The Emergent Basis of Expert Trust

Igor Douven, Nikolaus Kriegeskorte

DOI: 10.1017/psa.2025.10186
This is a manuscript accepted for publication in *Philosophy of Science*.
This version may be subject to change during the production process.

The Emergent Basis of Expert Trust*

Igor Douven† Nikolaus Kriegeskorte‡

Abstract

Goldman (2001) asks how novices can trust putative experts when background knowledge is scarce. We develop a reinforcement-learning model, adapted from Barrett, Skyrms, and Mohseni (2019), in which trust arises from experience rather than prior expertise labels. Agents incrementally weight peers who outperform them. Using a large dataset of human probability judgments as inputs, we simulate communities that learn whom to defer to. Both a strictly individual-learning variant and a reputation-sharing variant yield performance-sensitive deference, the latter accelerating convergence. Our results offer an empirically grounded account of how communities identify and trust experts without blind deference.

Keywords: agent-based modeling; expertise; reinforcement learning; social epistemology; testimony; trust.

1 Introduction

Goldman (2001, 2021) raises the question of how a novice faced with a putative expert may come to trust the expert's testimony. He offers a range of strategies (e.g., examining track records, and attending to disciplinary consensus or argumentative strength) that the novice could use. However, he also notes that these strategies rely on background information that will not always be available to novices. This raises the *problem of blind trust*, which is the concern that a novice may sometimes have to accept putative expert testimony in the absence of any good reason to think that the so-called expert is trustworthy.

Not all agree that this is properly called a *problem*. According to Burge (1993), Foley (1994, 2001), and others, accepting testimony on the basis of an expert's say-so can be *justified*, for—they claim—we are prima facie entitled to accept what someone asserts, unless given reason to doubt. But in typical cases in which we seek expert advice, we need more than a mere "absence of reason to doubt" before following the advice. Goldman speculates that at least sometimes the "more" could come from inductive inferences we draw concerning someone's testimonial trustworthiness, but Coady (1992) and others (including Burge and Foley) seem right that we often lack sufficient evidence to warrant such inferences.

This paper approaches the problem of how to come to trust experts from a different angle. Rather than beginning with labeled experts and asking how we might recognize them or recognize those whose advice we should heed, we show how the status of expert may emerge from epistemic interactions not involving any inductive reasoning. Our goal is to thereby provide a unified answer to both the question of

*Supplementary Materials, containing the Julia code (Bezanson et al., 2017) used for the simulations and analyses we report as well as high-resolution versions of all figures, can be downloaded from this repository: https://osf.io/ytg8k/?view_only=374bff118f2943a9ba1c6fb849205f8f.

†SND / CNRS / Sorbonne University, Paris, France, igor.douven@sorbonne-universite.fr.
‡Zuckerman Mind Brain Behavior Institute, Columbia University, NYC, USA, nk2765@columbia.edu.

how to justify trusting certain experts and the question of how experts come to be recognized as such in the first place. To achieve this goal, we recruit an agent-based model developed by Barrett, Skyrms, and Mohseni (2019), who show how epistemic networks can, through trial and error, "self-assemble" in ways that benefit all their members.

We adapt and extend this model in several ways. First, we apply it to a large empirical dataset of probability assignments by hundreds of participants to hundreds of statements with known ground truth. This allows us to define the actual epistemic reliability of each agent not by stipulation (as Barrett and colleagues do), but by empirical performance. Second, we compare two variants of the model, one in which each agent privately reinforces its own successful interactions, and one in which the agents collectively create a shared pool of reputational information reflecting their aggregated experiences. Finally, in neither variant of our model can agents observe nature at all; they are strictly dependent on one another. By contrast, Barrett and colleagues allow both observation of nature (success tied to epistemic reliability) and communication (success tied to social reliability).

We find that, over successive rounds of interaction, agents increasingly defer to those with greater overall expertise (operationalized in terms of classification correctness and Brier score, calculated from their performance on the entire dataset). Even without explicit reasoning or track record evaluation, the model reliably amplifies the influence of the more epistemically successful agents. When reputational information is pooled—as is done in one version of the model—this effect becomes even stronger, leading to a further improvement in epistemic performance of the community as a whole and a more accurate convergence on its epistemically most successful members. The model thus offers an empirically grounded account of how trust in expertise can emerge from simple, local learning mechanisms. Importantly, the trust that arises in this model is not blind, but is earned through interaction, even if agents themselves are not *reasoning* about track records or other truth indicators.

The remainder of this paper is organized as follows. Section 2 starts by describing Barrett et al.'s model and then introduces our model, in its two variants. Section 3 describes the dataset of probability assignments that we ran our model on and also states the details of the agent-based simulations that we performed. Section 4 presents the results of these simulations and discusses their philosophical implications. In Section 5, we conclude by stating limitations of our current approach and mentioning avenues for future research.

Before starting, we should emphasize that our aim in this paper is purely descriptive: to show how performance-sensitive deference can emerge from simple reinforcement-style learning without explicit reasoning about reliability or track records. This complements—but does not settle—the normative question of when such trust is justified. Barrett (2024) argues that Humean custom can bootstrap "learning how to learn," where simple reinforcement guides the adoption of more sophisticated rules without presupposing rational justification (see also Douven 2025). Our results fit that picture: local reinforcement suffices for communities to center reliable members, even if normative justification requires further argument (e.g., meta-inductive defenses, in the manner of Schurz 2012, 2019).

2 Modeling reinforcement and recognition

Reinforcement learning (RL) is a branch of machine learning in which agents learn to make decisions by interacting with an environment and receiving feedback, often as rewards or penalties, for their actions (Roth & Erev 1995; Erev & Roth 1998; Sutton & Barto 2018; Haas 2022). Gradually, agents adjust their behavior to favor actions that in the past led to better outcomes, thereby learning an effective policy. RL

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

has numerous known applications, for example, in robotics (e.g., Haarnoja et al. 2018), autonomous driving (e.g., Tesla's FSD v12), medical decision-making (Coronato et al. 2020), and training large language models (Havrilla et al. 2024; DeepSeek-AI 2025; Xu et al. 2025). Most famously, DeepMind used RL to train its AlphaGo program, which defeated the world champion Go player in 2016 (Silver et al. 2016).

Here, we are interested in an application developed by Barrett, Skyrms, and Mohseni (2019), which implements RL in an agent-based model to study the self-assembly of epistemic networks through local interaction dynamics. We build on this model to address the research questions mentioned in the introduction—how experts can be identified and how trust in them can emerge through simple interactions, without the need for blind trust or track record keeping. We first describe Barrett et al.'s model and then detail our specific adaptations and extensions.

The core idea of Barrett et al.'s model is that agents build up dispositions to consult one another through trial-and-error learning. In their setup, each agent is equipped with an "urn" containing identifiers for all agents, itself, and for nature. Observing nature succeeds with probability equal to an agent's *epistemic reliability*, and communication succeeds with probability equal to the *social reliability* (which may be set to 1 in the perfect-communication case). Each round, the agent draws a ball from its urn to decide whether to consult nature directly or to ask an agent for their belief. If the result of the consultation (or observation) is successful—that is, if it leads to a true belief (which is taken to be revealed to the agent)—the agent reinforces the chosen action by adding to the urn another ball with the same name on it as on the ball that was drawn. Over time, successful sources of information (whether nature or specific agents) become increasingly likely to be consulted again, resulting in the emergence of a social network with weaker and stronger connections (with strength determined by consultation frequency).[1]

Note that the learning process in Barrett et al.'s model is local and nondeliberative. Agents do not reason about who is reliable or maintain explicit beliefs about the accuracy of others. They simply reinforce what works. But while simple, doing so helps them group into social networks in which less reliable inquirers defer to those who, through repeated success, have proven to be good informants. Most notably, by forming these networks agents tend to become more epistemically successful than they would have been if they acted on their own.

The model we develop in this paper is inspired by Barrett et al.'s approach but departs from it in important respects, both theoretical and methodological. First, we replace the abstract and randomly assigned reliabilities of Barrett et al.'s agents with empirical performance data drawn from a large experimental dataset. In this dataset, hundreds of participants assigned probability judgments to hundreds of factual statements whose truth values are known to the experimenters (see Sect. 3). From these judgments, we calculate each participant's overall classification correctness and average Brier score across all statements, providing concrete, empirically-grounded measures of their epistemic reliability. The agents in our model are "copies" of these participants (in that the agents each reflect all the judgments of a different participant), so that we simulate interactions among agents whose behavior is grounded in real-world human data rather than among agents with stipulated probabilities.

Second, our model focuses exclusively on social deference rather than on the choice between observing nature and consulting others. In Barrett et al.'s model, each agent faces a decision between gathering new evidence (observing nature) and consulting a peer. In our setting, the data do not include independent

[1]This urn learning model follows the reinforcement tradition of Roth–Erev, in which choices that "worked" are reinforced by increasing their sampling weight, and where the process is non-Markovian in that it aggregates the entire history of successes (Roth & Erev 1995; Erev & Roth 1998). In the Sutton–Barto approach (Sutton & Barto 2018), by contrast, RL typically tracks value functions or policies over explicit state–action spaces. Human and animal learning often exhibits the former (simpler) kind of RL (Schultz, Dayan, & Montague 1997; Niv 2009).

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

evidence-gathering actions. Rather, each agent is associated with a fixed set of responses to factual statements (as determined by the aforementioned dataset), and learning is simulated by agents using the urn mechanism to choose whose judgment to defer to on a given claim. Unlike the agents in Barrett et al.'s model, our agents do not pass along beliefs they acquired from third parties. Each reinforcement step compares the same claim: an agent's own probability versus a consulted agent's probability, with reinforcement to the better of the two.

Third, we add to the model the ability to mark reputation via a collective urn. In Barrett et al.'s model, each agent has a private urn, which is strictly updated on the basis of the agent's own interactions. In our second variant, we allow the contents of the urns of all agents to be merged into a collective urn from which each agent samples. This pooling simulates a form of social reputation building, allowing agents to benefit from the experiences of others and not just their own. An agent who is widely trusted and performs well will appear frequently in the collective urn, increasing their chances of being consulted even by agents who, so far, have not interacted with them. In other words, successful agents not only get reinforced by individual agents, but also acquire a reputation within the group.[2]

In short, while the logic of reinforcement remains central, our model extends Barrett et al.'s framework along three dimensions: (i) by grounding agent performance in empirical data; (ii) by isolating social dynamics; and (iii) by adding reputation marking. The resulting model, in either variant (with and without reputation marking), is simple, yet allows for an empirically grounded study of how expert trust might emerge from epistemically basic social interactions.

3 Experimental setup

To test how expert trust may emerge from social interaction, we simulate a sequence of peer selection processes based on real-world human judgment data. Here, we describe the dataset on which the simulation is based, the mechanics of agent interaction, and the metrics used to assess the degree to which the model successfully identifies and reinforces the most epistemically reliable agents.

Data are from the study reported in Stinson et al. (2025).[3] They consist of the probability judgments from 376 participants (all US citizens) of 1,200 general knowledge claims (e.g., "The first handheld calculator was built before 1975," "Mexico was the main trading partner of the US in the 1990s"), half of which are true and half false. More specifically, the claims were so chosen that they fall into six categories (history, geography, science, sports and leisure, social sciences and politics, arts and entertainment), with 200 claims per category (100 true, 100 false). They were evaluated in six online sessions, in each of which participants rated 200 unique claims, presented in randomized order. (See Stinson et al., 2025, for a more detailed description of the data and the experimental procedure.)

Before running the simulations, we calculated two key global performance metrics for each of the 376 participants based on their probability assignments to the 1,200 claims: (1) classification correctness, which is the proportion of claims where the participant's probability assignment interpreted as a binary prediction (true if > 0.5 and false otherwise) matched the actual truth value of the claim; and (2) the average Brier score, which is the mean squared difference between the participant's assigned probabilities and the claims' actual truth values (0 for false, 1 for true). Figure 1 shows density plots of the distributions

[2]We like to think of this pooling operation as a minimal, tractable proxy for common ways in which communities share reputational information, such as leaderboards, citation indexes, and other publicly accessible quality indicators; but other interpretations are possible as well.

[3]Stinson et al.'s (2025) study is concerned with different ways of aggregating probabilities and not related to our current topic.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

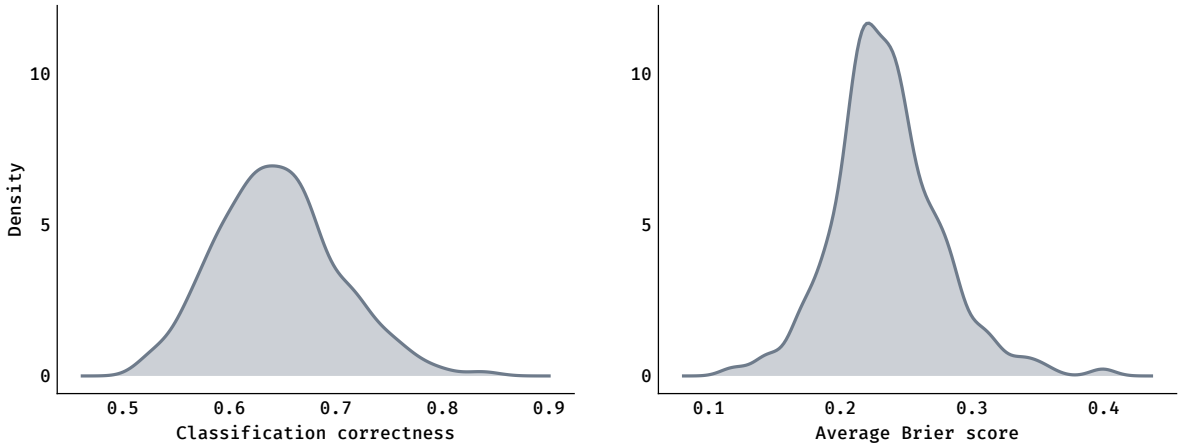

Figure 1: Distributions of overall epistemic performance. Classification correctness (left) and average Brier scores (right) for the 376 agents, calculated across all 1,200 claims.

of these overall scores for all participants (for detailed statistics, see the Supplementary Materials). It should be stressed that these scores are not directly accessible to agents during the simulation but serve only as our external benchmarks for epistemic success.

In the simulation, each participant becomes an agent, and for any of the 1,200 claims, the probability an agent assigns to that claim represents the response it would give if consulted by another agent and asked how probable, according to the consulted agent, the claim is. Each agent has an urn that plays a role in how they select others to consult. Initially, the urn contains one instance of every agent's ID (including the agent's own), indicating a uniform prior disposition to consult any agent. As agents interact, the urn is updated via reinforcement: successful consultations increase the frequency of the consulted agent (i.e., of their ID) in the urn, making them more likely to be drawn again in the future.

The simulation proceeds in discrete rounds (up to 1,000 in our analyses). In each round, every agent (1) randomly selects a factual claim from the dataset; (2) draws an agent (the "informant") from an urn; (3) the Brier score of both its own probability assignment and the consulted informant's probability assignment for the selected claim is computed (after the true answer for that claim is revealed for the purpose of this comparison); (4) reinforces success: if the informant's Brier score for that claim is lower (i.e., better) than its own, the consulting agent adds another instance of the informant's ID to its private urn. If its own score was better or equal, it adds another instance of its own ID to its private urn. In the individual learning variant, agents sample informants from their own private urns throughout. In the reputation marking variant, the contents of all private urns are pooled into a collective urn at the start of each round, and agents sample their informants from this shared collective urn.

Algorithm 1 shows the pseudocode for the variant with reputation marking. That for the simpler, individual learning variant is obtained by omitting the collective pooling step (lines 10–12) and having each agent k sample directly from its own urn $Agents[k].urn$ (in line 15). It should be emphasized that the *score_history* and *self_score* containers (lines 7–8 and updated in lines 21, 24, and 27 of the pseudocode) store Brier scores only for the purpose of our data analysis (see Sect. 4); past scores are never accessed or used by the agents themselves.[4]

[4] Readers interested in the full computational details of the simulations are referred to the Julia code in the Supplementary Materials.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

Algorithm 1 Agent-based reinforcement learning with collective reputation

procedure CollectiveUrnSimulation(N_P, N_C, N_R, Ratings, TruthValues)
  ▷ *N_P: number of participants; N_C: number of claims; N_R: number of runs; Ratings: $N_C \times N_P$ matrix of participant ratings; TruthValues: N_C-element vector of truth values of claims* ◁
  Initialize *Agents* ← array of N_P Agent structures
  for *i* from 1 to N_P **do** ▷ *Initialize each agent*
    Agents[*i*].*id* ← *i*
    Agents[*i*].*urn* ← list containing integers from 1 to N_P ▷ *Initially, urn contains all agent IDs once*
    Agents[*i*].*score_history* ← empty list ▷ *For analysis: agent's actual Brier scores*
    Agents[*i*].*self _score* ← empty list ▷ *For analysis: Brier scores if agent were fully self-reliant*
  for *run* from 1 to N_R **do** ▷ *Main simulation loop*
    CollectiveUrn ← empty list
    for *i* from 1 to N_P **do** ▷ *Pool all tickets from individual urns*
      Append all elements of *Agents*[*i*].*urn* to *CollectiveUrn*
    for *k* from 1 to N_P **do** ▷ *Each agent k makes a decision*
      claim_idx ← RandomInteger(1, N_C) ▷ *Select a random claim*
      informant_id ← RandomSample(*CollectiveUrn*) ▷ *Sample informant from collective urn*
      truth ← TruthValues[*claim_idx*] ▷ *0 or 1*
      rating_own ← Ratings[*claim_idx*, *Agents*[*k*].*id*]
      rating_select ← Ratings[*claim_idx*, *informant_id*]
      score_own ← (*rating_own* − *truth*)^2 ▷ *Brier score for own judgment*
      score_select ← (*rating_select* − *truth*)^2 ▷ *Brier score for selected informant's judgment*
      Append *score_own* to *Agents*[*k*].*self _score* ▷ *For analysis*
      if *score_select* < *score_own* **then** ▷ *Selected informant was better*
        Append *informant_id* to *Agents*[*k*].*urn* ▷ *Reinforce selected informant*
        Append *score_select* to *Agents*[*k*].*score_history* ▷ *For analysis*
      else ▷ *Own judgment was better or equal*
        Append *Agents*[*k*].*id* to *Agents*[*k*].*urn* ▷ *Reinforce self*
        Append *score_own* to *Agents*[*k*].*score_history* ▷ *For analysis*
  return *Agents* ▷ *Final state of all agents (urns, score histories)*

The main question we are aiming to answer is whether agents who are epistemically more successful (according to their overall classification correctness and Brier scores on the full dataset) come to be more frequently consulted. We use both measures of overall success. More specifically, to evaluate whether the model identifies the most successful agents, we calculate, after each round, the correlation between an agent's consultation frequency—how often they have been consulted by others, up to the given round—and their overall classification correctness, as well as the correlation between consultation frequency and their overall average Brier score. Note that for classification correctness higher is better while for average Brier scores lower is better. Thus, high positive correlation with classification correctness and high negative correlation with average Brier scores indicate that the most epistemically successful agents are also the most trusted. In addition to this, we track the relative centrality of experts versus that of non-experts by identifying the five most and the five least successful agents and compare how often, on average, each group is consulted by others.

Finally, to assess the added value of the reputational mechanism, we compare simulations using only individual urns with those in which there is a collective urn. For each variant of the model, we compute mean correlation values over multiple simulation runs (N = 1000) and test whether the differences are statistically significant using two-sample t-tests, also reporting Cohen's d values for effect size.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

4 Results

To address the question of the emergence of expert trust, we start by examining the model variant where agents learn strictly from their own interactions, updating only their private urns. We already saw the distribution of epistemic success scores on the full dataset of 1,200 claims in Figure 1. Figure 2 illustrates for a single, typical simulation run the evolution of the correlation between agents' consultation frequency and their pre-calculated, overall epistemic performance. As the number of runs increases, the correlation between how often an agent is consulted and their overall classification correctness steadily rises, here approaching approximately .81 toward the end. Conversely, the correlation between consultation frequency and overall average Brier score becomes increasingly negative, reaching about −0.53 in this run. This indicates that the simple reinforcement mechanism, even when localized to individual experience, enables the community to gradually identify, and hence to consult more frequently, its more epistemically successful members.

For the same simulation, the left panel of Figure 3 visualizes the relationship between an agent's overall classification correctness and their final consultation frequency after 1,000 runs, and the right panel does the same for average Brier scores; LOESS smoothers are added to highlight the trends, showing a strong positive trend on the left—agents scoring better on classification correctness tend to be selected more often—and a negative one on the right, indicating that agents with higher average Brier scores tend to be selected less often.

This trend of increasing deference to better-performing agents is also evident when comparing explicitly defined groups. Figure 4 tracks the average number of times the top 5 most expert and bottom 5 least expert agents are consulted by others, where expertise is understood both in terms of classification correctness (results shown in left panel) and in terms of average Brier scores (right panel). On both interpretations, the experts and non-experts are consulted with increasing frequency over the 1,000 runs, but the former come to be consulted much more frequently than the latter, resulting in a clear and widening gap emerging between the two groups.

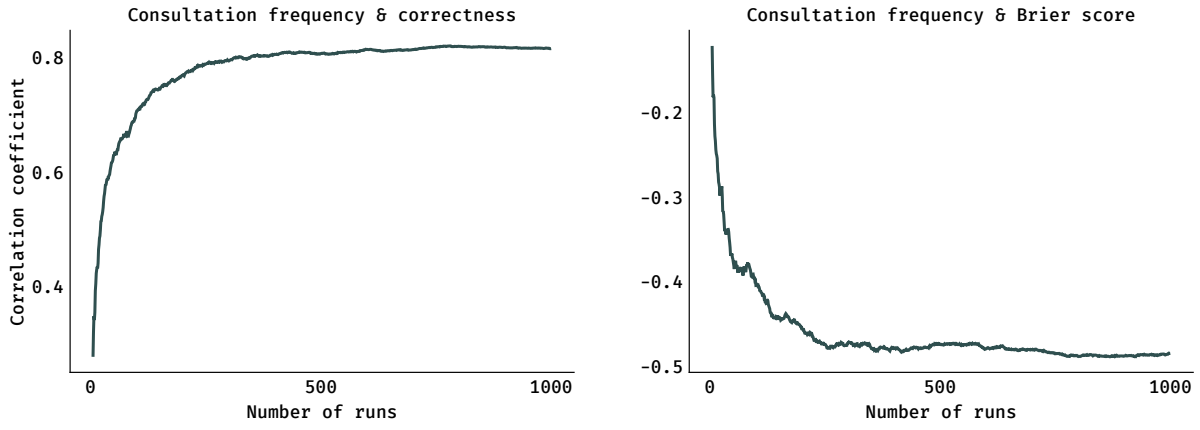

Figure 2: Evolution of performance-sensitive deference in a single simulation run (individual learning model). Correlation between agents' consultation frequency (up to the current run) and their overall classification correctness (left) and overall average Brier score (right), across 1,000 runs. Higher positive correlation with correctness and higher negative correlation with Brier score indicate better identification of expertise.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

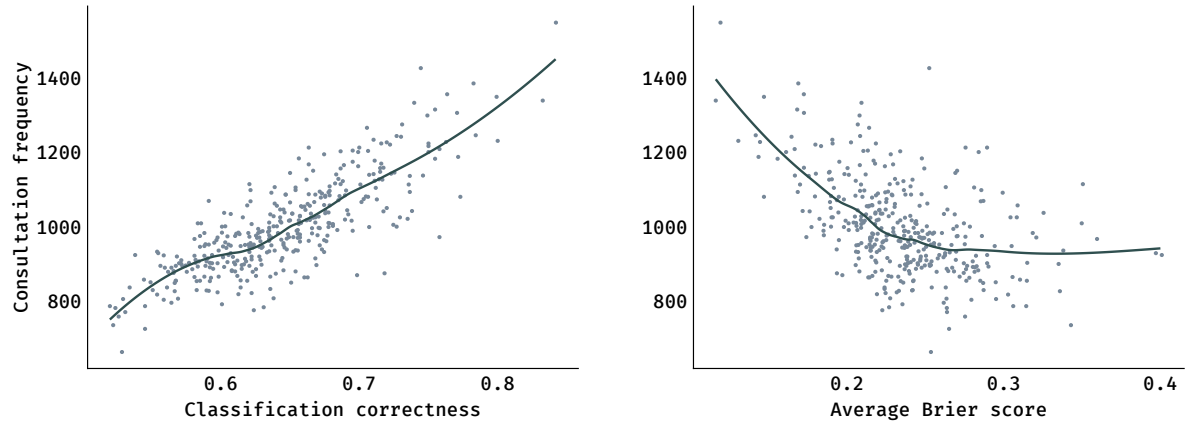

Figure 3: Relationship between overall epistemic performance and final consultation frequency after 1,000 runs in a single simulation (individual learning model). Left panel showing consultation frequency versus overall classification correctness; right panel showing consultation frequency versus overall average Brier score (with smoothers added to highlight trends).

Figure 5 illustrates this process of developing consultation preferences over time, focusing on a small subset of agents (to avoid clutter). More specifically, it presents a series of snapshots of the evolving consultation network at different stages of the simulation, where each panel focuses on the top 3 consultation choices made by 10 randomly preselected agents (which in the simulation that was used to produce the figure happened to be agents 36, 62, 78, 138, 170, 191, 218, 226, 237, and 329). Nodes (circles with IDs) represent individual agents, with node size proportional to the agent's global reputation, that is, how many times they have been reinforced as a good informant by any agent in the entire community up to that simulation run. Arrows indicate a directed consultation preference, an arrow from agent A to agent

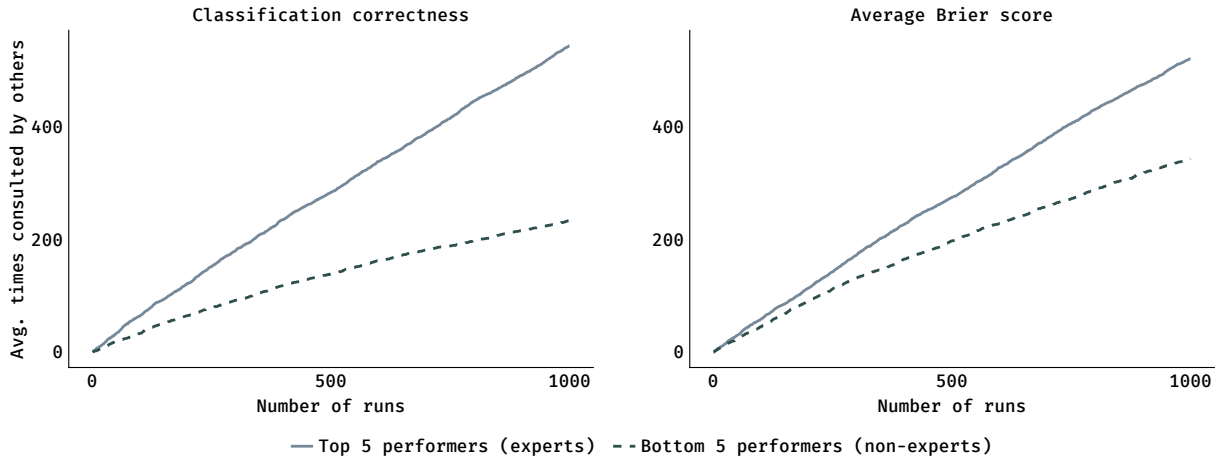

Figure 4: Emergence of expert–non-expert distinction in a single simulation run (individual learning model). Average number of times the top 5 (solid light-gray lines) and bottom 5 (dashed dark-gray lines) performers are consulted by others over 1,000 runs. Performance is based on overall classification correctness (left) and overall average Brier score (right). (NB: For Brier scores, lower is better, so "top 5" here means lowest 5 Brier scores.)

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

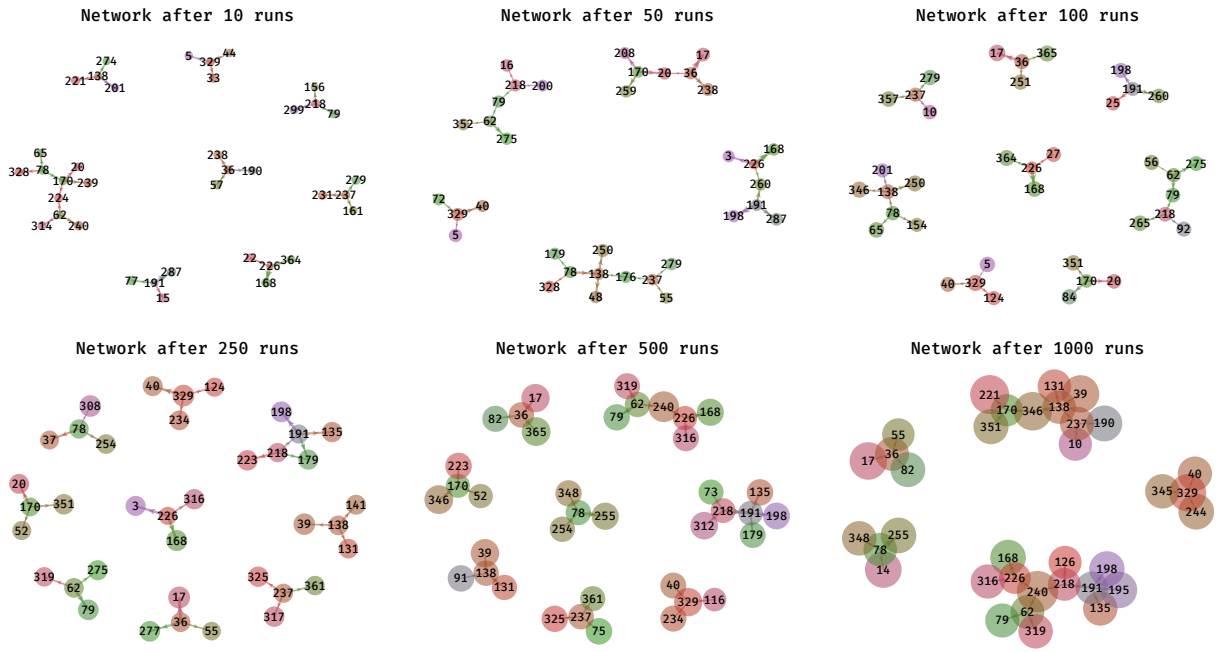

Figure 5: Snapshots of the evolving consultation network (individual learning model, single simulation run). Each panel shows the top 3 consultation choices made by 10 preselected agents at different simulation stages. Node size is proportional to the agent's global consultation frequency by all other agents. Arrows indicate directed consultation preference from a pre-selected agent. (Colors facilitate distinguishing among agents but are otherwise meaningless.)

B meaning that A (one of the 10 preselected) frequently chose agent B as a better informant. (Node colors serve only to distinguish individual agents.) The progression from one panel to the next shows how, from initially more diffuse interactions, patterns of deference emerge, with connections increasingly forming around more "popular" (frequently reinforced) nodes. We see, for instance, how at the end of the process agents 62, 191, and 218, which were among the preselected agents, are in the same clique, because agent 191 was among the agents most frequently consulted by agent 218, and agents 62 and 218 both frequently consulted agent 240 (which was not included in the preselection).

Although it is clear that a network forms, it is yet to be determined whether it does so to good effect. That it actually benefits agents can be seen in Figure 6, which shows the average difference, at each run, between the 50-run rolling mean of Brier scores agents *would* have achieved if they had always only consulted themselves (this is what is stored in the *self_score* container in Algorithm 1) and the 50-run rolling mean of the Brier scores they actually achieved (which are stored in the *score_history* container). Note that the improvement that comes from trusting others and building a network is most pronounced in the initial segment of the 1,000 runs, but even though it then diminishes somewhat, it remains substantial. The reason why the improvement is greatest at the start is that consulted agents get reinforced when they do better on a selected claim than the agent itself. This doing better will sometimes just be a lucky hit, of course, in which case the agent may reinforce another agent whose overall scores are actually worse than its own, meaning that later selections of that agent—which are now more likely to occur—may not help the agent to improve its actual score over its "self score."[5]

[5]It is to be noted that these consultation networks are *not* guaranteed to be optimal: simple reinforcement can "satisfice" (in

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

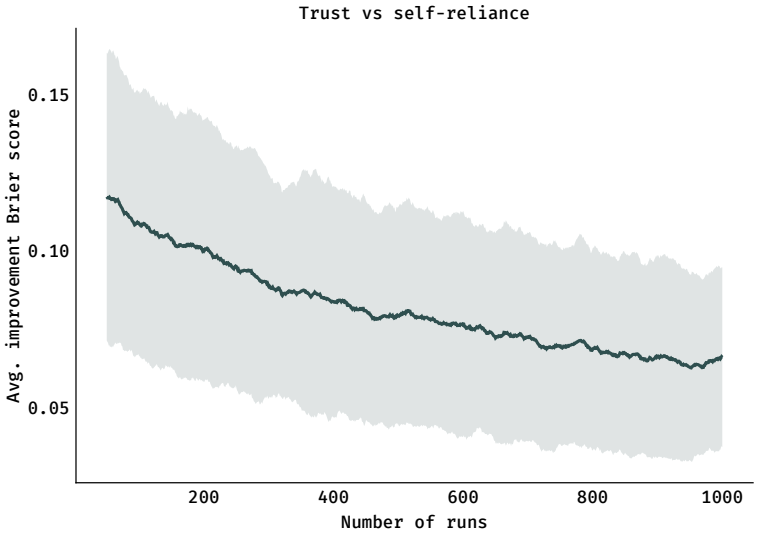

Figure 6: Average epistemic improvement from social learning in a single simulation run (individual learning model), showing the 50-run rolling mean (± SEM across agents at each run) of the difference between agents' Brier scores if relying on self-judgment ("self score") and their actual achieved Brier scores through consultation.

Turning now to the impact of shared reputational information, we investigate this by comparing the individual learning model with the collective urn model, in which (it will be recalled) agents sample from a pooled urn reflecting the aggregated reinforcement history of the entire community. We ran 1,000 full simulations (each lasting 1,000 runs) for both models.

The top row of Figure 7 shows the average correlation (averaged over the 1,000 simulations) between consultation frequency and overall classification correctness (left) and overall average Brier score (right), for both the individual (red solid lines) and collective/reputational (blue dashed lines) urn models. Both panels show that the collective urn model leads to a faster convergence on epistemically successful agents. Absolute differences may not seem large for classification correctness or average Brier scores, but two-sample t-tests performed after each run reveal that they were virtually always significant. Moreover, calculating Cohen's d values also after each run shows that the effect of adding reputational marking actually tends to be very large, as can be seen in the bottom row of Figure 7. (Conventionally, Cohen's d values above 0.8 count as large.)

These results demonstrate that while simple individual reinforcement learning can lead to the emergence of trust in more competent peers, the introduction of a mechanism to share reputational information significantly improves the speed and accuracy of this process. The community as a whole becomes better at identifying its most reliable epistemic members when individual learning experiences about who to trust are made publicly available.

the sense of Simon, 1982) on locally good but suboptimal patterns (e.g., a chain where an intermediate node is preferred over a still more reliable node). With some tuning, one may be able to prove convergence toward optimal structures in our setting by adapting results on reinforcement learning (see, e.g., Beggs 2005). We leave such theorems to future work and here focus on the descriptive emergence of performance-sensitive deference.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

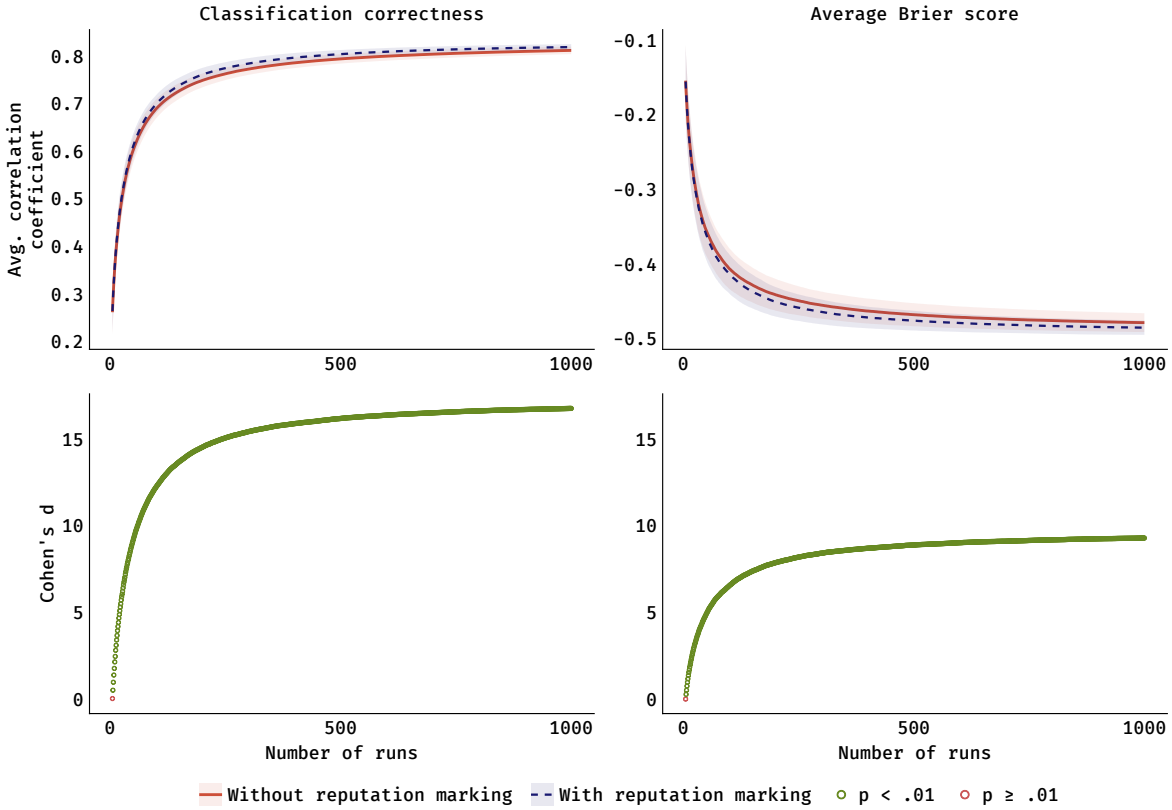

Figure 7: Impact of reputation on identifying expertise. Top row showing average correlation (± SEM across 1,000 simulations) between agent consultation frequency and overall classification correctness (left) and overall average Brier score (right), over 1,000 runs, for models with individual urns (red solid lines) and collective urns (blue dashed lines). Bottom panels show corresponding effect sizes (mostly very large here) for the difference between the two models (green markers indicate $p < .01$).

5 Conclusion

We studied the emergence of expert trust through simple reinforcement learning mechanisms by adapting an agent-based model from Barrett, Skyrms, and Mohseni (2019) and applying it to a large dataset of probability judgments. Our findings show that communities can develop patterns of trust in their more epistemically competent members without evaluating track records or having access to background information on reliability, or in fact without any member ever having to trust another member *blindly*.

We showed that even in a model of strict individual learning, where agents only reinforce choices based on their private interaction outcomes, a significant correlation emerges between an agent's overall epistemic performance (their classification correctness and average Brier score on the full dataset) and how frequently they are consulted, meaning that epistemically more competent agents become more central in a dynamically forming network. This addresses a fundamental question that arises about expert trust, which even precedes how novices might choose among already identified "putative experts," to wit, how certain individuals may come to be recognized as expert-like resources in the first place. Furthermore, introducing a mechanism for sharing reputational information—by pooling individual

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

reinforcement histories into a collective urn—had a clear positive impact on the speed and accuracy of this process of expertise identification.

Most importantly, while admittedly abstracting from many complexities of real-world expert assessment, our model offers an empirically grounded, bottom-up account of how expert trust can emerge without recruiting any sophisticated cognitive machinery. The trust that emerges is not based on explicit inductive reasoning about past success rates by individual agents, but is an indirect outcome of a simple behavioral reinforcement process, akin to habit formation.

Most work in network epistemology, particularly the agent-based modeling tradition, has centered on questions of the division of cognitive labor and the optimal structure of communication networks. Researchers have asked, for example, how a community of scientists should distribute their efforts among different research programs to maximize collective success; what the ideal level of connectivity is in a scientific community; and how different communication structures affect a group's ability to converge on the truth (Zollman 2007; Weisberg & Muldoon 2009; Rosenstock, Bruner, & O'Connor 2017; O'Connor & Weatherall 2019; Šešelja 2021; Ferrari, Lammers, & Wenmackers Forthcoming). Our work, in contrast, addresses a question that is arguably prior to these concerns, viz., how the very networks of trust and deference that underpin scientific collaboration and communication emerge to start with. One avenue for future research would be to combine our model of emergent trust with existing models of the division of cognitive labor and then investigate how the dynamic formation of trust networks, as modeled in our paper, affects a community's ability to solve complex problems or efficiently explore an "epistemic landscape." Such a combined model could help us answer whether communities that are good at identifying and trusting their most competent members are also more successful at allocating their cognitive labor effectively.

Reinforcement learning encompasses a great number of algorithms (see Sutton & Barto 2018), many of them more sophisticated than the one we adapted from Barrett and colleagues' work. Future research could consider using more complex (though still relatively simple) algorithms and conduct behavioral experiments to compare the ability of different such algorithms to model the emergence of expert trust. For example, in Figure 6 we saw how emerging reliance on others helps to improve the epistemic positions of agents, but also that this improvement decreases somewhat over time. As explained, that is because agents also tend to carry along some "dead wood" in their urns: peers who are, *overall*, epistemically inferior to an agent may nevertheless be reinforced by that agent when a claim is selected on which, coincidentally, the peer happens to do better. One could think of a further adaptation of the Barrett et al. model that allows agents to get rid of such dead wood, for instance, by endowing them with a second urn to which they add a ball with an ID from a peer occurring in their first urn each time that peer is sampled and thus consulted on a given claim but then turns out to have a higher Brier score than the agent itself. Once a peer has (say) three balls with its ID occurring in the second urn, the agent could purge the peer from the first urn (whether completely—by removing *all* balls with that peer's ID—or only partially—by removing *some* balls—if there is a difference). Note that such an adaptation would still not engage agents in any form of inductive reasoning, nor would it introduce a need for blind trust at any point.

There are other ways in which our approach could be refined. For instance, the current model (in either variant) does not look at the costs of consultation, or ways in which agents might argue for their positions or disclose their evidential sources—which might all figure in a decision whether or not to reinforce an agent. Addressing such issues would require the introduction of richer agent structures, which is a possibility we flag here only to set it aside as work for another day.

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

Meanwhile, our findings offer a proof of concept that simple and local learning rules may suffice to allow a basis for expert trust to emerge within a community of interacting agents. Nothing said in the foregoing goes against the main gist of Goldman's work on experts nor, to our knowledge, against that of any other main account of expertise. In particular, if there is a set of "putative experts" already in place, laypeople (or novices) may well want to decide *which* expert's advice to follow (in case not all experts agree) on the basis of the experts' track records and other quality indicators, as recommended by Goldman (2001, 2021) and others (e.g., Schurz 2012, 2019; Friedman & Šešelja 2023; though see Huang Forthcoming for critical discussion). What we hope to have offered is an approach to expertise that *complements* existing accounts, by highlighting how communities of initially *undifferentiated* agents—no one has been labeled as a putative expert yet—can self-organize into networks that effectively identify and place central their most capable members, and where all trust involved is *earned* and not *blind*.

Acknowledgments

For valuable comments on previous versions of this paper, we are greatly indebted to Christopher von Bülow and to two anonymous referees for this journal.

Funding statement

None to declare.

Declarations

None to declare.

References

Barrett, Jeffrey A. 2024. "Humean Learning (How to Learn)." *Philosophical Studies* 181:281–97. https://doi.org/10.1007/s11098-023-02086-3

Barrett, Jeffrey A., Skyrms, Brian, & Mohseni, Aydin. 2019. "Self-Assembling Networks." *British Journal for the Philosophy of Science* 70:301–25. https://doi.org/10.1093/bjps/axx039

Beggs, Alan W. 2005. "On the Convergence of Reinforcement Learning." *Journal of Economic Theory* 122:1–36. https://doi.org/10.1016/j.jet.2004.03.008

Bezanson, Jeff, Edelman, Alan, Karpinski, Stefan, & Shah, Viral B. 2017. "Julia: A Fresh Approach to Numerical Computing." *SIAM Review* 59:65–98. https://doi.org/10.1137/141000671

Burge, Tyler. 1993. "Content Preservation." *Philosophical Review* 102:457–88. https://doi.org/10.2307/2185680

Coady, Cecil A. J. 1992. *Testimony*. Oxford: Clarendon Press.

Coronato, Antonio, Naeem, Muddassar, De Pietro, Giuseppe, & Paragliola, Giovanni. 2020. "Reinforcement Learning for Intelligent Healthcare Applications: A Survey." *Artificial Intelligence in Medicine* 109:101964. https://doi.org/10.1016/j.artmed.2020.101964

DeepSeek-AI. 2025. "DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning." *arXiv* 2501.12948, https://arxiv.org/abs/2501.12948.

Douven, Igor. 2025. "Adaptive Updating: Ecological Rationality Meets Reinforcement Learning." *Minds & Machines* 35:32. https://doi.org/10.1007/s11023-025-09735-y

Erev, Ido, & Roth, Alvin E. 1998. "Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique Mixed-Strategy Equilibria." *American Economic Review* 88:848–81.

Ferrari, Sacha, Lammers, Wouter, & Wenmackers, Sylvia. Forthcoming. "How the Structure of Scientific Communities Could Impact the Public Uptake of Uncertain Science." *Philosophy of Science*. https://doi.org/10.1017/psa.2025.11

Foley, Richard. 1994. "Egoism in Epistemology." In F. Schmitt (Ed.), *Socializing Epistemology* (pp. 53–73). Lanham, MD: Rowman & Littlefield.

Foley, Richard. 2001. *Intellectual Trust in Oneself and Others*. Cambridge: Cambridge University Press.

Friedman, Daniel C., & Šešelja, Dunja. 2023. "Scientific Disagreements, Fast Science and Higher-Order Evidence." *Philosophy of Science* 90:937–57. https://doi.org/10.1017/psa.2023.83

Goldman, Alvin. 2001. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63:85–110. https://doi.org/10.1111/j.1933-1592.2001.tb00093.x

Goldman, Alvin. 2021. "How Can You Spot the Experts? An Essay in Social Epistemology." *Royal Institute of Philosophy Supplements* 89:85–98. https://doi.org/10.1017/S1358246121000060

Haarnoja, Tuomas, Zhou, Aurick, Abbeel, Pieter, & Levine, Sergey. 2018. "Soft Actor–Critic: Off-Policy Maximum Entropy Deep Reinforcement Learning with a Stochastic Actor." *Proceedings of the 35th International Conference on Machine Learning* (PMLR 80; pp. 1861–70). Stockholm.

Haas, Julia. 2022. "Reinforcement Learning: A Brief Guide for Philosophers of Mind." *Philosophy Compass* 17:e12865. https://doi.org/10.1111/phc3.12865

Havrilla, Alex, Du, Yuqing, Raparthy, Sharath C., Nalmpantis, Christoforos, Dwivedi-Yu, Jane, Zhuravinskyi, Maksym, Hambro, Eric, Sukhbaatar, Sainbayar, & Raileanu, Roberta. 2024. "Teaching Large Language Models to Reason with Reinforcement Learning." *arXiv* 2403.04642, https://arxiv.org/abs/2403.04642.

Huang, Alice C. W. Forthcoming. "Track Records: A Cautionary Tale." *British Journal for the Philosophy of Science*. https://doi.org/10.1086/728459

Niv, Yael. 2009. "Reinforcement Learning in the Brain." *Journal of Mathematical Psychology* 53:139–54. https://doi.org/10.1016/j.jmp.2008.12.005

O'Connor, Cailin, & Weatherall, James O. 2019. *The Misinformation Age: How False Beliefs Spread*. New Haven, CT: Yale University Press.

Rosenstock, Sarita, Bruner, Justin, & O'Connor, Cailin. 2017. "In Epistemic Networks, Is Less Really More?" *Philosophy of Science* 84:234–52. https://doi.org/10.1086/690717

Roth, Alvin E., & Erev, Ido. 1995. "Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term." *Games and Economic Behavior* 8:164–212. https://doi.org/10.1016/S0899-8256(05)80020-X

Schultz, Wolfram, Dayan, Peter, & Montague, P. Read. 1997. "A Neural Substrate of Prediction and Reward." *Science* 275:1593–99. https://doi.org/10.1126/science.275.5306.1593

Schurz, Gerhard. 2012. "Meta-Induction in Epistemic Networks and the Social Spread of Knowledge." *Episteme* 9:151–70. https://doi.org/10.1017/S1742360012000064

Schurz, Gerhard. 2019. *Hume's Problem Solved: The Optimality of Meta-Induction*. Cambridge, MA: MIT Press.

Šešelja, Dunja. 2021. "Exploring Scientific Inquiry via Agent-Based Modelling." *Perspectives on Science* 29:537–57. https://doi.org/10.1162/posc_a_00382

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press

Silver, David, Huang, Aja, Maddison, Chris J., et al. 2016. "Mastering the Game of Go with Deep Neural Networks and Tree Search." *Nature* 529:484–89. https://doi.org/10.1038/nature16961

Simon, Herbert A. 1982. *Models of Bounded Rationality*. Cambridge, MA: MIT Press.

Stinson, Patrick, van den Bosch, Jasper, Jerde, Trenton, & Kriegeskorte, Nikolaus. 2025. "Collective Inference of the Truth of Propositions from Crowd Probability Judgments." *arXiv* 2501.04983, https://arxiv.org/abs/2501.04983.

Sutton, Richard S., & Barto, Andrew G. 2018. *Reinforcement Learning: An Introduction* (2nd ed.). Cambridge, MA: MIT Press.

Weisberg, Michael, & Muldoon, Ryan. 2009. "Epistemic Landscapes and the Division of Cognitive Labor." *Philosophy of Science* 76:225–52. https://doi.org/10.1086/644786

Xu, Fengli, Hao, Qianyue, Zong, Zefang, et al. 2025. "Towards Large Reasoning Models: A Survey of Reinforced Reasoning with Large Language Models." *arXiv* 2501.09686, https://arxiv.org/abs/2501.09686.

Zollman, Kevin J. S. 2007. "The Communication Structure of Epistemic Communities." *Philosophy of Science* 74:574–87. https://doi.org/10.1086/525605

https://doi.org/10.1017/psa.2025.10186 Published online by Cambridge University Press