.
**Dowe D.L., Gardner S. and Oppy G.** (2007). 'Bayes not Bust! Why Simplicity Is No Problem for Bayesians.' *The British Journal for the Philosophy of Science* **58**(4), 709–54.
**Draper D.** (1995). 'Assessment and Propagation of Model Uncertainty.' *Journal of the Royal Statistical Society, Series B (Methodological)* **57**(1), 45–97.
**Elliott K.C.** (2011). 'Direct and Indirect Roles for Values in Science.' *Philosophy of Science* **78**(2), 303–24.
**Eva B.** (2019). 'Principles of Indifference.' *The Journal of Philosophy* **116**(7), 390–411.
**Faulkner P.** (2011). *Knowledge on Trust*. New York: Oxford University Press.
**Faulkner P.** (2015). 'The Attitude of Trust is Basic.' *Analysis* **75**(3), 424–29.
**Faulkner P.** (2017). 'The Problem of Trust.' In P. Faulkner and T. Simpson (eds), *The Philosophy of Trust*, pp. 109–28. New York: Oxford University Press.
**Feldman R.** (2007). 'Reasonable Religious Disagreements.' In L.M. Antony (ed), *Philosophers Without Gods*, pp. 194–214. New York: Oxford University Press.
**Fetzer J.** (1977). 'Reichenbach, Reference Classes, and Single Case 'Probabilities".' *Synthese* **34**(2), 185–217.
**Forster M. and Sober E.** (1994). 'How to Tell When Simpler, More Unified, or Less Ad Hoc Theories Will Provide More Accurate Predictions.' *The British Journal for the Philosophy of Science* **45**(1), 1–35.
**Fricker E.** (1994). 'Against Gullibility.' In B.K. Matilal and A. Chakrabarti (eds), *Knowing from Words: Western and Indian Philosophical Analysis of Understanding and Testimony*, pp. 125–61. Dordrecht: Kluwer Academic Publishers.
**Gigerenzer G. and Goldstein D.G.** (1996). 'Reasoning the Fast and Frugal Way: Models of Bounded Rationality.' *Psychological Review* **103**(4), 650–69.
**Goldberg S.** (2019). 'Against Epistemic Partiality in Friendships: Value-Reflecting Reasons.' *Philosophical Studies* **176**(8), 2221–42.
**Goldenberg M.** (2016). 'Public Misunderstanding of Science? Reframing the Problem of Vaccine Hesitancy.' *Perspectives on Science* **24**(5), 552–81.
**Goldman A.** (2001). 'Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* **63**(1), 85–110.
**Hájek A.** (2007). 'The Reference Class Problem is Your Problem Too.' *Synthese* **156**(3), 563–85.
**Hawley K.** (2014a). 'Trust, Distrust, and Commitment.' *Noûs* **48**(1), 1–20.
**Hawley K.** (2014b). 'Partiality and Prejudice in Trusting.' *Synthese* **191**(9), 2029–45.
**Hawley K.** (2019). *How To Be Trustworthy*. New York: Oxford University Press.
**Hodges J.S.** (1987). 'Uncertainty, Policy Analysis, and Statistics.' *Statistical Science* **2**(3), 259–91.
**Hume D.** (1748/[1999]). *An Enquiry Concerning Human Understanding*. New York: Oxford University Press.
**Intemann K.** (2005). 'Feminism, Underdetermination, and Values in Science.' *Philosophy of Science* **72**(5), 1001–12.
**Jackson E.** (2025). 'Permissivist Evidentialism.' In K. McCain, S. Stapleford, and M. Steup (eds), *Evidentialism at 40: New Arguments, New Angles*, pp. 24–40. New York: Routledge.
**Jackson E. and LaFore G.** (2024). 'Permissivism, Underdetermination, and Evidence.' In M. Lasonen-Aarnio and C. Littlejohn (eds), *The Routledge Handbook of the Philosophy of Evidence*, pp. 358–70. New York: Routledge.
**James W.** (1897). 'The Will to Believe.' *The Will to Believe and Other Essays in Popular Philosophy*, pp. 1–31. New York, NY: Longmans, Green, & Co.
**Jones K.** (1996). 'Trust as an Affective Attitude.' *Ethics* **107**(1), 4–25.
**Jones K.** (2012). 'Trustworthiness.' *Ethics* **123**(1), 61–85.
**Jones K.** (2019). 'Trust, Distrust, and Affective Looping.' *Philosophical Studies* **176**(4), 955–68.
**Keller S.** (2004). 'Friendship and Belief.' *Philosophical Papers* **33**(3), 329–51.
**Kelly T.** (2013). 'Evidence Can Be Permissive.' In M. Steup, E. Sosa, and J. Turri (eds), *Contemporary Debates in Epistemology* (2nd ed.), pp. 305–11. Hoboken, NJ: Wiley-Blackwell.
**Keynes J.M.** (1921). *A Treatise on Probability*. London: Macmillan.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

**Kitcher P.** (2001). *Science, Truth, and Democracy*. New York: Oxford University Press.
**Lackey J.** (2008). *Learning From Words: Testimony as a Source of Knowledge*. New York: Oxford University Press.
**Lackey J.** (2021). 'Echo Chambers, Fake News, and Social Epistemology.' In S. Bernecker, A. Flowerree, and T. Grundmann (eds), *The Epistemology of Fake News*, pp. 206–27. New York: Oxford University Press.
**Levinstein B.A.** (2017). 'Permissive Rationality and Sensitivity.' *Philosophy and Phenomenological Research* **94**(2), 342–70.
**Longino H.** (2004). 'How Values Can Be Good for Science.' In P. Machamer and G. Wolters (eds), (2004). *Science, Values, and Objectivity*, pp. 127–42. Pittsburgh, PA: University of Pittsburgh Press.
**Longino H.** (2008). 'Values, Heuristics, and The Politics of Knowledge.' In M. Carrier, D. Howard, and J.A. Kourany (eds), *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, pp. 68–86. Pittsburgh, PA: University of Pittsburgh Press.
**Mason C.** (2023). 'Epistemic Partialism.' *Philosophy Compass* **18**(2), e12896.
**Merdes C., von Sydow M. and Hahn U.** (2020). 'Formal Models of Source Reliability.' *Synthese* **198**(S23), 5773–801.
**Olsson E.J.** (2011). 'A Simulation Approach to Veritistic Social Epistemology.' *Episteme* **8**(2), 127–43.
**Olsson E.J.** (2013). 'A Bayesian Simulation Model of Group Deliberation and Polarization.' In F. Zenker (ed), *Bayesian Argumentation: The Practical Side of Probability*, pp. 113–33. Dordrecht: Springer.
**Pettigrew R.** (2022). *Epistemic Risk and the Demands of Rationality*. New York: Oxford University Press.
**Quanbeck Z. and Worsnip A.** (2025). 'A Permissivist Alternative to Encroachment.' *Philosophers' Imprint* **25**(1), 1–22. https://doi.org/10.3998/phimp.3191.
**Reichenbach H.** (1949). *The Theory of Probability*. Berkeley, CA: University of California Press.
**Reid T.** (1764/[1997]). *An Inquiry into the Human Mind on the Principles of Common Sense*. Edinburgh: Edinburgh University Press.
**Rinard S.** (2014). 'The Principle of Indifference and Imprecise Probability.' *Thought* **3**(2), 110–14.
**Schoenfield M.** (2014). 'Permission to Believe: Why Permissivism Is True and What It Tells Us About Irrelevant Influences on Belief.' *Noûs* **48**(2), 193–218.
**Simion M.** (2021). 'Testimonial Contractarianism: A Knowledge-First Social Epistemology.' *Noûs* **55**(4), 891–916.
**Simpson R.M.** (2017). 'Permissivism and the Arbitrariness Objection.' *Episteme* **14**(4), 519–38.
**Simpson T.** (2012). 'What Is Trust?" *Pacific Philosophical Quarterly* **93**(4), 550–69.
**Sklar L.** (1975). 'Do Unborn Hypotheses Have Rights?" *Pacific Philosophical Quarterly* **62**(1), 17–29.
**Smith J.J.** (2020). 'Unacknowledged Permissivism.' *Pacific Philosophical Quarterly* **101**(1), 158–83.
**Sober E.** (2002). 'Instrumentalism, Parsimony, and the Akaike Framework.' *Philosophy of Science* **69**(S3), S112–S123.
**Sterner B. and Lidgard S.** (2024). 'Objectivity and Underdetermination in Statistical Model Selection.' *The British Journal for the Philosophy of Science* **75**(3), 717–39.
**Stroud S.** (2006). 'Epistemic Partiality in Friendship.' *Ethics* **116**(3), 498–524.
**Titelbaum M.G. and Kopec M.** (2019). 'When Rational Reasoners Reason Differently.' In M.B. Jackson and B.B. Jackson (eds), *Reasoning: New Essays on Theoretical and Practical Thinking*, pp. 205–31. New York: Oxford University Press.
**Tulodziecki D.** (2025). *Underdetermination and Theoretical Virtues*. New York: Cambridge University Press.
**Turnbull M.G.** (2018). 'Underdetermination in Science: What It Is and Why We Should Care.' *Philosophy Compass* **13**(2), e12475.
**Van Fraassen B.** (1989). *Laws and Symmetry*. New York: Oxford University Press.
**White R.** (2005). 'Epistemic Permissiveness.' *Philosophical Perspectives* **19**(1), 445–59.
**White R.** (2014). 'Evidence Cannot Be Permissive.' In M. Steup, J. Turrie, and E. Sosa (eds), *Contemporary Debates in Epistemology*, 2nd edition, pp. 312–323. Chichester: Wiley–Blackwell.
**Ye R.** (2021). 'The Arbitrariness Objection Against Permissivism.' *Episteme* **18**(4), 654–73.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

## Appendix

This appendix collects details referenced in the main text. Proof 1 establishes the relationship in Equation (7): $P(e|h,i) = 1 - \mathbb{M}[\tau(x|i)]$, while Calculations 1 and 2 work through the trust-reducing and trust-increasing interpretations of Anna's testimony that are referenced in Section 4.2.

**Proof 1.** We establish the relationship $P(e|h,i) = 1 - \mathbb{M}[\tau(x|i)]$, where $h$ represents the proposition that Mark is competent, $e$ represents Anna's countervailing testimony about Mark, $i$ represents Sarah's background information, and $\tau(x|i)$ is the probability density function representing Sarah's reliability distribution for Anna given background information $i$.

By the law of total probability for continuous random variables:

$$P(e|h,i) = \int_0^1 P(e|h,i,x)\tau(x|i)dx,$$

where $x$ represents a possible value of Anna's reliability.

Reliability in the model is defined as the probability that a speaker provides truthful testimony. We are computing $P(e|h,i,x)$, the probability of Anna's testimony $e$ given that $h$ is true. When $h$ is true, Anna's testimony $e$ is false. Thus, the probability that Anna provides this testimony is the complement of her reliability:

$$P(e|h,i,x) = 1 - x.$$

Substituting this relationship into our integral:

$$P(e|h,i) = \int_0^1 (1-x) \cdot \tau(x|i)dx.$$

We split the integral:

$$\int_0^1 (1-x)\tau(x|i)dx = \int_0^1 \tau(x|i)dx - \int_0^1 x\tau(x|i)dx.$$

Because $\tau(x|i)$ is a probability density function over $[0,1]$, we have $\int_0^1 \tau(x|i)dx = 1$. By the definition of expected value, $\int_0^1 x\tau(x|i)dx = \mathbb{M}[\tau(x|i)]$.

Therefore,

$$P(e|h,i) = 1 - \mathbb{M}[\tau(x|i)].$$

**Calculation 1: Trust-reducing interpretation.** Suppose that Sarah's prior credence in $h$ is 0.8, manifesting high trust in Mark. We are given that Sarah adopts reliability profile $\tau_a(x|i)$ with mean $\mathbb{M}[\tau_a(x|i)] = 0.8$. She then receives $e$, Anna's negative testimony about Mark. From the relationship established in Proof 1, $P(e|h,i) = 1 - \mathbb{M}[\tau_a(x|i)] = 1 - 0.8 = 0.2$. Recall that Anna testifies against $h$. When $\neg h$ is true (Mark is incompetent), reliable testimony would correctly report this. Therefore, $P(e|\neg h,i) = \mathbb{M}[\tau_a(x|i)] = 0.8$. By the law of total probability, $P(e|i) = P(e|h,i) \times P(h|i) + P(e|\neg h,i) \times P(\neg h|i)$. So $P(e|i) = 0.2 \times 0.8 + 0.8 \times 0.2 = 0.16 + 0.16 = 0.32$. We now apply Bayes' theorem: $P(h|e,i) = \frac{P(e|h,i) \times P(h|i)}{P(e|i)}$. Thus $P(h|e,i) = \frac{0.2 \times 0.8}{0.32} = \frac{0.16}{0.32} = 0.5$.

**Calculation 2: Trust-increasing interpretation.** Suppose that Sarah's prior credence in $h$ is 0.8, manifesting high trust in Mark. We are given that Sarah adopts reliability profile $\tau_b(x|i)$ with mean $\mathbb{M}[\tau_b(x|i)] = 0.3$. She then receives $e$, Anna's negative testimony about Mark. From the relationship established in Proof 1, $P(e|h,i) = 1 - \mathbb{M}[\tau_b(x|i)] = 1 - 0.3 = 0.7$. Recall that Anna testifies against $h$. When $\neg h$ is true (Mark is incompetent), reliable testimony would correctly report this. Therefore, $P(e|\neg h,i) = \mathbb{M}[\tau_b(x|i)] = 0.3$, reflecting the bias Sarah attributes to Anna. By the law of total probability, $P(e|i) = P(e|h,i) \times P(h|i) + P(e|\neg h,i) \times P(\neg h|i)$. So $P(e|i) = 0.7 \times 0.8 + 0.3 \times 0.2 = 0.56 + 0.06 = 0.62$. We now apply Bayes' theorem: $P(h|e,i) = \frac{P(e|h,i) \times P(h|i)}{P(e|i)}$. Thus, $P(h|e,i) = \frac{0.7 \times 0.8}{0.62} = \frac{0.56}{0.62} \approx 0.90$.

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press

## Author Biography

**William Conner** is a Postdoctoral Fellow in the Department of Philosophy at the University of North Carolina at Chapel Hill. His research focuses on formal and social epistemology, with particular interests in testimony, trust, and inquiry. He draws on tools and concepts from the philosophy of science and statistics to address these topics and investigate the limits of Bayesian epistemology.

**Cite this article:** Conner W. (2026). "Trust Issues: Reliability, Reference Classes, and the Underdetermination of Trust." *Episteme* 1–19. https://doi.org/10.1017/epi.2026.10126

https://doi.org/10.1017/epi.2026.10126 Published online by Cambridge University Press