The Mix Matters: Exploring the Interplay Between Epistemic and Zetetic Norms in Scientific Disagreement

Martin Justin[1], Dunja Šešelja[2], Christian Straßer[3], and Borut Trpin[4]

[1]*University of Maribor, martin.justin1@um.si*

[2]*Ruhr University Bochum, dunja.seselja@rub.de*

[3]*Ruhr University Bochum, christian.strasser@rub.de*

[4]*University of Maribor, LMU Munich, and University of Ljubljana, borut.trpin1@um.si*

Abstract

What is the rational response to a scientific disagreement? Many epistemologists argue that disagreement with an epistemic peer should generally lead to conciliation by lowering confidence in the disputed belief or even suspending judgment altogether. Although this conciliatory approach is widely regarded as a norm of individual rationality, its value in the context of collective scientific inquiry is less clear. Some have even raised concerns that conciliating in scientific disagreements may slow progress or reduce the efficiency of inquiry. In this paper, we introduce a novel agent-based model that captures key aspects of scientific disagreement by incorporating both epistemic norms, which govern belief revision, and zetetic norms, which guide how scientists pursue inquiry. Our results indicate that the effects of conciliating in the face of disagreement—whether detrimental or beneficial—depend on the zetetic norms that scientists follow. When they focus on exploiting the hypothesis that they believe is most likely to succeed, remaining steadfast is more effective. However, with exploratory scientists, conciliation does not negatively affect group performance. These findings highlight the critical role of zetetic norms in determining the rational response to disagreement in scientific practice.

1 Introduction

How should scientists respond when they find themselves in disagreement with epistemic peers—those who are equally informed, equally competent, and have access to the same evidence? This question has long been at the heart of debates in epistemology. The standard approach, embodied in the Conciliatory View, holds that rational agents should conciliate, or revise their beliefs, in the face of peer disagreement (Bogardus, 2009; Christensen, 2007; Elga, 2007; Feldman, 2005, 2006; Kornblith, 2010; Matheson, 2009). According to this view, disagreement with an epistemic peer provides a defeater for one's belief, requiring a reassessment of one's confidence and, often, a move toward consensus.

Yet, despite the intuitive appeal of this position, the Conciliatory View has faced significant challenges, especially in the context of scientific inquiry. This is an important challenge because scientists working on the same problem or in the same research field often disagree (Dellsén and Baghramian, 2021; Lawler, 2024). Notably, while most critiques of the Conciliatory View focus on the individual rationality of agents, a growing body of work in the philosophy of science has highlighted the broader implications of conciliation for group inquiry. These critiques suggest that conciliating in the face of disagreement can have detrimental effects on collective epistemic goals, such as promoting a diversity of approaches or fostering productive divisions of cognitive labor (Cruz and Smedt, 2013; Elgin, 2010).[1] Agent-based models, such as those developed by Douven (2010), have bolstered these arguments, showing that conciliation may indeed hinder scientific progress, particularly when the speed of inquiry is valued over accuracy.

In this paper, we argue that these critiques, while valuable, overlook a crucial dimension of the debate: the role of zetetic norms—norms of inquiry—alongside epistemic (or doxastic) norms. While epistemic norms concern what one should believe in light of disagreement, zetetic norms govern how inquiry itself should proceed. Despite their importance, zetetic norms have been underexplored in the context of scientific disagreement. We argue that this oversight may explain why the question of the rational response to peer disagreement, particularly in scientific contexts, remains unresolved. Our key insight is that the tension between conciliation and successful inquiry does not only arise due to different epistemic goals but emerges from the interplay between epistemic and zetetic norms.

Our contribution is twofold. First, we critically examine Douven's influential agent-based model of scientific disagreement and argue that it fails to accurately represent the conciliatory responses

1. Skipper and Steglich-Petersen (2020) make a similar point with respect to group deliberation. They show that conciliation can frustrate a group's favored trade-off between error-avoidance and truth-seeking in making a collective decision.

typically discussed in the literature on peer disagreement. Second, we introduce a novel agent-based model, grounded in the bandit modeling framework (Zollman, 2007, 2010), to explore how both epistemic and zetetic norms shape the rational response to disagreement in scientific communities. Our findings show that the negative impact of conciliation on inquiry, as suggested by Douven, only arises when agents prioritize exploitative inquiry strategies—those that aim to maximize immediate epistemic payoff by pursuing hypotheses believed to be most likely true. In contrast, when scientists adopt more exploratory inquiry strategies, conciliation does not impede the success of group inquiry.

This paper is structured as follows. In Section 2, we present and critically assess Douven's model of scientific disagreement. Section 3 introduces our alternative model, and presents the results of our simulations, which highlight the role of zetetic norms in mediating the effects of conciliation. In Section 4, we discuss the broader implications of our findings for the epistemology of peer disagreement, before concluding in Section 5 with reflections on the importance of integrating zetetic norms into normative accounts of rational disagreement.

2 An opinion dynamics model of peer disagreement

Douven's (2010) argument against the Conciliatory View is based on an agent-based model (ABM) of peer disagreement, also reproduced by Douven and Kelp (2011). ABMs are computer simulations that enable us to study how the norms guiding individual agents impact their collective performance (Šešelja, 2022). Here, we are interested in how different norms of peer disagreement, which guide the behavior of individual agents, affect the epistemic performance of their group as a whole.

Douven based his work on the Hegselmann-Krause model of opinion dynamics. The model, first introduced by Hegselmann and Krause (2002), simulates a group of agents who repeatedly aggregate their beliefs to study the circumstances that lead them to consensus, polarization, or further fragmentation. At the beginning of a simulation, each agent is assigned an opinion, represented by a real number between 0 and 1. The model is round-based. In each round, the agents exchange their opinions with others who have sufficiently similar views to their own. Each agent then updates their opinions by averaging them with opinions of those who fall into one's 'confidence interval'. More precisely, agent i with opinion x_i will consider agent j's opinion x_j only if $|x_i - x_j| \leq \varepsilon$, where ε is a parameter of the model. Equation 1 describes the update of agent i's opinion at round $u+1$:

weights

$$\underbrace{x_i(u+1)}_{\text{posterior opinion}} = \alpha \cdot \underbrace{\frac{1}{|X_i(u)|} \sum_{j \in X_i(u)} x_j(u)}_{\text{conciliation with peers}} + (1-\alpha) \cdot \underbrace{(\tau + rnd(\zeta))}_{\text{noisy truth signal}}, \tag{1}$$

where $X_i(u)$ contains all agents' opinions that fall within the agent's i confidence interval, determined by ε, and $|X_i(u)|$ is the cardinality of $X_i(u)$. In Douven's variation of the model, based on Hegselmann and Krause's (2006), scientists track a truth signal modeled by the parameter $\tau \in (0,1]$. In simulation runs, agents not only communicate with each other, but also get a noisy signal about τ. ζ represents noise in the truth signal—agents do not learn the value of the signal directly but draw pulls from a uniform distribution of the interval $[\tau - \zeta, \tau + \zeta]$. This is meant to represent the fact that "in real life, we have to deal with measurement errors and other facts that may make our data noisy" (Douven, 2010, p. 151). α is a parameter that determines how much weight agents assign to their (and other agents') belief(s) on the one hand and to the truth-signal on the other. When $\alpha = 1$, agents ignore the truth signal. In an extreme scenario, where $\varepsilon = 0$ and $\alpha = 1$, agents never change their opinions since they cannot learn from other agents or from the truth signal. When $0 < \alpha < 1$ agents update both on the basis of previous opinions and the truth signal. Finally, when $\alpha = 0$, they update their opinions only based on the truth signal, disregarding even their own prior opinions.

Douven offers two interpretations of what an opinion represents in his model. On the first, it reflects a belief about the true value of a parameter τ; splitting the difference then means that if one agent believes $\tau = x$ and another believes $\tau = y$, both adopt $(x+y)/2$ as their new belief (Douven, 2010, p. 140). On the second interpretation, opinions represent credences in a hypothesis H, so disagreement is resolved by averaging the credences (e.g., from .6 and .4 to .5). The first interpretation—adjusting the content of a factual belief—deviates from how most epistemologists understand conciliation, which typically concerns revising one's confidence in a proposition, not its content. The second interpretation aligns more closely with standard views, modeling disagreement as a difference in first-order uncertainty. Still, a challenge remains: what exactly does the parameter τ represent if credences are to be interpreted as degrees of belief in a proposition? We return to this issue in Section 3.2, where we extend our model to accommodate first and higher-order uncertainty.

Douven uses combinations of different values ε and α to model various ways of reacting to peer disagreement. On the one hand, agents who do not learn from others ($\varepsilon = 0$) and consider their priors and the new information about the world to an equal extent ($\alpha = 0.5$) represent reasoners who follow the Steadfast View. On the other hand, agents who learn from others ($\varepsilon = 0.1$) and give more weight to the aggregate of their and their peers' priors ($\alpha = 0.9$) represent reasoners who follow the Conciliatory View.

In his model Douven compares how successful groups of different agents are in converging to the truth signal. His results, replicated in Figure 1, indicate a trade-off between speed and accuracy. Steadfast agents "get within a moderate distance of the truth relatively quickly, but for them that is about as good as it gets in terms of truth closeness" (Douven, 2010, p. 151). On the other hand, conciliatory agents move closer to the true value more slowly, but "in the somewhat longer run they are on average much closer to the truth" (Ibid.). More precisely: "communities of agents who may obtain imprecise information about τ end up on average being closer to τ for higher values of both

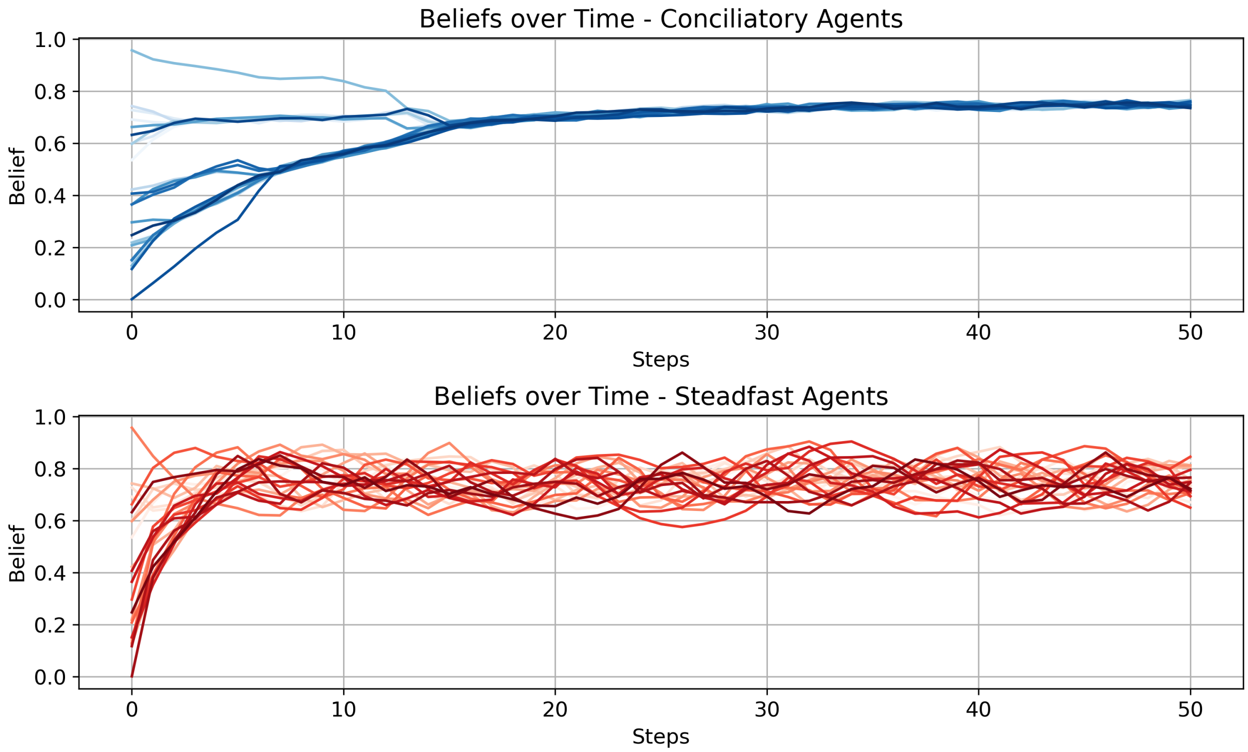

Figure 1: Reproduction of the results by Douven (2010). The x-axes represent time as steps in the simulation, while the y-axes represent opinions. Each line represents how one agent's opinions change over time. All charts in this section are of this format. While each chart only shows the initial 50 rounds of the given simulation, the results remain stable for up to 10,000 rounds.

α and ε. However, with lower values for these parameters, they on average tend to get faster to a value that is at least moderately close to τ" (Douven and Kelp, 2011, p. 276). See also Figure 1.

These results show that conciliating with disagreeing peers can hurt the speed of inquiry. As Douven argues, this implies that an appropriate response to peer disagreement depends on whether we prioritize speed of inquiry over accuracy of opinions. Since our priorities may depend on context, these results suggest that there are no "general and illuminating things to be said about what we ought or ought not to do in cases of peer disagreement" (Douven, 2010, p. 156).[2]

This is a notable result already because it introduces the assessment of epistemic norms from the perspective of collective inquiry. That said, the above model has important limitations. First, the

2. Langhe (2013) presents another ABM based on the Hegselmann-Krause model of opinion dynamics, to study disagreements between agents working within different epistemic systems. The latter concern methodological principles and considerations that scientists use when studying the same problem, for example, Marxist and neoclassical approaches to economy. Since this presents an explicit departure from the peer disagreement debate, we do not discuss this model here.

agent's reasoning is skewed towards recent evidence, and as such not entirely rational.[3]

In addition, the performance of agents largely depends on the value of α, which is problematic since the specific α-values, used to generate the above results, are not well motivated by the literature on peer disagreement. We elaborate on these points in the Appendix.[4]

Here, we wish to draw attention to a more general issue. Arguments against conciliation in the context of group inquiry often make the mistake of conflating the dimensions of belief revision and of inquiry. Cruz and Smedt (2013, p. 175), for example, argue that remaining steadfast can lead scientists "to uncovering new evidence" and directs them "to a closer scrutiny of existing evidence and assumptions". However, both uncovering new evidence and scrutinizing existing evidence are forms of inquiry, not belief states. When Cruz and Smedt (2013) talk about the epistemic value of sustained disagreement, they do not refer to the value of remaining steadfast in one's beliefs but to the value of remaining steadfast in one's line of pursuit or inquiry.

However, it is not self-evident that scientists do or should simply inquire what they believe to be the true or most likely hypothesis.For example, scientists can suspend their judgment about a hypothesis because a peer disagrees with them, but at the same time continue to gather evidence on it (Fleisher, 2018). In addition, there has recently been a shift of attention to norms of inquiry in epistemology more broadly. Some philosophers have argued that to better understand knowledge and belief formation, we should consider not only norms of belief update but also the inquiry process and the norms that guide it (Falbo, 2023; Friedman, 2020; Thorstad, 2021). This shift has sometimes been termed the "zetetic turn" in epistemology.

Although Douven's model does not aim at studying norms of inquiry, it exhibits a similar limitation. It represents the dimension of inquiry in a simplified manner, where all agents gather information from the truth signal in the same way. This raises the question of whether using different modeling assumptions to represent, on the one hand, the dimension of belief formation and, on the other hand, the dimension of inquiry would yield the same insights about the impact of conciliation

3. The inclusion of the truth signal is what sets Douven's model apart from simple contagion models (Oliveira et al., 2022; Piedrahita et al., 2018), which generally do not incorporate an evidential aspect and rather view belief change as a process of social transmission or imitation.

4. In recent years, Douven presented several more complex models that improve upon this simple one. Douven (2019) and Douven and Hegselmann (2021) introduced agents who are rational Bayesian reasoners, while Douven and Hegselmann (2022) combined bounded-confidence social learning with the bandit problem. Most recently, Douven (2025) presented an extension of the Hegselmann-Krause model where agents' reasoning is modeled by artificial neural nets. These models are much more sophisticated than the one presented above. However, they do not discuss peer disagreement explicitly.

on collective inquiry.

3 A bandit model of peer disagreement

To better understand how conciliation affects group inquiry, we present a novel ABM of peer disagreement based on the bandit modeling framework. In this section, we will first give a short introduction to bandit models in general, motivate our choice of this modeling framework, and then present our model in more detail.

3.1 A brief primer on bandit models

Bandit models, as now commonly employed in the philosophy of science based on the seminal work by Bala and Goyal (1998), consist of two model entities, bandits (or slot machines), which represent scientific theories and agents, who stand for scientists. Each bandit, B_i, has two possible outcomes, success and failure, and a predetermined success rate, $0 \leq s_i \leq 1, s_i \in \mathbb{R}$. Agents don't know these values but can assess them by testing the bandits. Thus, every agent A_j, has a subjective probability of the success rate of each bandit, c^i_j. For a given bandit B_i and an agent A_j, the value of this subjective probability is given as a beta distribution, where both the mean and the variance play an important part.[5] The mean of the distribution represents the agent's numerical assessment of the bandit. This is shown shown by Equation 2,

$$c^i_j = \alpha^i_j / (\alpha^i_j + \beta^i_j), \tag{2}$$

where α^i_j is the number of successful tests of the bandit, and β^i_j is the number of unsuccessful tests. The variance, on the other hand, represents agent's confidence in this assessment. For example, two agents, A_1 and A_2, might assign the same success rate to the bandit, $c^i = 0.5$, but where $\alpha_1 = \beta_1 = 10$ while $\alpha_2 = \beta_2 = 10,000$. In this case, the variance of A_2's distribution is much smaller, making the agent much more "resilient" in the sense that new evidence would cause smaller belief changes (Skyrms, 1977). Initially, the agents' prior subjective probabilities are generated by randomly drawing their α's and β's from a uniform distribution on a given interval.

5. In contrast to Zollman (2010), Zollman (2007) does not rely on beta distributions to model agents' beliefs. In that model, agents already receive information about the payoff of an action, which is randomly drawn from a distribution, and their beliefs are updated via Bayes' rule (see Rosenstock et al., 2017 for a detailed presentation of the model used in Zollman, 2007). O'Connor and Weatherall (2018) use the same procedure but also add agents who update via Jeffrey conditionalization.

In addition to testing the bandits, agents can also share information among themselves. The way agents exchange information is determined by the network structure of the model. The three most common network structures are the *complete graph*, in which all agents are connected, *the circle*, in which every agent is connected only with its two neighbors, and *the wheel*, which is similar to the circle but has one additional fully connected agent in the center of the network. Of course, different networks can also be randomly generated based on a desired density (Zollman, 2007). Agents use this shared information to update their subjective probabilities in the same way as described in Equation 2.

Simulations proceed in rounds where each round consists of four steps:

1. Agents choose a bandit.
2. Agents test the chosen bandit.
3. Agents share their data.
4. Agents update their credences (i.e., their subjective probabilities).

The simulation ends when a stopping condition is met. The latter depends on the specific set-up and goals of a concrete model. For example, in Zollman (2007), agents already start with a correct belief about the worse of the two bandits, so choosing that bandit gives them no new information. Thus, he implements two stopping conditions: the simulation stops (1) if all agents converge to the worse of the bandits or (2) if all agents come to form the correct belief with a probability greater than 0.9999. Meanwhile, Wu (2023), whose model is similar to Zollman (2007) but also allows for permanent polarization between two subgroups of agents, adds an additional condition. In her model, the simulation also stops if all agents in one subgroup become highly confident in the correct belief while all agents in the other subgroup converge to the worse bandit.

A simulation run has three possible outcomes: (1) *correct convergence*, (2) *incorrect convergence*, (3) *polarization*. The first outcome is achieved if, at the end of the simulation, all agents choose to test the bandit with the highest probability of success. The second outcome is achieved if, at the end of the simulation, all agents choose to test one of the other, sub-optimal bandits. The last outcome is achieved if, at the end of the simulation, agents still decide to test different bandits. The first outcome represents successful collective inquiry. More precisely, the epistemic performance of a group can be calculated as the number of simulations that end with correct convergence divided by the number of all simulation runs. Notice that this metric doesn't distinguish between incorrect convergence and polarization; both are simply understood as failures. Alternatively, the epistemic performance of a group can also be measured in terms of the speed of convergence, i.e., by the number of rounds it takes for a group to converge, correctly or incorrectly.

Since their introduction to the philosophy of science by Kevin Zollman (2007), bandit models have been used to study a wide variety of problems, from demographic diversity (Fazelpour and Steel, 2022), conformity (Weatherall and O'Connor, 2021), and epistemic benefit of marginalized social positions (Wu, 2023) and confirmation bias (Gabriel and O'Connor, 2024) to scientific polarization (O'Connor and Weatherall, 2018; Weatherall and O'Connor, 2021), epistemic effects of biased research (Holman and Bruner, 2017; Holman and Bruner, 2015; Weatherall et al., 2020), diagnosticity of evidence (Michelini et al., 2023), restricting dual-use research (Wagner and Herington, 2021) etc. As we show below, they are also useful for studying peer disagreement in the context of group inquiry. In contrast to the model by Douven (2010) and Douven and Kelp (2011), they allow a clear separation between different stages of inquiry, namely, choosing a hypothesis, collecting data, and communicating with other group members.

This is important for two reasons. First, it means that updating beliefs based on first-order evidence and sharing higher-order evidence of disagreement can be implemented as two separate, sequential steps. This more closely resembles situations described in the literature on peer disagreement. Second, it allows for an independent representation of epistemic (norms that deal with agents' beliefs) and zetetic norms (norms that deal with pursuits), which are not distinguished in models based on the Hegselmann-Krause's framework (Šešelja, 2021). The first kind of norm, which includes Epistemic norms, including norms of responding to peer disagreement, can be implemented by changing the agents' updating function. Zetetic norms, on the other hand, can be represented by modifying how agents decide which bandit to test at the beginning of each round.

3.2 Modeling peer disagreement with bandits

To model peer disagreement with a bandit model, we extended the basic framework in two ways. First, in the extended model, agents also share their credences about bandits (i.e., *beliefs*).[6] To do this, the protocol of each round (see Sec. 3.1) is extended in the following way:

1. Agents choose a bandit, according to one of the zetetic norms (explained below).
2. Agents test the chosen bandit.
3. Agents share their data.
4. Agents update their credences (subjective probabilities).
5. Agents share their credences about both bandits.
6. Agents update their credences based on beliefs of other agents, according to one of the epistemic norms (explained below).

Modeling epistemic norms. The above adjustments allow us to include different norms of responding to disagreement in the model, specifically two norms that follow from the Conciliatory View (which we will call *Belief Conciliation* and *Resilience Conciliation*) and the norm that follows

6. Santana (2021) similarly used a bandit model to study intragroup disagreements. However, in his model, agents learn about each other's pursuits (i.e., which bandit they are testing), not beliefs.

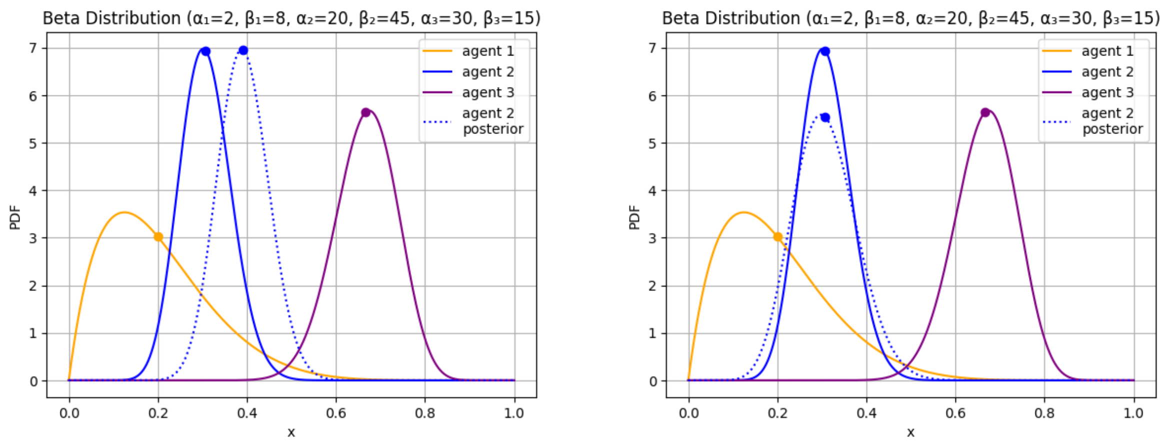

Figure 2: Update in a scenario with three agents and their beliefs in a given bandit represented by the underlying beta-distribution. The dotted line is the belief of agent 2 after the update. Left: Update according to Belief Conciliation. Right: Update according to Resilience Conciliation.

from the Steadfast View (which we will call *the Steadfast Norm*). Let's consider in more detail how these norms are implemented in the model.

The Steadfast View requires agents to stick to their beliefs in the face of disagreement. Agents in the model who act in accordance with the steadfast norm, simply ignore the beliefs of other agents. In other words, they act exactly the same way as agents in the base bandit model.

The Conciliatory View requires agents to dissolve their disagreements by meeting disagreeing peers halfway. As there are different ways to interpret what "meeting one halfway" means, different norms can follow from this view. We implemented two such norms. First, Belief Conciliation is a norm that requires one to adjust one's beliefs (credences) about the bandits. For this, agents take an average of their own and other agents' credences about both bandits following Equation 3,

$$c_{avg} = \frac{1}{N} \cdot \sum_{i=1}^{N} c_i, \tag{3}$$

where N represents the number of neighbors and c_i represents the belief of a given agent, A_i.

Since agents' beliefs are given as a beta distribution, agents must change their α- and β-values after adopting a new belief. This is done in such a way that the variance of the distributions remains the same, and therefore also the agents' resilience towards new evidence (see below for more on this notion). For an agent A_i the new α_i, i.e. α_i' and the new β_i, i.e. β_i', are given by Equation 4:

$$\alpha_i' = \left(\frac{1-c_i'}{\sigma_i^2} - \frac{1}{c_i'}\right) \cdot c_i'^2 \qquad \text{and} \qquad \beta_i' = \alpha_i' \cdot \left(\frac{1}{c_i'} - 1\right). \tag{4}$$

We provide an example of an update in Figure 2.[7]

This implementation of the Conciliatory Norm is an extension of *Straight Averaging*, as presented by Jehle and Fitelson (2009, p. 284). As Jehle and Fitelson (2009) point out, Straight Averaging is not without problems—it may create incoherent sets of beliefs once applied to scenarios that involve both peer beliefs and beliefs that are not governed by proper peerhood.[8] However, they also admit that it is the most natural formal interpretation of the norm that follows from the Conciliatory View of peer disagreement, which is also employed in Douven's model.

One might worry that our model does not adequately capture the conciliationist perspective in the peer disagreement literature, since our agents appear to reconcile by averaging point estimates of a bandit's payoff—thereby adjusting the content of their beliefs rather than their uncertainty. However, the agents' credences can be naturally reinterpreted as expressing uncertainty about the proposition "the next pull from this bandit will be a 1." On this interpretation, when agents split differences, they are conciliating over differing degrees of belief in a future event, aligning more closely with the way conciliation is typically understood in epistemology: as an adjustment of uncertainty in light of peer disagreement rather than a direct adjustment of factual estimates.

Second, we implement an alternative interpretation of the Conciliatory View: Resilience Conciliation. This norm, first suggested by Steglich-Petersen (2019), states that agents could respond to defeating higher-order evidence not by changing their beliefs but by undermining their beliefs' *resilience*. By resilience of beliefs, Steglich-Petersen (2019, p. 214) refers to the beliefs' sensitivity

7. Note that with *Belief Conciliation*, agents share only the mean of the beta distribution representing their beliefs. While sharing their full beliefs would also include sharing the variance (see Section 3.1), we have omitted this step since all agents have access to the same evidence, making the differences in the variance between agents, which result from their priors, minimal. For example, consider an extreme case where agent A starts with $\alpha_0 = \beta_0 = 1$ and agent B starts with $\alpha_0 = \beta_0 = 250$ for the same bandit. Here, A's variance is approximately 16.6 times higher than B's. However, this difference gets much smaller when they receive evidence. Let's say there are only three agents in the model, each taking 1000 pulls from the bandit, and, for simplicity, we assume that the means stay 0.5. Then, after the update A's $\alpha_1 = \beta_1 = 1501$ and B's $\alpha_1 = \beta_1 = 1750$, meaning that A's variance is already only 1.17 times higher. The difference gets even smaller each round or with each additional agent sharing more evidence.

8. This can lead to incoherence because the averaging procedure only governs updates to the disputed propositions. For example, two agents might be epistemic peers with respect to p and average their credences in p accordingly. However, because the rule is silent on how to update credences in other propositions related to p, those may remain unchanged, resulting in a probability distribution that no longer sums to 1.

to new evidence. He argues that following this norm is rational in cases where you get some higher-order evidence that suggests that you might have made a mistake but not how you should revise your belief. While traditional cases of peer disagreement are not like that—you should move in a specific direction, i.e., closer to your peer—group disagreements can be. Imagine that you are a scientist working in a larger group. In a situation where some of your colleagues think that p while others support an alternative theory, there is no clear direction in which to move your belief. Resilience Conciliation seems to be a rational response in such cases.

While Steglich-Petersen (2019) does not present a precise formal account of this idea, there is an intuitive way to represent the resilience of beliefs in our model. Since agents' assessments of the bandits are represented using beta distribution, we can make an assessment more or less resilient to new evidence by changing the α and β values (i.e., the number of times an agent has tested a bandit) while keeping the mean of the distribution (i.e., the beliefs about the success rate of a bandit) the same. Two agents A_1 and A_2 can have the same belief about a given bandit, but if A_1 based it on only 10 tests while A_2 based it on 1000 tests, A_1's belief will be more susceptible to change under new evidence. Using this simple mechanism, we model Resilience Conciliation in the following way. After updating on evidence, agents observe each other's beliefs about both bandits. If they disagree, they increase their variance proportionally to the size of the disagreement. Their credences remain unchanged where their α and β values change in accordance with Equations 5 and 6, respectively.

$$\alpha_i' = c_i \cdot \Big((\alpha_i + \beta_i) - w \cdot \sum_{j \neq i} |c_i - c_j|\Big), \tag{5}$$

$$\beta_i' = (1 - c_i) \cdot \Big((\alpha_i + \beta_i) - w \cdot \sum_{j \neq i} |c_i - c_j|\Big). \tag{6}$$

These two equations first lower the sum of an agent's α and β values for a given bandit proportional to the weighted distance in credences with the disagreeing agent and then recalculate them to keep the credence unchanged. The weighting factor w is needed since the distances in credences in the model are quite small, while the α and β values are much larger. w represents the significance that agents assign to their disagreements—larger w values means that they take even very small differences in credences as important. In Figure 2 we provide a simple example of an update.

Modeling zetetic norms. The second extension in our model concerns the way agents decide which bandit to test at the beginning of each round. This dimension of their behavior represents different norms of inquiry or zetetic norms. In most of the existing bandit models, agents follow the so-called myopic strategy and simply pursue the bandit they think has the highest success rate. We will call this the *Credence-Based Norm of inquiry*. Despite this norm being the most commonly used in the literature, alternative strategies have also been suggested. Kummerfeld and Zollman (2016), for example, describe the so-called *epsilon greedy strategy* in which agents can test different bandits in the same rounds. Frey and Šešelja (2020), on the other hand, introduce a strategy called *rational inertia*, in which agents do not immediately abandon their current bandit in face of evidence showing the superiority of the rival, but stick to it for a certain number of rounds.

In our model, we implement these three ways of conducting an inquiry. Since the implementation of the Credence-Based Norm is trivial—in each round, agents simply choose to test the bandit they believe has the highest success rate—we will only discuss the other two in a bit more detail.

Let us first look at the Epsilon Greedy Norm for inquiry. This norm is similar to the Credence-Based Norm in that agents mostly pursue the bandit they believe has the highest probability of success. Unlike with the Epsilon Greedy Norm, there is a small chance, $\varepsilon \in (0, 0.5)$, that the agents will test one of the other bandits. Since there are only two bandits in our model, the agents test the preferred bandit with $1 - \varepsilon$ probability and the other with ε probability (Kummerfeld and Zollman, 2016, p. 1061). ε is a parameter of the model. See Figure 3 (left) for an illustration.

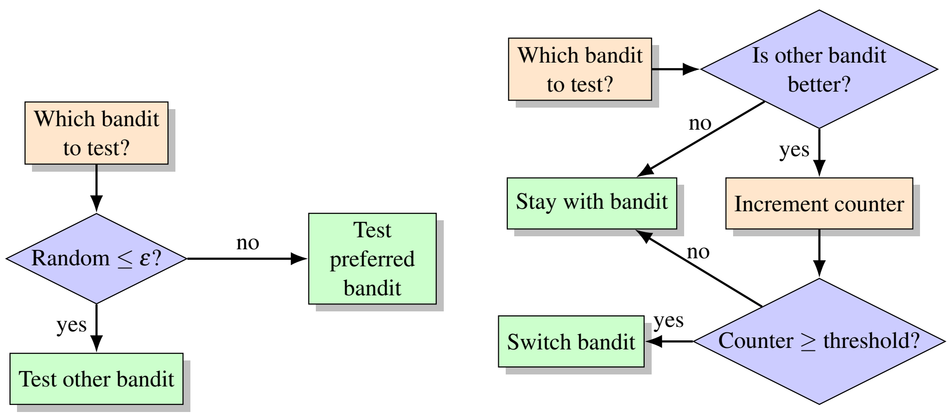

Figure 3: Decision procedures of the Epsilon-Greedy Norm (left) and Rational Inertia Norm (right).

In contrast to the epsilon-greedy agents who engage in exploratory behavior, agents who follow the Rational Inertia Norm for inquiry are exploration-averse. These agents decide to test the other bandit only after it has proven to have a better success rate for a certain number of rounds. Otherwise, they stubbornly stick to the bandit they have been testing before (Frey and Šešelja, 2020). More precisely, when deciding which bandit to test, they first check which of the two bandits has, in their opinion, a higher probability of success. If that is the bandit they have tested in the previous round, they simply continue to do so. If they think the other bandit is better, they still decide to test the same bandit as in the previous round, but they also start a counter. If the bandit that is not tested, turns out to be better for a certain number of rounds, they finally switch. The number of rounds that they wait is a parameter of the model. See Figure 3 (right) for an illustration.

Table 1 lists all the norms implemented in our model.

Epistemic Norms	Zetetic Norms
Steadfast Norm Credence Conciliation Resilience Conciliation	Credence-Based Norm Epsilon Greedy Norm Rational Inertia Norm

Table 1: List of epistemic and zetetic norms implemented in the model

3.3 Parameters used in the simulations

To test how different norms of responding to peer disagreement affect group performance, we compare how agents following different epistemic norms perform under different norms of inquiry. Specifically, we run simulations for all possible combinations of the three epistemic norms with the three zetetic norms. We do so for groups of 3 to 12 agents, who communicate through a fully connected network. Other parameters are kept constant throughout all simulation runs. The success rates of the two bandits are fixed at 0.501 and 0.500, the number of tests each agent performed per round is 1000, and agents' prior α and β are taken from a uniform distribution of a closed interval between 1 and 250. Each simulation is carried out for $10,000$ rounds with no additional stopping conditions, where each round consists of the six steps outlined above.

The values for the success rate of the two bandits, the tests per round, and the number of rounds are modeled after Zollman (2010). The prior values are chosen to ensure the right amount of disagreement between the agents at the beginning of the simulation. Since prior values determine the possible range of the initial α- and β-values of the agents, smaller values mean more similar prior subjective probabilities of the agents. To ensure some initial disagreement, we choose values that are a bit higher than usual—Zollman (2010), for example, uses $[0,4]$—but not so high that they cause agents to remain polarized for the whole simulation length of 10,000 rounds (which could happen if they follow the Credence-Based Norm).

Three parameters are specific to our zetetic norms. For the *Resilience Conciliation* norm, we set $w = 10{,}000$, modeling agents who lower confidence even in response to small disagreements. For the *Epsilon Greedy* norm, we set $\varepsilon = 0.005$, indicating infrequent exploration. For the *Rational Inertia* norm, we set $i = 10$, following Frey and Šešelja. We also test the *Belief Conciliation* norm with bounded confidence, where agents update only if $|c_i - c_j| \leq 0.1$. This yields 240 parameter combinations, each run $1{,}000$ times.

4 The tension between conciliation and successful inquiry

In this section, we will discuss our results, relate them to the broader literature on scientific disagreements, and compare them to the findings in Douven (2010) and Douven and Kelp (2011).

4.1 Steadfast Norm helps myopic agents

Let us first look at Figure 4, which shows the epistemic performance of conciliatory and steadfast agents under the Credence-Based Norm, representing myopic agents.

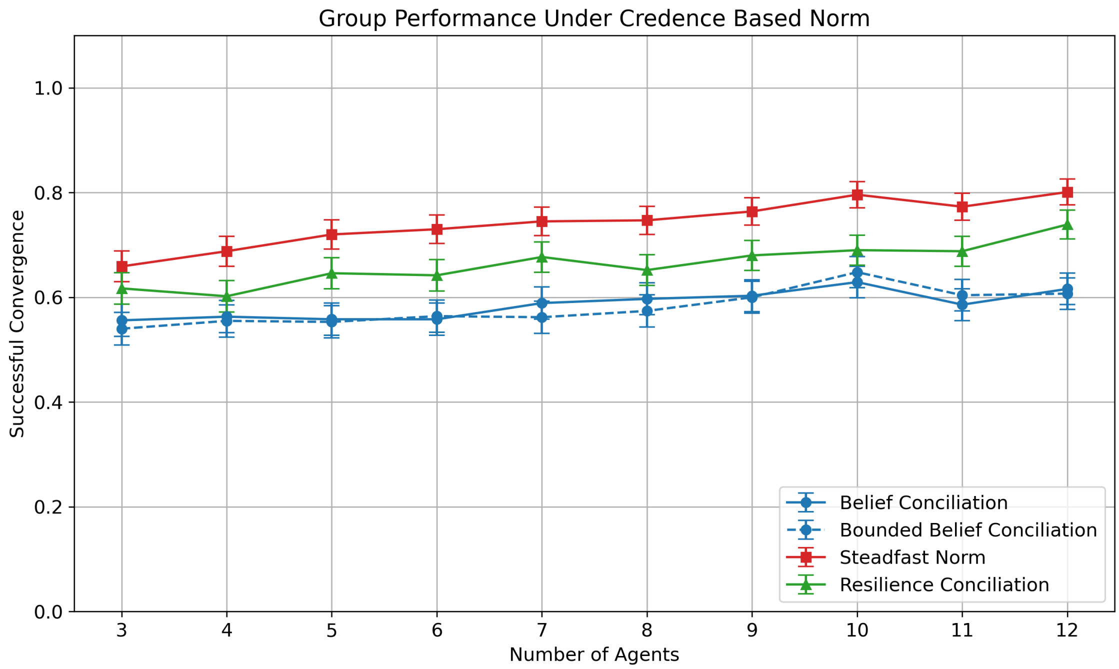

Figure 4: Epistemic performance of steadfast and conciliatory agents under the Credence-Based Norm at Round 10,000. The x-axis represents different group sizes and the y-axis the frequency of correct convergence. The error bars represent the 95% confidence interval. All of the following charts are of this format.

What we can see is that agents who split the difference with their peers perform significantly worse than agents who simply ignore their disagreeing peers. There is a natural explanation for this effect. Agents who follow the Belief Conciliation resolve their disagreements by aggregating their credences every round. Under the Credence-Based Norm for inquiry, they always test the bandit they think has the highest probability of success. Because conciliatory agents never disagree, this means that they always test the same bandit. This makes them much more vulnerable to misleading evidence. Consider the following scenario: It can happen that, in one of the early rounds, agents get evidence that convinces most of them that Bandit 2 is the better of the two (although it is actually worse). When they aggregate their credences at the end of the round, they all end up adopting this belief. Consequently, they all start testing Bandit 2 and do not get any new information about Bandit 1. Thus, they are unlikely to ever learn that Bandit 1 is actually the better one. A similar explanation holds for agents who follow Resilience Conciliation. By lowering their confidence, these agents make their beliefs much more prone to change and acutely responsive to evidence. Consequently, they all quickly adopt similar beliefs, although they don't explicitly aggregate them.

The negative effect of agents' homogeneity on the epistemic performance of a group as a whole is a well-established in bandit models (Zollman, 2007, 2010). It has also been illustrated in other contexts, such as the division of cognitive labor in scientific communities (Kitcher, 1990). As a consequence, remaining steadfast in the face of disagreement can be seen as a way of ensuring transient diversity and improving the inquiry of a group. This adds to other such mechanisms discussed in Smaldino et al. (2024) and Wu and O'Connor (2023).

The result that conciliation hurts inquiry also connects well with some previous work in the philosophy of science. Using a case study from paleontology, Cruz and Smedt (2013) have, for example, argued that sustaining disagreement, instead of suspending judgment, is epistemically beneficial in science. Specifically, they argue that remaining steadfast in their belief can encourage scientists to look for and generate new evidence. In addition, it can help them to re-evaluate the evidence and overcome confirmation bias. They write: "Epistemic peer disagreement has the advantage that it forces one to pay attention to anomalous data that one initially failed to detect or had glossed over as a result of confirmation bias. Moreover, disconfirmation bias leads scientists to be critical and especially vigilant to their opponents' arguments and evidence" (Cruz and Smedt, 2013, p. 176). This insight is confirmed by the results summarized in Figure 4: agents who remain steadfast in the face of disagreement are less likely to succumb to misleading evidence.

These results can also be interpreted as an example of the Independence Thesis (Mayo-Wilson et al., 2011). The Independence Thesis is a conjunction of two claims: (1) that rational individuals can form irrational groups and (2) that irrational individuals can form rational groups. As argued above, conciliating with disagreeing peers can be considered as an epistemically rational response to peer disagreement. Agents who follow Belief or Resilience Conciliation are thus individually rational, but, as a group, they frequently end up epistemically worse off. On the other hand, steadfast agents can be seen as individually irrational, but this in turn makes their collective inquiry better off.

4.2 ...But not under other zetetic norms

The negative effect of conciliation on group inquiry is not robust under different norms of inquiry. Figure 5 shows that under the Epsilon Greedy Norm, conciliatory and steadfast agents perform equally well. This is because the negative effects of the homogeneity of conciliatory agents are counteracted by their exploratory inquisitive behavior. In other words, although conciliatory agents all have the same or at least very similar beliefs, they still get enough information about both bandits from their exploratory behavior to prevent wrong convergence. In addition, it turns out that increased diversity, furnished by the Steadfast Norm, brings no additional benefits.

One could argue that these results do not show much. Epsilon greedy behavior is, in general, a very good strategy for agents solving a bandit problem (Kummerfeld and Zollman, 2016). This

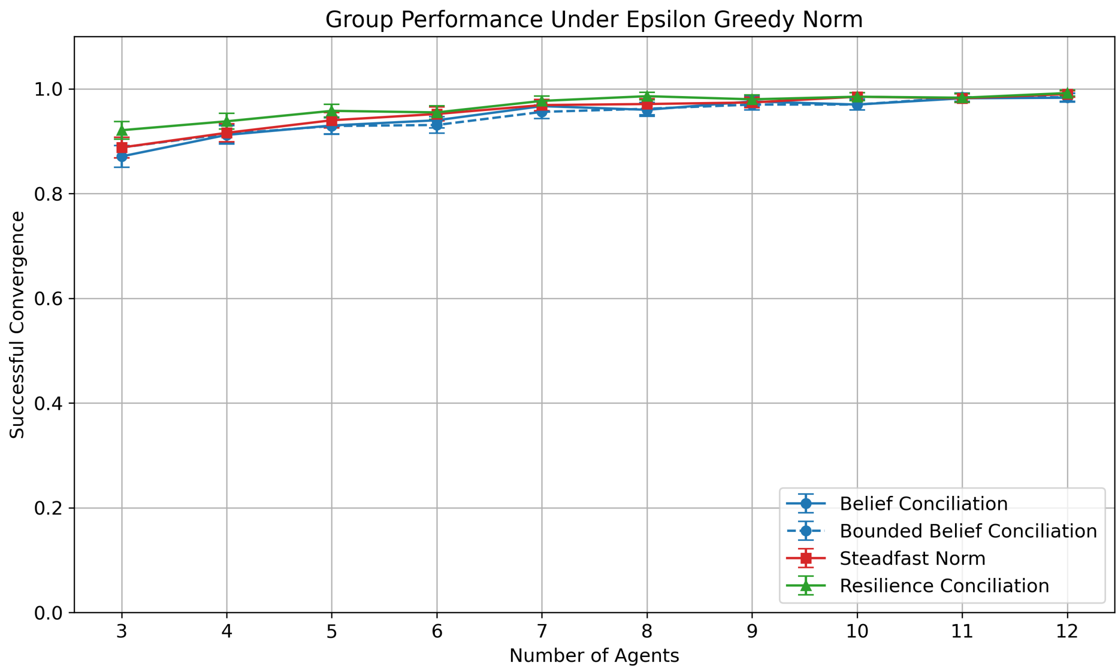

Figure 5: Steadfast and conciliatory agents under the Epsilon Greedy Norm for $\varepsilon = .005$.

means that, given enough time, agents who inquire in this way will be able to reach correct convergence irrespective of the epistemic norm at work. Therefore, the fact that conciliatory and steadfast agents perform equally well does not show anything interesting—*most* agents would reach correct convergence given enough time. To address this concern, we opted for a very small ε value. This way, agents' inquiry is meaningfully different from the one under the Credence-Based Norm but not so much as to make the problem too easy for them. The latter is shown by the fact that for most group sizes agents did not always correctly converge. Consequently, the results summarized in Figure 5 are not simply due to the problem being too easy for the agents. That is, a conciliatory population is able to catch up with a steadfast one not because of easy inquiry, but because of their inquisitive behavior which is now more exploratory than under the Credence-Based Norm.

Finally, we present the results for the Rational Inertia Norm in Figure 6. First, we see that both conciliatory and steadfast agents perform better under the Rational Inertia Norm than under the Credence-Based Norm. This was expected since, by preventing agents from switching bandits immediately, the Rational Inertia Norm acts as a source of diversity. It also aligns with the results in Frey and Šešelja (2020). Second, in contrast to Epsilon Greedy, there is a difference between the performance of both kinds of agents. At least for a relatively small i-value, steadfast agents perform the best and there is thus still a tension between the Conciliatory View and the collective inquiry.

Now, combining the insights of the previous discussion, we can predict that with higher i-values, the difference between the epistemic performance of steadfast and conciliatory agents decreases.

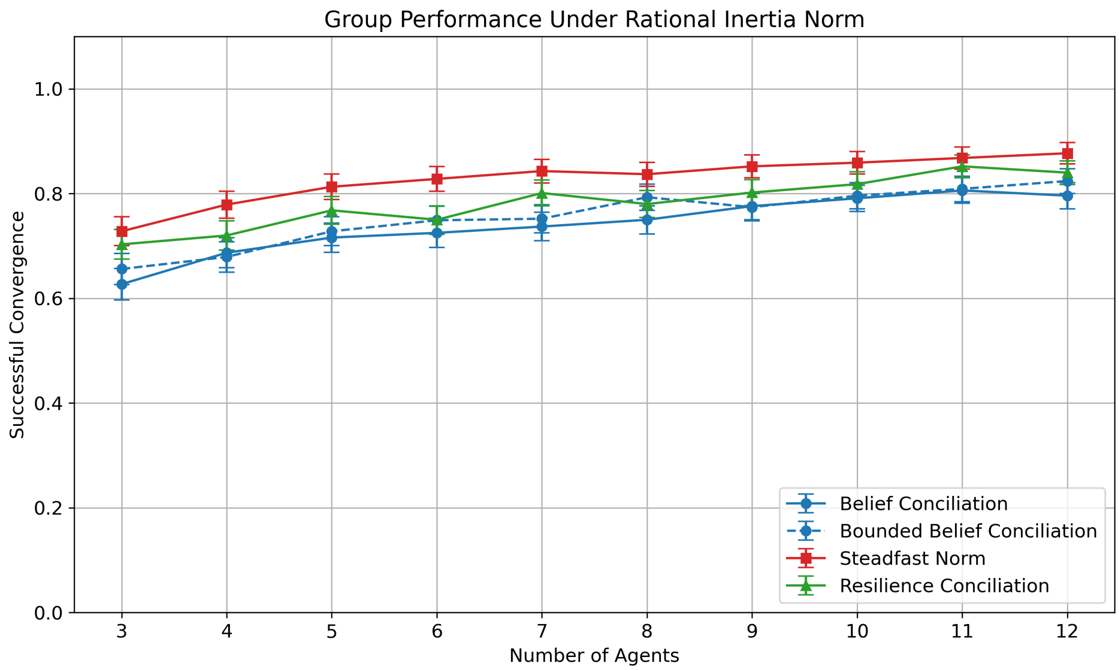

Figure 6: Steadfast and conciliatory agents under the Rational Inertia Norm for $i = 10$.

Forcing agents to stick to one bandit before switching is a way of ensuring diversity; making agents stick for longer thus ensures more diversity. Furthermore, if diversity is already ensured by agents' inquisitive behavior, their cognitive diversity—i.e., differences in their credences—matters less. This is indeed borne out in the data. At $i = 50$, for example, a 10-member group of conciliatory agents reaches the correct convergence 93.1 ($\pm$1.6) % of the time, while steadfast agents do in 95.7 ($\pm$1.3) %. The difference becomes even smaller if we further increase the i-value.

In sum, these results show that responding to peer disagreement in accordance with the Conciliatory View can hurt the success of collective inquiry. However, the results also suggest that this negative effect does not extend to all cases of disagreement. Agents who follow the Belief or Resilience Conciliation do not perform worse than steadfast agents across all zetetic norms—when agents follow the Epsilon Greedy Norm for inquiry, for example, conciliating does not have a negative impact on their group performance.

4.3 Discussion

Now, we can ask how this compares with the results presented by Douven (2010) and Douven and Kelp (2011). To recall, they discovered a trade-off between fast but less accurate steadfast agents and slower but more accurate conciliatory agents. Based on this result, they similarly argued that choosing the best response to peer disagreement is context dependent. Although we do not replicate this exact trade-off—for example, under the Credence-Based Norm, agents who follow

Belief Conciliation converge faster but are less accurate—our results seem to confirm that choosing an appropriate norm of peer disagreement is context-dependent. But there is a crucial difference between our results and Douven and Kelps' ones. In their case, deciding between the two norms of responding to peer disagreement depends on whether we prioritize speed or accuracy of inquiry. Our results, on the other hand, show that even if we only focus on one outcome—in our case, correct convergence, i.e., accuracy—the results depend on the underlying zetetic norm. In other words, our model shows that the decision about an appropriate response to disagreement is context-dependent in a deeper sense: it depends not only on our epistemic goals (do we want speed or accuracy?) but also on the underlying inquisitive behavior of scientists.

This is important because it highlights a broader methodological point. It shows that the epistemic effects of norms of responding to peer disagreement cannot be meaningfully studied independently from other kinds of norms, such as norms of inquiry. We therefore disagree with Douven (2010) that there are no "general and illuminating things to be said about what we ought or ought not to do in cases of peer disagreement". Our results do not discredit this possibility wholeheartedly; rather, they show that norms of responding to peer disagreement are closely intertwined with norms of inquiry and should thus be studied together with them. This does not exclude the possibility that, when considered together, one such combination of epistemic and zetetic norms would turn out to be the best for achieving our epistemic goals.

Consequently, our results show that idealizing away different zetetic norms is not inconsequential modeling decision. Idealization is of course crucial in modeling (and scientific representation in general). "Toy" (Reutlinger et al., 2018) or "minimalist" (Weisberg, 2007) models, such as Douven's, Zollman's, or indeed our model, allow us to "offer simple explanatory hypotheses about the causal dependencies" underlying the immensely complex target phenomena (Frey and Šešelja, 2020). However, as Frey and Šešelja point out, it is important to ask whether the results of these models remain robust under different idealizing assumptions about the underlying causal dependencies. Our results show that zetetic norms significantly impact the results and should thus be considered when deciding which features to present in the model.

This is important since many current models share the assumption that we can abstract away from different zetetic norms. This is true for Douven (2010) and Douven and Kelp (2011) but also for many bandit models which only focus on the "myopic" agents who follow what we call Credence-Based Norm (see, for example, Holman and Bruner, 2015; Weatherall and O'Connor, 2021; Wu, 2023; Zollman, 2007, 2010). Nevertheless, we do not wish to claim that our model is in any respect more descriptive than those mentioned here. It is still highly idealized and simple.

For example, our three zetetic norms are abstract and idealized versions of norms that guide actual scientific inquiry. In real life, zetetic norms may both look different and vary from one field to another due to various factors constraining scientists' approach to inquiry. Some of these factors are practical. Some research, for example, in experimental physics requires enormous investments in equipment, making an exploratory inquisitive strategy prohibitively expensive. In contrast, in some other fields, like cognitive psychology, the cost of conducting experiments may be comparably low, allowing researchers to pursue multiple lines of research at the same time. Other factors that constrain the inquiry are ethical. For instance, it is sometimes unethical for medical researchers to continue to conduct trials and treat patients with a drug they believe to be much less effective than an alternative drug. While these complexities show that scientists may not be able to simply replace one zetetic norm with another, our main interest in this paper was to investigate whether zetetic norms, in principle, make a difference for the impact of epistemic norms on the collective inquiry, rather than to determine the optimal combination of these two kinds of norms.

These issues also concern a broader methodological challenge: how much detail should we build into models of scientific inquiry? Overly simple models may miss crucial interactions between norms, while overly detailed models risk being unwieldy or context-specific. Our modeling approach aims to strike a middle ground, showing that even basic representations of zetetic norms can reveal dependencies between inquiry strategies and belief-updating norms. This balance, we suggest, is key to drawing meaningful lessons from formal models of scientific inquiry.[9]

5 Conclusion

In this paper, we argued that the existing criticism of the Conciliatory View on peer disagreement among scientists overlooks a crucial dimension of the debate: the role of zetetic norms—norms of inquiry—alongside epistemic (or doxastic) norms. The Conciliatory View states that in cases of peer disagreement, one should dissolve the disagreement by meeting the disagreeing peer halfway. An important argument against the view, brought up by Douven (2010) and Douven and Kelp (2011), states that conciliation can be detrimental to achieving our epistemic goals. They support their argument by developing an agent-based model of peer disagreement. In Section 2, we presented their model and highlighted some shortcomings. In Section 3, we then presented a novel model of peer disagreement based on the bandit modeling framework. Our model improves on previous work in two ways. First, it allows for implementing norms of peer disagreement that align better with the literature by explicitly separating between belief updating based on first-order evidence and sharing higher-order evidence as two sequential steps in the simulation. Second, it provides an example of modeling zetetic norms (norms of inquiry) in addition to epistemic norms. In Section 4, we discussed the results of the model and embedded them in the broader literature on disagreements.

9. We thank anonymous reviewers for encouraging us to discuss the issues related to idealizing assumptions about the zetetic norms and to simplicity and idealization in general.

We found that norms of responding to peer disagreement cannot be studied independently from norms that guide agents' inquiry. This connection between the two kinds of norms has been overlooked both in the literature on peer disagreement and in the study of scientific inquiry more broadly. Thus, our paper presents an exemplary study of how zetetic and epistemic norms interact in the context of group inquiry. Looking ahead, the model can be further extended by incorporating and comparing additional norms of responding to peer disagreement, as well as additional zetetic norms. For example, on the former, it would be interesting to test the norm according to which agents not only remain steadfast but also boost their confidence upon discovering *agreement* from others. Concerning the latter, it would be worthwhile exploring how zetetic norms proposed in the literature on the pursuit-worthiness of scientific theories perform in this model (e.g. Šešelja et al. (2012)). For instance, a zetetic norm following Laudan (1977) suggests that one should pursue a theory that exhibits a higher rate of problem solving than its rivals. This could be captured by agents pursuing a bandit that has exhibited the highest growth in its success rate.

Going beyond the bandit models, the approach taken here can be fruitfully employed to study the interaction between epistemic and zetetic norms in other frameworks. For instance, a similar study could be conducted using the argumentative landscape model presented by Borg et al. (2019, 2018). In that model, a scientific theory is represented as a set of arguments, modeled in an abstract way. While an individual theory is conflict-free, meaning that no argument in a theory attacks another in the same theory, arguments from different theories attack each other. By exploring the argumentative landscape and discovering new arguments, agents gradually learn which of the theories is more "defensible", and therefore superior to its rivals. Although the authors do not describe them as such, the model already contains both epistemic and zetetic norms. On the one hand, agents can employ different norms to evaluate which theory they should accept. On the other hand, agents can employ different strategies for deciding which theory to pursue, that is, when to switch from one theory to another and explore that one instead. We take the first to be a clear example of epistemic and the latter an example of zetetic norms.

Another area of study where the interaction between the two kinds of norms could be explored is the problem of the division of cognitive labor. For example, the epistemic landscape model has been frequently used to study different norms to ensure an optimal division of labor in exploring an unknown epistemic landscape (Alexander et al., 2015; Pöyhönen, 2017; Thoma, 2015; Weisberg and Muldoon, 2009). This model can also be used to study both zetetic norms—e.g., Weisberg and Muldoon have compared agents follow others and agents who work independently in deciding which part of the landscape to explore—and epistemic norms—e.g., different ways in which agents evaluate the significance of the patch on the landscape they are currently exploring. Similarly, the dynamic could be explored in the NK-landscape models. In the model by Wu (2024), agents employ both social learning and individual inquiry to explore the highly multidimensional landscape. While her paper focuses on comparing the two different social learning strategies, in a footnote, Wu already suggests implementing alternative search rules, which could be understood as alternative zetetic norms. Thus, the approach of studying the combinations of these norms presented in this paper could provide a new perspective on this and similar debates.

References

Alexander, Jason Mc Kenzie, Johannes Himmelreich, and Christopher Thompson (2015). "Epistemic landscapes, optimal search, and the division of cognitive labor". In: *Philosophy of Science* 82 (3), pp. 424–453. DOI: 10.1086/681766.

Bala, Venkatesh and Sanjeev Goyal (1998). "Learning from Neighbours". In: *Review of Economic Studies* 65.3, pp. 595–621. DOI: 10.1111/1467-937X.00059.

Bogardus, Tomas (2009). "A Vindication of the Equal-Weight View". In: *Episteme* 6.3, pp. 324–335. DOI: 10.3366/E1742360009000744.

Borg, AnneMarie, Daniel Frey, Dunja Šešelja, and Christian Straßer (2019). "Theory-choice, transient diversity and the efficiency of scientific inquiry". In: *European Journal for Philosophy of Science* 9.2, p. 26. DOI: 10.1007/s13194-019-0249-5.

Borg, AnneMarie, Daniel Frey, Dunja Šešelja, and Christian Straßer (2018). "Epistemic Effects of Scientific Interaction: Approaching the Question with an Argumentative Agent-Based Model". In: *Historical Social Research* 43.1, pp. 285–309.

Christensen, David (2007). "Epistemology of Disagreement: The Good News". In: *The Philosophical Review* 116.2, pp. 187–217. DOI: 10.1215/00318108-2006-035.

Cruz, Helen De and Johan De Smedt (2013). "The value of epistemic disagreement in scientific practice. The case of Homo floresiensis". In: *Studies in History and Philosophy of Science Part A* 44.2, pp. 169–177. DOI: 10.1016/j.shpsa.2013.02.002.

Dellsén, Finnur and Maria Baghramian (2021). "Disagreement in science: introduction to the special issue". In: *Synthese* 198.S25, pp. 6011–6021. DOI: 10.1007/s11229-020-02767-0.

Douven, Igor (2010). "Simulating peer disagreements". In: *Studies in History and Philosophy of Science Part A* 41.2, pp. 148–157. DOI: 10.1016/j.shpsa.2010.03.010.

— (2019). "Optimizing group learning: An evolutionary computing approach". In: *Artificial Intelligence* 275, pp. 235–251. DOI: 10.1016/j.artint.2019.06.002.

— (2025). "Social Learning in Neural Agent-Based Models". In: *Philosophy of Science* 92 (1), pp. 141–161. DOI: 10.1017/psa.2024.33.

Douven, Igor and Rainer Hegselmann (2021). "Mis- and disinformation in a bounded confidence model". In: *Artificial Intelligence* 291, p. 103415. DOI: 10.1016/j.artint.2020.103415.

Douven, Igor and Rainer Hegselmann (2022). “Network effects in a bounded confidence model”. In: *Studies in History and Philosophy of Science* 94, pp. 56–71. DOI: 10.1016/j.shpsa.2022.05.002.

Douven, Igor and Christoph Kelp (2011). “Truth Approximation, Social Epistemology, and Opinion Dynamics”. In: *Erkenntnis* 75.2, pp. 271–283. DOI: 10.1007/s10670-011-9295-x.

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

Elgin, Catherine Z. (2010). “Persistent Disagreement”. In: *Disagreement*. Oxford University Press, pp. 53–68. DOI: 10.1093/acprof:oso/9780199226078.003.0004.

Falbo, Arianna (2023). “Should epistemology take the zetetic turn?” In: *Philosophical Studies* 180.10-11, pp. 2977–3002.

Fazelpour, Sina and Daniel Steel (2022). “Diversity, Trust, and Conformity: A Simulation Study”. In: *Philosophy of Science* 89.2, pp. 209–231. DOI: 10.1017/psa.2021.25.

Feldman, Richard (2005). “Respecting the Evidence”. In: *Philosophical Perspectives* 19.1, pp. 95–119. DOI: 10.1111/j.1520-8583.2005.00055.x.

— (2006). “Epistemological Puzzles about Disagreement”. In: *Epistemology Futures*. Oxford University PressOxford, pp. 199–215. DOI: 10.1093/oso/9780199273317.003.0013.

Fleisher, Will (2018). “Rational endorsement”. In: *Philosophical Studies* 175.10, pp. 2649–2675. DOI: 10.1007/s11098-017-0976-4.

Frey, Daniel and Dunja Šešelja (2020). “Robustness and Idealizations in Agent-Based Models of Scientific Interaction”. In: *The British Journal for the Philosophy of Science* 71.4, pp. 1411–1437. DOI: 10.1093/bjps/axy039.

Friedman, Jane (2020). “The Epistemic and the Zetetic”. In: *The Philosophical Review* 129.4, pp. 501–536. DOI: 10.1215/00318108-8540918.

Gabriel, Nathan and Cailin O’Connor (2024). “Can Confirmation Bias Improve Group Learning?” In: *Philosophy of Science* 91 (2), pp. 329–350. DOI: 10.1017/psa.2023.176.

Hegselmann, Rainer and Ulrich Krause (2002). “Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation”. In: *Journal of Artificial Societies and Social Simulation* 5.3.

— (2006). “Truth and Cognitive Division of Labour: First Steps Towards a Computer Aided Social Epistemology”. In: *Journal of Artificial Societies and Social Simulation* 9.3, p. 10.

Holman, Bennett and Justin Bruner (2017). “Experimentation by Industrial Selection”. In: *Philosophy of Science* 84.5, pp. 1008–1019. DOI: 10.1086/694037.

Holman, Bennett and Justin P. Bruner (2015). “The Problem of Intransigently Biased Agents”. In: *Philosophy of Science* 82.5, pp. 956–968. DOI: 10.1086/683344.

Jehle, David and Branden Fitelson (2009). “What is the “Equal Weight View”?” In: *Episteme* 6.3, pp. 280–293. DOI: 10.3366/E1742360009000719.

Kelly, Thomas (2010). “Peer Disagreement and Higher-Order Evidence”. In: *Disagreement*. Oxford University Press, pp. 111–174. DOI: 10.1093/acprof:oso/9780199226078.003.0007.

Kitcher, Philip (1990). “The Division of Cognitive Labor”. In: *The Journal of Philosophy* 87.1, p. 5. DOI: 10.2307/2026796.

Kornblith, Hilary (2010). “Belief in the Face of Controversy”. In: *Disagreement*. Oxford University Press, pp. 29–52. DOI: 10.1093/acprof:oso/9780199226078.003.0003.

Kummerfeld, Erich and Kevin J. S. Zollman (2016). “Conservatism and the Scientific State of Nature”. In: *The British Journal for the Philosophy of Science* 67.4, pp. 1057–1076. DOI: 10.1093/bjps/axv013.

Langhe, Rogier De (2013). “Peer disagreement under multiple epistemic systems”. In: *Synthese* 190.13, pp. 2547–2556. DOI: 10.1007/s11229-012-0149-0.

Laudan, Larry (1977). *Progress and its Problems: Towards a Theory of Scientific Growth*. London: Routledge & Kegan Paul Ltd.

Lawler, Lisa (2024). “On the nature of scientific disagreement”. In: *Routledge Handbook of Philosophy of Disagreement*. Ed. by M. Baghramian, J.A Carter, and R Rowland. Routledge. Forthcoming.

Lyon, Aidan (2014). “Why are Normal Distributions Normal?” In: *The British Journal for the Philosophy of Science* 65.3, pp. 621–649. DOI: 10.1093/bjps/axs046.

Matheson, Jonathan (2009). “Conciliatory Views of Disagreement and Higher-Order Evidence”. In: *Episteme* 6 (3), pp. 269–279. DOI: 10.3366/E1742360009000707.

Mayo-Wilson, Conor, Kevin J. S. Zollman, and David Danks (2011). “The Independence Thesis: When Individual and Social Epistemology Diverge”. In: *Philosophy of Science* 78.4, pp. 653–677. DOI: 10.1086/661777.

Michelini, Matteo, Javier Osorio, Wybo Houkes, Dunja Šešelja, and Christian Straßer (2023). “Scientific Disagreements and the Diagnosticity of Evidence: How Too Much Data May Lead to Polarization”. In: *Journal of Artificial Societies and Social Simulation* 26 (4). DOI: 10.18564/jasss.5113.

O’Connor, Cailin and James Owen Weatherall (2018). “Scientific polarization”. In: *European Journal for Philosophy of Science* 8.3, pp. 855–875. DOI: 10.1007/s13194-018-0213-9.

Oliveira, Jaqueline Faria de, Humberto Torres Marques-Neto, and Márton Karsai (2022). “Measuring the effects of repeated and diversified influence mechanism for information adoption on Twitter”. In: *Social Network Analysis and Mining* 12.1, p. 16.

Piedrahita, Pablo, Javier Borge-Holthoefer, Yamir Moreno, and Sandra González-Bailón (2018). “The contagion effects of repeated activation in social networks”. In: *Social networks* 54, pp. 326–335.

Pöyhönen, Samuli (2017). “Value of cognitive diversity in science”. In: *Synthese* 194 (11), pp. 4519–4540. DOI: 10.1007/S11229-016-1147-4/FIGURES/5.

Reutlinger, Alexander, Dominik Hangleiter, and Stephan Hartmann (2018). “Understanding (with) Toy Models”. In: *The British Journal for the Philosophy of Science* 69.4, pp. 1069–1099. DOI: 10.1093/bjps/axx005.

Rosenstock, Sarita, Justin Bruner, and Cailin O’Connor (2017). “In Epistemic Networks, Is Less Really More?” In: *Philosophy of Science* 84.2, pp. 234–252. DOI: 10.1086/690717.

Santana, Carlos (2021). “Let’s not agree to disagree: the role of strategic disagreement in science”. In: *Synthese* 198.S25, pp. 6159–6177. DOI: 10.1007/s11229-019-02202-z.

Šešelja, Dunja (2021). “Some lessons from simulations of scientific disagreements”. In: *Synthese* 198.S25, pp. 6143–6158. DOI: 10.1007/s11229-019-02182-0.

— (2022). “Agent-based models of scientific interaction”. In: *Philosophy Compass* 17.7. DOI: 10.1111/phc3.12855.

Šešelja, Dunja, Laszlo Kosolosky, and Christian Straßer (2012). “Rationality of Scientific Reasoning in the Context of Pursuit: drawing appropriate distinctions”. In: *Philosophica* 86, pp. 51–82.

Skipper, Mattias and Asbjørn Steglich-Petersen (2020). “When Conciliation Frustrates the Epistemic Priorities of Groups”. In: *The Epistemology of Group Disagreement*. Ed. by Fernando Broncano-Berrocal and Adam Carter. Routledge.

Skyrms, Brian (1977). “Resiliency, propensities, and causal necessity”. In: *The Journal of Philosophy* 74.11, pp. 704–713.

Sliwa, Paulina and Sophie Horowitz (2015). “Respecting all the evidence”. In: *Philosophical Studies* 172.11, pp. 2835–2858. DOI: 10.1007/s11098-015-0446-9.

Smaldino, Paul E., Cody Moser, Alejandro Pérez Velilla, and Mikkel Werling (2024). “Maintaining Transient Diversity Is a General Principle for Improving Collective Problem Solving”. In: *Perspectives on Psychological Science* 19 (2), pp. 454–464. DOI: 10.1177/17456916231180100.

Steglich-Petersen, Asbjørn (2019). “Higher-Order Defeat and Doxastic Resilience”. In: *Higher-Order Evidence*. Oxford University PressOxford, pp. 209–225. DOI: 10.1093/oso/9780198829775.003.0010.

Thoma, Johanna (2015). “The epistemic division of labor revisited”. In: *Philosophy of Science* 82 (3), pp. 454–472. DOI: 10.1086/681768.

Thorstad, David (2021). “Inquiry and the epistemic”. In: *Philosophical Studies* 178.9, pp. 2913–2928.

Wagner, Elliott and Jonathan Herington (2021). “Agent-Based Models of Dual-Use Research Restrictions”. In: *The British Journal for the Philosophy of Science* 72 (2), pp. 377–399. DOI: 10.1093/bjps/axz017.

Weatherall, James Owen and Cailin O'Connor (2021). "Conformity in scientific networks". In: *Synthese* 198.8, pp. 7257–7278. DOI: 10.1007/s11229-019-02520-2.

Weatherall, James Owen, Cailin O'Connor, and Justin P. Bruner (2020). "How to Beat Science and Influence People: Policymakers and Propaganda in Epistemic Networks". In: *The British Journal for the Philosophy of Science* 71.4, pp. 1157–1186. DOI: 10.1093/bjps/axy062.

Weatherall, James Owen and Cailin O'Connor (2021). "Endogenous epistemic factionalization". In: *Synthese* 198 (S25), pp. 6179–6200. DOI: 10.1007/s11229-020-02675-3.

Weisberg, Michael (2007). "Three Kinds of Idealization". In: *Journal of Philosophy* 104.12, pp. 639–659. DOI: 10.5840/jphil20071041240.

Weisberg, Michael and Ryan Muldoon (2009). "Epistemic Landscapes and the Division of Cognitive Labor". In: *Philosophy of Science* 76.2, pp. 225–252. DOI: 10.1086/644786.

Wu, Jingyi (2023). "Epistemic advantage on the margin: A network standpoint Epistemology". In: *Philosophy and Phenomenological Research* 106 (3), pp. 755–777. DOI: 10.1111/phpr.12895.

— (2024). "Better than Best: Epistemic Landscapes and Diversity of Practice in Science". In: *Philosophy of Science* 91 (5), pp. 1189–1198. DOI: 10.1017/psa.2023.129.

Wu, Jingyi and Cailin O'Connor (2023). "How should we promote transient diversity in science?" In: *Synthese* 201.2, p. 37. DOI: 10.1007/s11229-023-04037-1.

Zollman, Kevin J. S. (2007). "The Communication Structure of Epistemic Communities". In: *Philosophy of Science* 74.5, pp. 574–587. DOI: 10.1086/525605.

— (2010). "The Epistemic Benefit of Transient Diversity". In: *Erkenntnis* 72.1, pp. 17–35. DOI: 10.1007/s10670-009-9194-6.

Appendix

This appendix expands on our critique of the model presented in Douven (2010). We argue that the model has two limitations: first, agents do not update their beliefs in a rational way, and second, the conciliatory norm is not represented in an adequate way. We discuss each of these points in turn.

Let's first recall how agents take the truth-signal into account. In each turn, they simply draw a new value from a uniform distribution between $\tau - \zeta$ and $\tau + \zeta$. Then, they give it a certain weight and sum it with their weighted prior belief (or, in the case of conciliatory agents, a weighted aggregate of their prior belief and the prior beliefs of their peers). Consequently, with each update, agents only consider the newest draw from the truth-signal. This doesn't mean that they simply discard their old evidence—they keep it as part of what formed their prior belief. However, it entails that for the agents, every new piece of evidence has the same impact on one's belief about τ no matter how much evidence they have previously gathered.[10]

We find this assumption inadequate for modeling rational reasoners. For one, it does not apply to scientists; they do not simply disregard the amount of existing evidence when considering the results of a new study. For example, for a scientist, acting as Douven's agent, a new study showing that $p = 0.9$ would have the same impact on her belief about p as a long series of consistent studies showing that $p = 0.9$. To relax this assumption, we modified a model so that agents remember previous draws from the truth signal and update using an average of the new draw and a certain number of previous ones. The number of the previous draws—the size of agents' memories—is a parameter of the model and can be modified.

Furthermore, we believe that modeling the noise of the truth signal using a uniform distribution is unrealistic. Rather, we model noise as normally distributed. Although, as Lyon (2014) points out, "normal distributions are not as normal as we once thought they were," the fact that it is found in nature makes it a better contender for modeling this kind of idealized noisy truth signal.

Figure 7 compares conciliatory and steadfast agents with memory who draw from a normally distributed noisy truth signal. We can still see the trade-off between speed and accuracy.[11] However, in comparison to the original results (Figure 1), the difference in accuracy is less pronounced. Specifically, we see that the erratic behavior of steadfast agents in the original model was an artifact

10. Take as an example an agent i who received the sequence $.5, .5, .5, .9$ in rounds 1–4 from the noisy truth signal. We suppose that i has no other agents in their confidence interval and that i starts with the belief .5. In Douven's model the sequence of beliefs is $x_i(1) = \frac{.5+.5}{2} = .5 = x_i(2) = x_i(3)$ and $x_i(4) = \frac{.5+.9}{2} = .7$. In contrast, a Bayesian agent will average all previous data to obtain $x_i(4) = \frac{.5+.5+.5+.9}{4} = .6$.

11. Code for the modified model is publicly available at: [REDACTED].

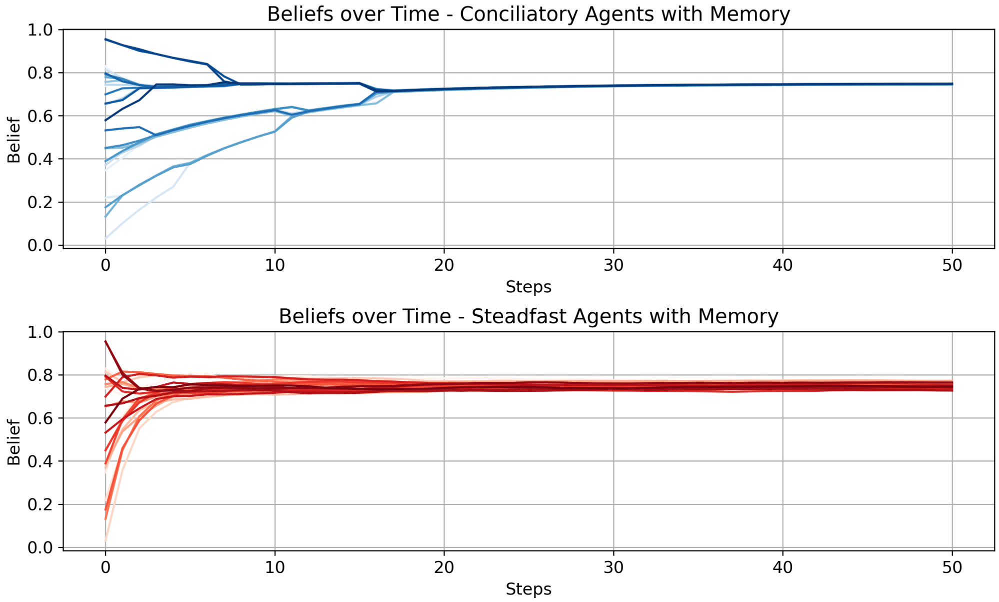

Figure 7: Comparison of conciliatory and steadfast agents with memory and normally distributed noisy data.

of implementing reasoning with noisy truth signal in an overly simplified manner.

That being said, Douven's main result concerning the trade-off between speed and accuracy remains robust (even if less pronounced) under these changes. However, the model has another issue. This has to do with the way different norms for responding to peer disagreement are implemented in the model. As mentioned above, Douven represents conciliatory and steadfast agents by manipulating α- and ε-values in Equation 1. Conciliatory agents split the difference with their close peers ($\varepsilon = 0.1$) and give more weight to this aggregate belief ($\alpha = 0.9$). Steadfast agents, on the other hand, ignore other agents with different beliefs ($\varepsilon = 0$) and give equal weight to their prior belief and the information from the truth signal ($\alpha = 0.5$).

Using different ε-values is an intuitive way of modeling various ways of responding to peer disagreement: conciliatory agents take their peers into account, while steadfast agents do not. What about using different α-values? At face value, this also seems reasonable. After all, the Conciliatory View does require that reasoners prioritize higher-order evidence, coming from the testimony of others to first-order evidence they gathered in their own research. In addition, some philosophers, for example Kelly (2010) and Sliwa and Horowitz (2015), argue that the Conciliatory View requires reasoners to ignore first-order evidence. Since, as Jehle and Fitelson (2009) emphasize, it is not clear what updating rule follows from the Conciliatory View, this issue is open to interpretation, with Douven taking one possible take on it.

While we agree that Douven's model represents conciliatory agents who put much more weight on higher-order evidence than on evidence each agent gathered on their own, this doesn't seem to adequately capture conciliation in case of scientific disagreements. Scientists conciliating in this way would barely consider evidence acquired through their own research. Instead, they would form their views mainly on the basis of their peers' beliefs, which is a rather extreme kind of conciliation. This suggests taking lower values for α as a more suitable assumption to model scientific disagreements.

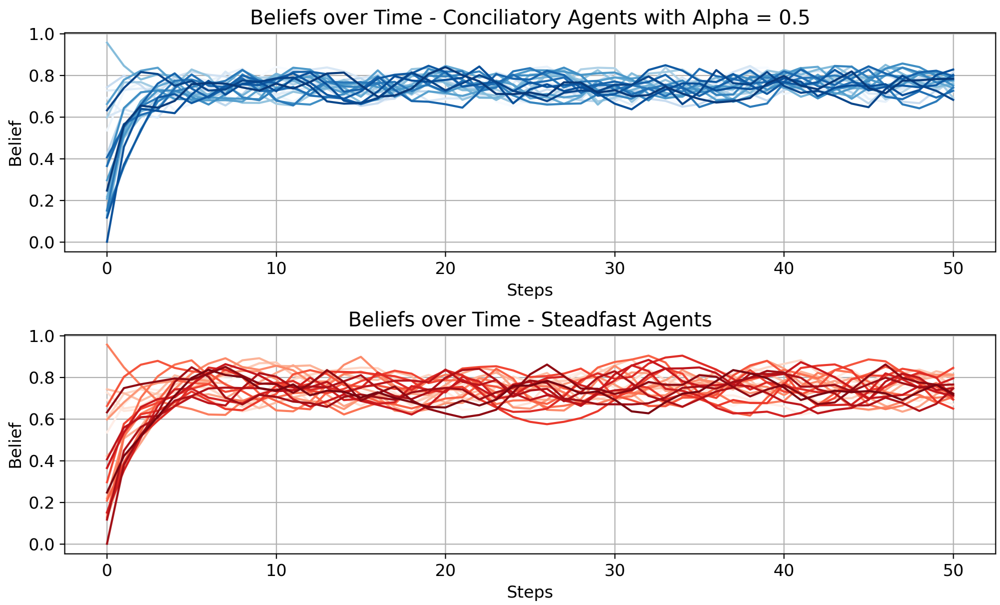

Figure 8: Comparison between Douven's conciliatory and steadfast agents with α-value fixed at 0.5.

Figure 8 compares the performance of conciliating ($\varepsilon = 0.1$) and steadfast ($\varepsilon = 0$) agents under the same α-value. The figure shows that if we hold α fixed between both kinds of agents, conciliating doesn't have downsides: conciliatory agents are at least as accurate as steadfast ones while speed of accuracy depends mainly on α. The trade-off between speed and accuracy, reported by Douven (2010) and Douven and Kelp (2011) thus seems to be simply a consequence of using different α-values for different kinds of populations.