Evidential Dependence and Polarization in Social Networks

Leon Assaad (L.Assaad@campus.lmu.de)
Munich Center for Mathematical Philosophy, LMU Munich
Ludwigstraße 31, 80539 Munich, Germany
David Young (d.young@bbk.ac.uk), Ulrike Hahn (u.hahn@bbk.ac.uk)
Centre for Cognition, Computation, and Modelling, Birkbeck, Univ. of London
Malet Street, London, WC1E 7HX, UK

Abstract

It is an inherent part of communication across social networks that the same evidence can reach us by different paths. This creates the possibility of inadvertent double counting of evidence. Recent work has shown that some degree of double counting seems to be an inevitable part of group deliberation. This makes it imperative to understand what impact evidence-reduplication has on beliefs or opinions across social networks. In this paper, we use agent-based modelling to show how and why double counting gives rise to polarization.

Keywords: Agent-Based Model; Polarization; Bayesian Inference; Deliberation; Social Networks

Introduction

When faced with a situation in which multiple sources report the sighting of a bear in the neighbourhood, it seems intuitively obvious that our degree of belief or confidence that there is, indeed, a bear should be different depending on whether these sources have each independently spotted a bear, or whether their reports stem ultimately from a further single, underlying source.

Experimental research has become increasingly interested in the extent to which lay reasoners are sensitive to this distinction (Madsen et al., 2020; Pilditch et al., 2020; Whalen et al., 2018; Yousif et al., 2019). The empirical conclusion from that body of work, to date, suggests that lay reasoners show some sensitivity to whether or not evidence sources are dependent or independent, but fail to reckon fully with the intricacies that can emerge here.

What has remained obscured in that literature is the extent to which the problem of evidential dependency is, arguably, both *ineliminable* and *unsolvable*. Evidential dependence is ubiquitous because we are typically part of wider social networks, even where we are ostensibly communicating only in dyads (Hahn, 2023). Kathy and Priya may be speaking only to each other in a given moment, but each also has other people they speak to, potentially about the same topic, and some of those people, in turn, might be direct contacts.

In short, information exchange across a social network happens both across direct and indirect paths that link agents within a wider network. Information flows across all of these paths to create informational dependence. As the simple example of the bear illustrates, a failure to reckon with the fact that essentially the same information has reached us via multiple routes will lead to assuming there is more evidence than there actually is. In effect, unnoticed evidential dependence creates over-weighting or 'double counting'.

The first problem in trying to deal appropriately with the problem of evidential dependence is that *in practice* we typically do not know (or even have a way of knowing) the structure of the wider network across which information flows (Merdes et al., 2021). More surprising is the fact that—even if we did—there is no normative tool that would allow us to calculate the 'correct' discounted evidential value. It is clear that we should somehow make adjustments to the evidential value in order to account for the dependence, but by how much, exactly, should we adjust that value?

It turns out that not even an ideal Bayesian agent who knows the network structure can calculate the impact of dependence in general. Bayesian Belief Networks were, in fact, designed precisely to allow normative Bayesian computations for dependent evidence (Pearl, 1988), and are the best tool that we have. However, recent analysis shows that Bayesian Belief Networks cannot fully capture evidential dependence on social networks (Pilditch et al., 2025). For one, there is a mismatch in the semantics of Bayesian Networks in that they capture the perspective of an external observer, not a node that is itself within the network. Given that our best tool does not provide a solution here, there is consequently a deep sense in which the problem of reckoning with dependence to avoid double-counting is (currently) unsolvable.

All of this means that there will inevitably be 'double counting errors' brought about by communication across social networks (and with it dependence) which itself is ubiquitous. This makes it imperative to understand the consequences of failure to reckon with evidential dependence.

In this paper, we explore those consequences by means of an agent-based simulation. As we will see, evidential dependence is a powerful source of polarization. This has only been hinted at in past work (Hahn, 2023). There are no simulations that provide qualitative and quantitive insight into this phenomenon. This matters because consensus is seen as desirable across many deliberation contexts (Landemore & Page, 2015). And there is not only widespread concern about polarization as a threat to democratic societies (Sunstein, 2018), there is also a widespread suspicion that changes to our information environment, in particular the rise of online social media and its displacement of traditional media outlets, have fuelled increases in polarization.

However, it remains unclear how that might be. In fact,

In G. Boleda, I. Dautriche, M. Maldonado, C. Saldana and T. Torrent (Eds.), *Proceedings of the 48th Annual Conference of the Cognitive Science Society*

most models of opinion or belief dynamics across social networks inevitably converge, unless specific disruptive mechanisms are added to their trajectory (Bramson et al., 2017; Mäs & Flache, 2013). Many putative causal factors that could cause such disruption have been suggested in the literature: divergent perceptions of source reliability (Henderson & Gebharter, 2021; Olsson, 2020), affiliative sorting (Mäs & Flache, 2013), the selective sharing of information (Assaad & Hahn, 2024; Kopecky, 2024), algorithmically induced filter-bubbles (Pariser, 2011), memory limitations (Singer et al., 2019), divergent perceptions of how the evidence relates to the causal structure of the world (Jern et al., 2014), and irrationalities such as motivated reasoning and cognitive biases (Anderson, 2021; Kahan, 2013; Taber et al., 2009).

The potential role of *informational dependence* (Hahn, 2023) as a factor has entered the discussion only recently. There have, however, been no attempts to date to explore this possibility in simulations. The present paper addresses that gap with agent-based simulations that explicitly manipulate informational dependence. These make more precise how and why informational dependence generates polarization.

Modelling of Polarization

Imagine a group of agents collecting and exchanging information to deliberate whether a certain hypothesis H is true (H) or false ($\neg H$). Each forms a degree of belief $P(H) \in [0, 1]$ about the hypothesis based on the received information. This scenario, much of the modelling literature assumes, captures the essential form of real group deliberation and we will make the same assumption (e.g., Baccini et al., 2023).

Polarization is then understood as belief dispersion induced by deliberation (cf. Bramson et al., 2017, p. 120). If the group arrives (more or less) at one degree of belief, we speak of convergence on consensus. All other belief distributions show belief dispersion of varying degrees. Perfect bi-polarization (i.e., one half of the group believing 1, the other 0) is the most extreme case of dispersion, though such perfect splits are unlikely to obtain in models, and even less so in the real world. How do computational models reconstruct polarization among rational agents? As already noted, the literature contains a plethora of models (e.g., Axelrod, 1997; Hegselmann, Krause, et al., 2002). The specific modelling route we pursue here involves Bayesian agents. Bayesian reasoning is an established standard for rational reasoning (Hartmann, 2021), so it constitutes a 'best case' starting point. In past research on polarization, Bayesian models have featured in two different strands: bandit models and source reliability simulation models.

Bandit models (Bala & Goyal, 1998) use one or multi-armed bandits where each 'arm' represents a theory. Agents test these theories' success rates, aiming to identify the arm with the highest expected payoff. They compute beliefs about the best theory through Bayesian conditionalization based on their own outcomes and those of their neighbours. The standard two-armed bandit models used in the study of scientific communication (Zollman, 2007, 2010) do not lead to polarization. Polarization can, however, be obtained by including a conformity bias parameter (Weatherall & O'Connor, 2021) or by allowing evidence from disagreeing neighbours to be seen as increasingly uncertain (O'Connor & Weatherall, 2018).

While agents in bandit models exchange evidence directly, Bayesian source reliability models involve agents exchanging *testimony*: agents assert to other agents the truth or falsity of the hypothesis (once their degree of belief exceeds some threshold) and that assertion itself is (socially obtained) evidence alongside private evidence ('from the world') that agents collect. Crucially, the evidential value of such testimonial evidence rests on perceived source reliability: trusted sources have greater impact on recipient's beliefs. Diverging perceptions of trust may give rise to bi-polarization (Olsson, 2020) whereby (typically) a majority converges on the truth with a stranded minority firmly believing the opposite. Many subsequent studies have explored this approach (Fränken & Pilditch, 2021; Hahn et al., 2024; Merdes et al., 2021; Pallavicini et al., 2021) and the underlying drivers of source reliability (Collins et al., 2018).

The simulations in this paper share the Bayesian approach, but like bandit models, involve agents (faithfully) sharing evidence, not testimony.

Simulations

For our simulations, we use the NormAN framework (short for Normative Argument Exchange across Networks, developed in Assaad et al., 2023—a freely available[1] agent-based simulation environment). NormAN follows on from earlier models of argument exchange such as Mäs and Flache, 2013. In it, rational Bayesian agents exchange evidence directly, interpret evidence uniformly (without trust updating), recognize indexed pieces evidence—so they never double-count any—and have perfect memory.

Deliberation is based on a set of evidence pieces: each piece is initialized as 'true' or 'false' and points towards or against the truth of the discussed hypothesis H (e.g., $\{E_1, \neg E_2, \ldots, \neg E_n\}$). Agents become aware of the true values of these pieces of evidence either through individual inquiry, or through communication. Communication involves choosing and directly passing on specific pieces of evidence, i.e., argument exchange. Most of the work using NormAN to date has explored the consequences of different communication rules (Schöppl, 2025; Schöppl & Hahn, 2024). By contrast, our simulations explore differences in the nature of the evidence as it presents itself to agents.

In order to study the impact of 'double counting' errors brought about by non-independence it is important that we start the simulations with independent pieces of evidence. While the NormAN framework is designed to help study argument exchange about problems involving complex relationships between individual pieces of evidence (not only arguments for and against, but arguments that relate to each other,

[1] https://github.com/NormAN-framework/base-model

such as under-cutters or defeaters, Assaad et al., 2025), this particular feature is not immediately relevant to the current investigation. Consequently, we started with model variant developed by Assaad and Hahn, 2024 that involves random draws of independent pieces of evidence.[2] To that code, we added the ability to vary the percentage of the evidence that was indexed, and hence recognisable to agents as (potential) reduplication.[3]

Whereas the earlier code indexed all pieces of evidence, our code allows some pieces of evidence to remain un-indexed. This means multiple copies of that same piece of evidence received via communication are not recognized as 'the same piece of evidence'. Consequently, rather than updating their beliefs only once, agents update in response to each copy. In our simulations, we varied the indexable percentage between 0% (no recognition) and 100% (complete indexing) where the percentages correspond to the proportion of the unique arguments that have received an index (rounded to the nearest integer, with 0.5 being rounded up).

Agents receive evidence in two ways: individual inquiry or communication with link-neighbours. Since our simulations seek to examine the impact of evidential dependence via communication, however, we disabled individual inquiry in our model runs. Instead, we simply initialized the networks by providing each agent with their own single piece of unique evidence. Following Assaad and Hahn, 2024, we used symmetric evidence sets: each pro-H item is mirrored by an equally strong con-H item, such that the product of their likelihood ratios is 1. There are equally many strong arguments for both sides, randomly distributed among agents. Lastly, H was initialized as 'true' (H) with a base rate of 50%.

Over the model run, agents communicate with their neighbours by randomly selecting a piece of evidence from the evidence they have received so far. While random sharing is inefficient in as much as the same evidence may be shared repeatedly, it is an unbiased sharing rule that—under conditions of full indexing—means that all agents in a fully connected network will receive all pieces of evidence, and therefore converge on the same posterior degree of belief if the simulation is run long enough. It thus provides a useful baseline. Functionally, one might take it to map onto communicative contexts such as posting on social media where one does not know enough about one's audience to maintain a stable model of what they know, so the communication mode is essentially 'broadcast'.

Agents collect pieces of evidence by adding the respective truth values to their memory store and aggregate them via Bayes' rule (e.g., $P(H|E_1, \neg E_4)$). Each agent starts with the same prior degree of belief $P(H) = 0.5$ (a 'neutral' belief, Baccini et al., 2023). Finally, agents communicate with one another across symmetrical links determined by the social network. For our simulations we used a generic small world network (Watts & Strogatz, 1998), with a re-wiring parameter of 0.1 and degree set by $k = 2$ using the options the NormAN framework provides.

Our dependent variable of interest was polarization. We considered two measures, the variance and the absolute average deviation (AAD) from the mean population belief—as per Bramson et al., 2017. The results below show the mean final AAD values of simulation runs (alternative measures are available on OSF). AAD values range from 0—all agents share the same degree of belief—to 0.5, where agents are perfectly bi-polarized: half believe $P(H) = 0$, the other $P(H) = 1$. As independent variables, we manipulated network size (the number of agents) and the indexable percentage, simulating all factorial combinations of five network sizes (10, 20, 50, 100, 200) and six levels of indexable % (0, 30, 50, 70, 90, 100). The model was run for 200 time steps. Each combination was simulated 200 times, for a total of 9000 runs.

Results

Figure 1 shows the results of our basic simulations. The five panels of Fig. 1 represent different network sizes (N= 10, 20, 50, 100, and 200). Along the x-axis in each panel is the indexable percentage among the available pieces of evidence. As can be seen clearly, polarization increases with network size: for small N, networks mostly converge, for the largest networks in the simulation, none of them do. This matches the findings of Hahn et al., 2024 that polarization increases with network size even for agents in Olsson's model of testimony, and for agents that do not engage in dynamic updating of trust. More importantly, for the present context, is the sharp contrast between 100% indexable evidence and all other percentages, regardless of exact level.

To follow up on this drop, we conducted a further simulation under the same conditions, contrasting just 100% versus 99% indexable material. The results of these additional 5000 simulation runs (500 per condition), confirm that the drop is apparent even at that small increment (see Fig. 2).

In fact, the simulations reveal two separate sources of increased belief dispersion: the increased dispersion for the larger networks reflects the fact that not all agents have seen all evidence by the end of the run. As the networks are initialized with one piece of evidence per agent, there is simply more evidence to distribute, yet the number of neighbours stays the same. This is seen clearly in the 100% indexable condition where there is no double-counting. The rate of that increase fits with the simulations on 'flooding the zone' conducted in the NormAN framework, which find linear increases in the time to convergence as a function of number of arguments (Hahn et al., 2025). This contrasts noticeably with the sharp jump in extremity of polarization seen once double-counting is introduced. In particular, this includes cases of extreme polarization, and includes cases where networks would no longer converge even if run indefinitely unless they are perturbed by new evidence or new connections. We illustrate how and why,

[2] Their model code is publicly available on The Open Science Framework (OSF) at https://osf.io/8qz9c/?

[3] Our model code, data and analysis scripts are publicly available on OSF at https://osf.io/yqaex/overview?view_only=13e891bd582d49a7869b914a3e53d64c

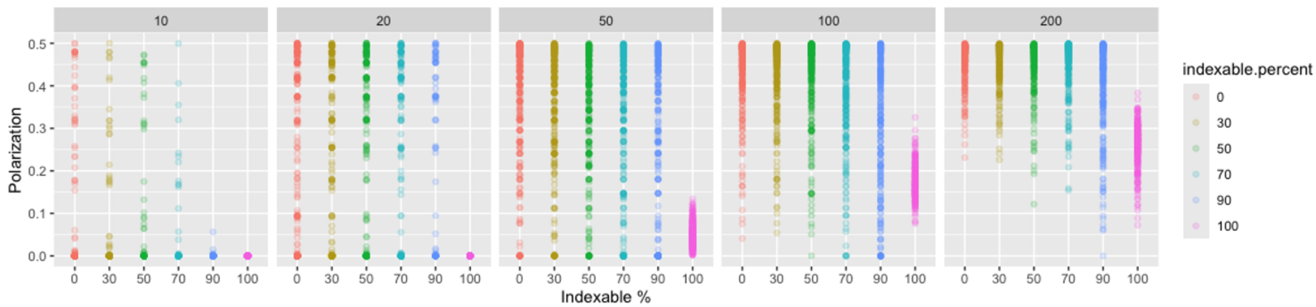

Figure 1: Dispersion (AAD) as a function of indexable evidence and network size. Each point shows the final AAD value from one simulation run (200 time steps); higher values indicate greater dispersion. Each panel corresponds to one network size (number of agents). The x-axis shows the percentage of total evidence that is indexable.

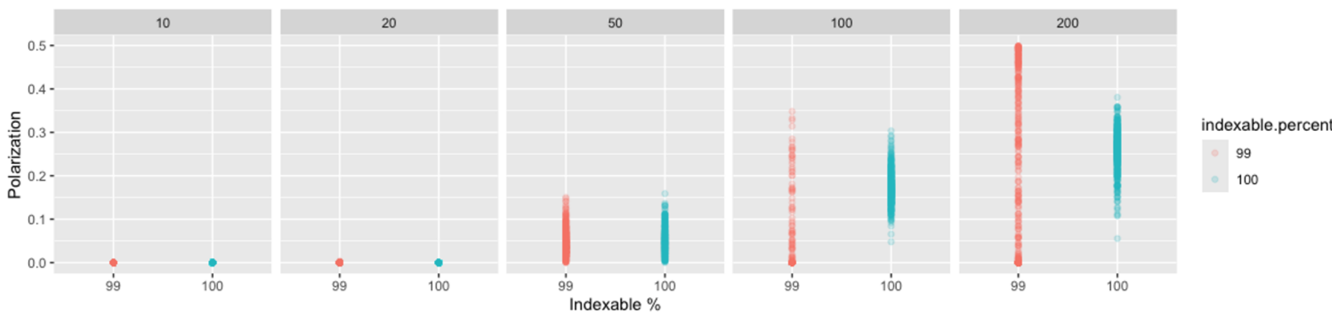

Figure 2: Dispersion (AAD) as a function of indexable evidence and network size. Each point shows the final AAD from one simulation run (200 time steps). Panels correspond to network size (number of agents). The x-axis shows the percentage of total evidence that is indexable. We compare 99% indexable (orange) to 100% (blue).

specifically, such cases arise in the next section.

Explaining the Effects

The only way our Bayesian agents in the simulation can diverge is because they possess different evidence. Fig. 3 shows the starting conditions of a sample run as seen in the NetLogo interface. In the top left corner is the distribution of sampled likelihood ratios, with evidence lending support to the hypothesis coloured in green, and the evidence against coloured red. There are 10 pieces of evidence in this instance, one for each of the 10 agents. At time step 0, that is, on intialization, each agent takes that piece of evidence and calculates their own posterior degree of belief. That belief distribution is shown in the histogram of the right hand panel, were each bar corresponds to one agent. The bottom left hand panel, finally, shows a histogram detailing how often each argument has been 'heard' across the entire population of agents. Each of the ten bars, from left to right, corresponds to one of the 10 likelihood ratios in the plot above it—all bar heights are equal, because, at initialization, each piece of evidence has been heard exactly once.

Fig. 4 then shows how the distribution of number of times heard has developed after a number of time steps invoving communication in the model (top row panels) and the resultant beliefs of the agents (bottom row panels). First, there is now considerable variation in the number of times each argument (piece of evidence) as been heard across all agents. However, the examples have been chosen to illustrate two qualitatively distinct cases in which that variability can manifest. On the left hand side is a run where a single piece of evidence in favour dominates (top left panel). This will push agents collectively toward one end of the belief distribution (bottom left panel). On the right hand side is a sample run where there are not only multiple peaks in that evidence-heard distribution, but peaks fall across either side of the neutral likelihood ratio 1. That is, there are peaks both for pieces of evidence *for* and *against*.

In this case, beliefs will diverge much more severely as seen in the corresponding belief distribution below (bottom right panel). This is because evidence that is double-counted can make whatever subset of evidence agents possess arbitrarily diagnostic. These extreme divergences may become stable in the sense that they will no longer shift over time. This is because different pieces of evidence are received via different paths. Those paths will vary in length, giving rise to different

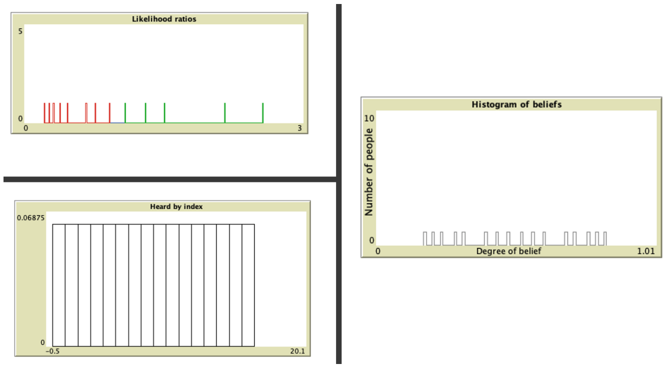

Figure 3: Simulation initialization: the evidence distribution in terms of Bayesian likelihood ratios (top left), the initial distribution of agent beliefs (right), and a histogram of how often each piece of evidence has been received ("heard") by agents (bottom left).

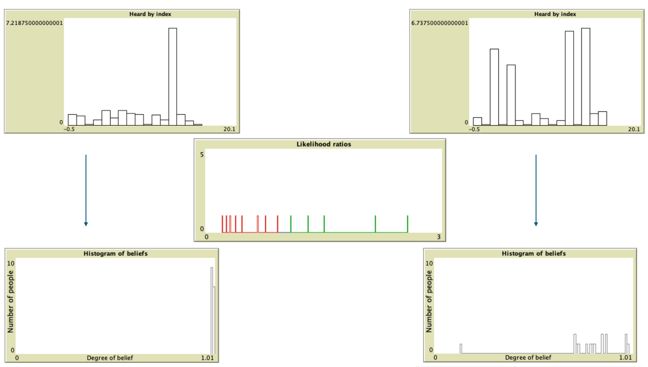

Figure 4: Simulation dynamics: Via deliberation, the distribution of how often each piece of evidence was heard becomes uneven (top row), yielding different belief outcomes (bottom row). Left: one pro-H item dominates, pushing beliefs towards $P(H) = 0$. Right: multiple peaks span both pro- and anti-H evidence, producing a divided belief distribution.

rates of reduplication, but unless the topology of the network itself is changed, the relative differences across individual pieces of evidence will be cyclical as determined by the length of those paths. In other words, though the numbers of times each piece of evidence is encountered will continue to increase over time, the systematic differences between them will remain (for a visual illustration see Hahn, 2023).

The differences in those times-heard histograms, finally, rest solely on the network topology in the current simulations (as agents randomly choose what to share). The only difference that emerges in final beliefs, and with that polarization, stems from the fact that, under perfect indexing, each piece of evidence is used to update an agent's beliefs at most once. So the different rates of reduplication are without consequence.

This suggests also that limiting re-sharing only to evidence that is not recognized as reduplication could only ever further *magnify* the pernicious effects of non-independence.

Conclusions

Some degree informational reduplication is a ubiquitous feature of communication on wider social networks. With no practical or formal tools at hand to deal with the issue, it is important to investigate potential group-level effects of double counting; not least because the type of polarization it creates is not generated by obviously irrational behaviour: many (formal) explanations trace polarization back to our epistemic irrationality—cognitive biases (Baccini et al., 2023), intellectual conformity (Weatherall & O'Connor, 2021) and trust dynamics (Fränken & Pilditch, 2021). Our results suggest that reasoners can diverge when evidential dependence is introduced by reduplication, even in the absence of blatant irrationalities (cf. Young et al., 2025).

The upshot is that contexts that create evidential dependence may be those where even idealized reasoners would come to diverge. Future (formal) work will thus be needed to study these contexts—for instance, larger networks and opaque online communication structures. Finally, if reduplication is a driver of polarization, then mitigating polarization may require the development of measures that reduce the risk of reduplication and evidential dependence.

Acknowledgments

L.A. gratefully acknowledges funding provided by the Konrad- Adenauer-Stiftung (PhD scholarship) and the support of the Chair of Philosophy of Science at the Munich Center for Mathematical Philosophy (LMU). D.Y. and U.H. acknowledge support by the European Union under the Horizon Europe project Perycles (Participatory Democracy that Scales: https://perycles-project.eu/). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency. Neither the European Union nor the granting authority can be held responsible for them.

Funded by
the European Union

References

Anderson, E. (2021). Epistemic bubbles and authoritarian politics. *Political epistemology*, 11–30.

Assaad, L., Fuchs, R., Jalalimanesh, A., Phillips, K., Schöppl, K., & Hahn, U. (2023). A bayesian agent-based framework for argument exchange across networks. *arXiv preprint arXiv:2311.09254*.

Assaad, L., Fuchs, R., Phillips, K., Schöppl, K., & Hahn, U. (2025). Capturing argument in agent-based models. *Topoi*.

Assaad, L., & Hahn, U. (2024). Rational polarization: Sharing only one's best evidence can lead to group polarization. *Proceedings of the annual meeting of the cognitive science society*, *46*.

Axelrod, R. (1997). The dissemination of culture: A model with local convergence and global polarization. *Journal of conflict resolution*, *41*(2), 203–226.

Baccini, E., Christoff, Z., Hartmann, S., & Verbrugge, R. (2023). The wisdom of the small crowd: Myside bias and group discussion. *JASSS*, *26*(4), 7.

Bala, V., & Goyal, S. (1998). Learning from neighbours. *The review of economic studies*, *65*(3), 595–621.

Bramson, A., Grim, P., Singer, D. J., Berger, W. J., Sack, G., Fisher, S., Flocken, C., & Holman, B. (2017). Understanding polarization: Meanings, measures, and model evaluation. *Philosophy of science*, *84*(1), 115–159.

Collins, P. J., Hahn, U., Von Gerber, Y., & Olsson, E. J. (2018). The bi-directional relationship between source characteristics and message content. *Frontiers in psychology*, *9*, 317842.

Fränken, J.-P., & Pilditch, T. (2021). Cascades across networks are sufficient for the formation of echo chambers: An agent-based model. *Journal of Artificial Societies and Social Simulation*, *24*(3).

Hahn, U. (2023). Individuals, collectives, and individuals in collectives: The ineliminable role of dependence. *Perspectives on Psychological Science*, 17456916231198479.

Hahn, U., Assaad, L., & Schöppl, K. (2025). Flooding the zone: An agent-based exploration. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *47*. https://escholarship.org/uc/item/8z32x661

Hahn, U., Merdes, C., & von Sydow, M. (2024). Knowledge through social networks: Accuracy, error, and polarisation. *Plos one*, *19*(1), e0294815.

Hartmann, S. (2021). Bayes nets and rationality. *The Handbook of Rationality*, 253.

Hegselmann, R., Krause, U., et al. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of artificial societies and social simulation*, *5*(3).

Henderson, L., & Gebharter, A. (2021). The role of source reliability in belief polarisation. *Synthese*, *199*(3), 10253–10276.

Jern, A., Chang, K.-M. K., & Kemp, C. (2014). Belief polarization is not always irrational. *Psychological review*, *121*(2), 206.

Kahan, D. M. (2013). Ideology, motivated reasoning, and cognitive reflection. *Judgment and Decision making*, *8*(4), 407–424.

Kopecky, F. (2024). Argumentation-induced rational issue polarisation: F. kopecky. *Philosophical Studies*, *181*(1), 83–107.

Landemore, H., & Page, S. E. (2015). Deliberation and disagreement: Problem solving, prediction, and positive dissensus. *Politics, philosophy & economics*, *14*(3), 229–254.

Madsen, J. K., Hahn, U., & Pilditch, T. D. (2020). The impact of partial source dependence on belief and reliability revision. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(9), 1795.

Mäs, M., & Flache, A. (2013). Differentiation without distancing. explaining bi-polarization of opinions without negative influence. *PloS one*, *8*(11), e74516.

Merdes, C., Von Sydow, M., & Hahn, U. (2021). Formal models of source reliability. *Synthese*, *198*, 5773–5801.

O'Connor, C., & Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science*, *8*, 855–875.

Olsson, E. J. (2020). Why bayesian agents polarize. In *The epistemology of group disagreement* (pp. 211–229). Routledge.

Pallavicini, J., Hallsson, B., & Kappel, K. (2021). Polarization in groups of bayesian agents. *Synthese*, *198*, 1–55.

Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. penguin UK.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco, CA: Morgan Kauffman.

Pilditch, T. D., Hahn, U., Fenton, N., & Lagnado, D. (2020). Dependencies in evidential reports: The case for informational advantages. *Cognition*, *204*, 104343.

Pilditch, T. D., Hahn, U., & Lagnado, D. (2025). The problem of dependency. *Synthese*, *205*(4), 143.

Schöppl, K. (2025). Industry influencing collective scientific reasoning: A bayesian, agent-based exploration. *Proceedings of the annual meeting of the cognitive science society*, *47*.

Schöppl, K., & Hahn, U. (2024). Exploring effects of self-censoring through agent-based simulation. *Proceedings of the annual meeting of the cognitive science society*, *46*.

Singer, D. J., Bramson, A., Grim, P., Holman, B., Jung, J., Kovaka, K., Ranginani, A., & Berger, W. J. (2019). Rational social and political polarization. *Philosophical Studies*, *176*, 2243–2267.

Sunstein, C. R. (2018). Republic: Divided democracy in the age of social media.

Taber, C. S., Cann, D., & Kucsova, S. (2009). The motivated processing of political arguments. *Political Behavior*, *31*, 137–155.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *Nature*, *393*(6684), 440.

Weatherall, J. O., & O'Connor, C. (2021). Conformity in scientific networks. *Synthese*, *198*(8), 7257–7278.

Whalen, A., Griffiths, T. L., & Buchsbaum, D. (2018). Sensitivity to shared information in social learning. *Cognitive science*, *42*(1), 168–187.

Young, D. J., Madsen, J. K., & De-Wit, L. H. (2025). Belief polarization can be caused by disagreements over source independence: Computational modelling, experimental evidence, and applicability to real-world politics. *Cognition*, *259*, 106126.

Yousif, S. R., Aboody, R., & Keil, F. C. (2019). The illusion of consensus: A failure to distinguish between true and false consensus. *Psychological Science*, *30*(8), 1195–1204.

Zollman, K. J. (2007). The communication structure of epistemic communities. *Philosophy of science*, *74*(5), 574–587.

Zollman, K. J. (2010). The epistemic benefit of transient diversity. *Erkenntnis*, *72*(1), 17–35.