

Modeling How False Beliefs Spread

Cailin O'Connor¹ and James Owen Weatherall²
Department of Logic and Philosophy of Science
University of California, Irvine

Introduction

Effective political decision making, like other decision making, requires decision-makers to have accurate beliefs about the domain in which they are acting. In democratic societies, this often means that accurate beliefs must be held by a community, or at least a significant portion of a community, of voters. Voters are tasked with scrutinizing candidates and possible policy proposals and, considering their own experiences, interests, goals, knowledge, and values, with deciding which of various ballot measures is most likely to bring about their desired outcomes.

For this process to make sense, voters need to have access to reliable information about what the current state of affairs is like – for instance, what problems need to be solved – and what the likely outcomes of policies will be. And yet, observation of contemporary political environments reveals widespread disagreement and polarization, not only about political values, but also about matters of fact, on topics ranging from the status of anthropogenic climate change, to the relative sizes of the crowds at Trump and Obama's inaugurations, to whether unemployment and the U.S. national debt increased or decreased under various U.S. presidential administrations. Irrespective of one's own beliefs or political affiliations, these disagreements reveal that false beliefs about matters of fact must be rampant, since after all, both sides cannot be correct on these matters.

Widespread false belief is often explained by appeal to individual epistemic factors, such as personal cognitive and epistemic biases, ignorance, and media illiteracy. But these individual factors can be at most only part of the story. The information needed to form reliable beliefs is difficult to come by, and generally requires a broad range of experiences and domain-specific knowledge that no individual can hope to achieve. In virtually all cases, individuals cannot gather and evaluate all, or even most, of the evidence and information that they need to make the sorts of decisions that democracy implicitly demands. Instead, they rely on others. Most of what we know, or believe, we have learned from other people.

Over the past four decades, observations of this sort, i.e., that most of our beliefs come from other people, have motivated the development of the field of social epistemology (see Goldman and O'Connor 2019 for more details). Social epistemology seeks to understand how social factors do and should influence belief. Insights gleaned in this field are of special importance for political epistemology, both because political decision making is often distributed across many members of a community and because, in democratic societies, voters must make decisions

¹ cailino@uci.edu

² james.owen.weatherall@uci.edu

about matters on which they have no personal experience, which may then affect others' lives more significantly than their own.

More recently, philosophers have begun to use formal methods, including mathematical models and computer simulations based on game theory, decision theory, and network science, to explore various aspects of social epistemology. A topic of particular focus has been the (formal) social epistemology of *false belief* (O'Connor & Weatherall 2019): that is, how do social factors contribute to the persistence and spread of false beliefs within a community, even in cases where evidence of their falsehood is available? Formal models are a particularly useful way of exploring this question because they allow one to step away from disputed and politically fraught questions concerning actual matters of fact, and ask how even ideal agents, in a context where there is no ambiguity regarding what the truth is, may nonetheless come to believe falsehoods.

This chapter will review the recent literature in the formal social epistemology of false belief. We will begin by introducing some further background on social epistemology. We will then introduce some basic ideas from the formal modeling literature, explaining how researchers use models to address the questions at hand and discussing various results. We then move on to what these models tell us about how politically or economically motivated actors shape public belief by exploiting social factors.

Social Epistemology

Traditional philosophical approaches to epistemology were largely individualistic in focus. But in the late 1980s, the field of social epistemology emerged, with a core mission to investigate social aspects of knowledge. A central tenet of this field is that belief and knowledge are deeply shaped by our social contexts and connections, and that to understand doxastic states in humans we need to attend to these social considerations (Goldman and O'Connor, 2019).

The key debates in social epistemology have been largely normative, in that they consider how beliefs or knowledge should be adopted in social contexts. For instance, the "problem of testimony" concerns what beliefs an agent should form in light of testimony from others, and whether these beliefs are justified. (See, for instance, Coady (1992); Fricker (1994).) The peer disagreement literature considers how individuals should change beliefs in light of disagreement (Christensen, 2007), and the judgement aggregation literature concerns how a group of agents should combine judgements (List and Petit, 2011). The "two experts, one novice" problem asks what an uninformed individual should do in the face of expert disagreement (Goldman, 2001). Other debates, such as that over collective agents, are less normative, but ask questions such as: what is it for a *group* to hold a belief, as opposed to an individual (Gilbert, 1989)?

Relatively less explored in this area are more empirically focused questions about how, in fact, beliefs spread between individuals. Why, for instance, do individuals actually make the judgement to trust some testimony, but not others? How do they decide what to believe when

experts disagree? How do groups, in fact, aggregate their judgements when making decisions? These sorts of questions have been widely addressed in the social sciences, and especially in psychology and behavioral economics. It will be beyond the purview of this paper to survey this literature, though it has been deeply influential in shaping the work we will discuss in the rest of the paper.

As noted, one off-shoot from social epistemology is the subfield sometimes called formal social epistemology. This subfield uses modeling methods and formal paradigms to address social aspects related to knowledge and belief. Much of this work addresses more descriptive/empirical questions such as: how does information flow in social networks? And, how do aspects of social learning influence this spread? In addition, it often seeks to make normative recommendations about the organization of epistemic communities by appeal to our empirical understanding of these communities, such as: how can we organize groups so as to best preserve accurate belief, given what we know about human social learning?

Much of the work in this subfield has been done by philosophers of science, who start with questions related to the spread of knowledge and theories in science. Given the interdisciplinary nature of the field, it is unsurprising that there is a great deal of overlap with work in the social sciences on belief and sociality. For the purposes of this paper, we will curtail our focus to one corner of this literature. In particular, we will look at some of the work using networks to consider the spread of true and false beliefs.³

Modeling the Spread of False Beliefs

One key insight into human knowledge is that most people derive most of their beliefs from others in their community. We all learn some things on our own - for instance that one's child reacts badly to grapefruit, or that the local squirrel likes to steal birdseed. But, especially when it comes to scientific beliefs, we learn a tremendous amount via social channels. There are different ways this happens. First, we learn from the testimony of others. I might tell you that a particular caterpillar is poisonous and if you trust my testimony you will adopt (and maybe pass on) this belief. Second, we share evidence with others. For instance, I might pass on a graph of a model charting the spread of a virus under different interventions. Given this socially spread evidence, your beliefs may change.

There are several approaches to modeling belief transmission. In some models, one has a collection of "agents" -- mathematical representations of idealized inquirers -- all of whom may interact with others to exchange information (e.g. Hegselman and Krause 2002). In other cases, the agents are arranged in a network, so that each agent can share information only with certain others, representing social connections. A *network* in this context consists in a collection

³ There is a related literature in formal epistemology that mostly uses proof-based methods to consider normatively ideal learning in social conditions. While we will not discuss this literature, interested readers can find an overview in Goldman and O'Connor (2019).

of *nodes*, associated with the models' agents. These nodes are connected by *edges* which represent social ties across which information might pass.

Two dominant network approaches to epistemology approximately track the two sorts of scenarios described in the first paragraph of this section. On one of these approaches, network models treat beliefs/ideas as virus-like. They pass from person to person directly, as an infection might. This best corresponds to a scenario where a belief spreads via testimony. The relevant modeling paradigm is sometimes referred to as the *diffusion* or *contagion approach*. (There are a number of variations and subtleties here, which we largely ignore.) The second approach models agents who both gather and share evidence with those in their social networks. This is sometimes called the *network epistemology approach* and corresponds to a scenario where beliefs spread socially, but the spread depends on the transmission of evidence rather than just testimony.⁴

The contagion/diffusion approach has proved a useful paradigm for modeling many aspects of belief spread (Rogers, 2010). It is particularly well-tuned for things like rumors (or “memes”): units of belief that are accepted and shared relatively non-reflexively, rather than subject to reasoning. We will not spend much time on this literature here, but will mention some findings to give the reader a sense of what it can do. A number of authors, such as Jackson and Rogers (2007), look at how network structure influences the diffusion of ideas. Golub and Jackson (2012), for example, consider how homophily -- disproportionate connectivity with in-group members -- could slow information spread.⁵ Some use this paradigm to consider the formation of epistemic networks. Barrett et al. (2019), for instance, consider what happens when epistemically focused agents choose to listen to good inquirers in a diffusion type network. Still others consider variants on learning rules. Granovetter (1978), for instance, presents an influential diffusion model where individuals only pick up new beliefs when a threshold percentage of neighbors hold them.

In this paper we focus more on the second type of model, where agents share evidence. This is particularly germane in thinking about the spread of scientific beliefs, which are often supported by data. An influential version of this sort of model was introduced by economists Bala and Goyal (1998). Kevin Zollman (2007, 2010) imported it to philosophy of science to study theory change in scientific communities.

How do these models work? They involve two elements: a network and a decision problem, called a *multi-armed bandit problem*. This sort of problem is analogous to choosing between slot machine (or “bandit”) arms that pay off at different rates. It is assumed that the agents involved are unsure about which choice (or “arm”) is the best one. They make this decision by

⁴ Some models include aspects of both elements. Golub and Jackson (2010), for instance, consider agents who gather evidence, but who develop beliefs by averaging those of their neighbors rather than by sharing evidence directly.

⁵ They use a variation on the diffusion idea, where agents adopt beliefs that average those of their neighbors.

actually testing their choices and seeing what happens. Over time their theories about the success of the arms changes as they gather evidence.

The social epistemology aspect of this model involves the sharing of this data through the network. Over subsequent rounds, agents gather data and share it with their neighbors, always updating their beliefs about the arms in light of their own results and the results shared with them by their neighbors. Many versions of the model involve what we might call “myopic Bayesians”. The agents are myopic because they tend to choose whichever arm they think best. They are Bayesians because they have credences about the arms, and use Bayes’ rule to update these credences in light of evidence. In other words, they display a sort of bounded rationality.

In these models, the social sharing of evidence means that individuals tend to converge with respect to belief. If I favor a poorer action, I do not personally gather data about better ones. But I may receive this data from my neighbors, and it will eventually be enough to guide me towards better actions. Then the evidence generated by my behavior may, in turn, convince my (other) neighbors to try a better action.

This sort of process means that in baseline versions of this model, communities always end up sharing beliefs. Consider the simple case where agents choose between two actions, A and B. These might represent doctors administering two different treatments for a given ailment. The doctors’ choices are guided by credences about which treatment succeeds more often. Suppose B has a higher probability of success than A.⁶ A typical model will start with agents who have random credences about whether B is better than A. In successive rounds, they try their preferred action, record the frequency with which that action succeeds (where those succeeded occur with some fixed probability, unknown to the agents), and share those results, always updating beliefs in light of that evidence and the evidence shared with them by their neighbors. Sometimes the community will end up with all agents having high credence that B is, in fact better. In other cases (though less often) the community ends up preferring A.

Why might they settle on the worse action -- in this case the treatment that works less often? This can result from strings of misleading data. Real evidence is often probabilistic, just like the evidence in the model. A medication that is generally successful may not work for everyone, whereas one that is less successful may still work sometimes. Likewise, action B may payoff less frequently than action A in a particular set of trials. This might cause all the agents in the community to turn to action A. If this happens, because all agents stop performing action B, they fail to learn more about it and remain in a suboptimal belief state.

What does this model tell us about social epistemology? First, it makes clear how evidence sharing can get a community on the same page, and how it can be beneficial. Without social influence, all individuals develop beliefs separately. Those who start with poor beliefs are never

⁶ And further suppose that when success happens it yields the same payoffs.

even exposed to evidence that might change them. This means that the actors in the network model tend to do better together than they do alone. But the model also illuminates why failures can happen. Group influence can lead an entire community to pre-emptively settle on a poor theory (even though individually some of them might have gotten things right).

In the rest of this section, we will describe some results derived from these models in the discipline of formal social epistemology. These results look at various features of the model, and at how those features influence, on average, the success of communities. They sometimes provide guidance into how real communities can fail (or succeed) in spreading belief.

Zollman (2007, 2010) considers the effect of connectivity on community success. In particular, he runs simulations of the models just described and compares those where actors are linked to relatively many vs. relatively few neighbors. In general, for such models, one might suppose that the more connections the better. Each neighbor shares only real data which, in general, tends to push one towards true beliefs. Paradoxically, Zollman finds that, in fact, less connected networks do better. This *Zollman effect* happens because tightly connected networks share data widely, meaning that a string of misleading evidence can convince everyone of the wrong belief. Less connected groups tend to preserve pockets of accurate belief, which can then lead the rest of the community to a better theory. Rosenstock et al. (2017) show that this effect only occurs for certain parameter values of the model, corresponding to cases where inquiry is difficult. In other words, when a community actually struggles to find truth, too much communication can be harmful.

Mayo-Wilson et al. (2011) use this model to defend a thesis central to the discipline of social epistemology, which is that the rationality of individuals and the rationality of communities can come apart. This *Independence Thesis* holds that sometimes irrational individuals can constitute effective communities of learners, while rational individuals fail to form a successful group. For example, if theory A is the most promising in a community, it may make sense for every individual to test it. But the community will do better if labor is divided and individuals test many, diverse possibilities and then share their knowledge. (See also Kitcher (1990).) Alternatively, irrationally stubborn actors can make a group more successful, by focusing on many different plausible theories and sharing knowledge about them.

O'Connor and Weatherall (2018) add to the base model a tendency of human learners that has been widely documented by psychologists. They suppose that the individuals place greater trust in data shared by those with similar beliefs. As noted, the base model always tends to consensus. But once this aspect of social trust is introduced, polarization, or the emergence of stable, opposing beliefs, becomes common. In particular, feedback loops emerge where individuals with different beliefs do not fully trust each others' evidence, and form increasingly distant beliefs until they no longer influence one other. Previous models have found similar feedback loops can lead to polarization (e.g., Hegselmann and Krause (2002), Olsson (2013)), but, remarkably, O'Connor and Weatherall show how this can happen even in the face of evidence gathering and otherwise rational belief updating.

In a follow-up paper, Weatherall and O'Connor (2020b) study how the same trust dynamics work when they are applied over multiple domains of belief. In this variation on the model, agents who agree on one topic place more trust in one another's evidence concerning other topics, even if there is no intrinsic relationship between the topics. They find that agents can form "epistemic factions", which are subcommunities with highly correlated beliefs, where those beliefs are polarized over the community as a whole. Remarkably, this occurs endogenously in the model, with no coordination. It suggests that some cases of correlated political polarization over multiple matters of fact, a phenomenon often attributed to shared values or ideology amongst members of the factions (e.g. by Lakoff (2010)) can arise without any such common cause. This can perhaps explain how, for instance, during the COVID-19 pandemic, beliefs about whether the antiviral hydroxychloroquine was an effective treatment for COVID-19 -- a matter with high stakes for the belief-holders and with no clear ideological basis -- became polarized in the United States along existing political lines.

Weatherall and O'Connor (2020a) look at another typical social behavior -- conformity bias -- and see how it influences these models. Agents who conform are more likely to engage in the behavior favored by their neighbors, even if it is not the one they would usually prefer. As Weatherall and O'Connor show, this addition to the model significantly hurts the knowledge-producing capacity of the community. Conforming individuals with accurate beliefs fail to spread useful data to their peers. And due to conformity, social networks with cliques can end up with polarization as well. Likewise Mohseni and Williams (2019) find that conformity bias in a network of Bayesian updaters hurts their ability to settle on accurate beliefs. In particular they find that highly connected individuals who constitute hubs in the network cause problems because many others try to conform with them, while they are strongly influenced by their connections.

As we have seen, these models have been used to inform a number of questions: What is the optimal communication structure? Do rational individuals necessarily make rational communities? How do social biases impede the spread of good beliefs in an epistemic network? These investigations, as noted, have a descriptive character. They seek to illuminate the workings of real communities (or at least to direct research towards plausible effects in real communities). Like more traditional work in social epistemology, though, they also answer normative questions, but at the community level. How do we design the best epistemic communities?

Propaganda in Epistemic Communities

Many widespread false beliefs appear to emerge and spread endogenously--that is, without any kind of intervention by actors who wish to influence public belief. For instance, during the late 19th century, there was a widespread belief in the northeastern United States that the so-called "tomato worm" was highly venomous, and that it could kill someone merely by "spitting" its venom onto their skin. Stories about the dangers of the worm, which was actually the

(harmless) larva of the hawkmoth, spread in newspapers, and some deaths were attributed to it. This belief apparently originated in error, but then propagated independently for years afterwards (Smith, 2001).

Such beliefs are often, though not always, harmless. (As an example of a harmful one, during the COVID-19 epidemic in Iran, viral rumors spread that drinking pure alcohol could cure COVID, leading to hundreds of deaths.) But many of the most politically salient examples of widely held false beliefs appear to have a different character. In these cases, there are often powerful politically and economically-motivated actors who have a vested interest in the public holding certain beliefs (or failing to come to hold certain beliefs).

For instance, although today there is broad scientific consensus, generally shared by the public, that tobacco products cause cancer, this has not always been the case. In fact, as historians of science Naomi Oreskes and Eric Conway (2010) describe, when evidence began to come to the public's attention in the early 1950s that tobacco use may have severe adverse health effects, the tobacco industry responded with a number of interventions designed to combat public uptake of this evidence. (See also Brandt (2007) and Michaels (2008) for further material on this case and several others.)

There is considerable reason to think that these efforts succeeded. For instance, cigarette sales briefly dipped (for the first time in decades) following the publication of a widely read *Reader's Digest* article, "Cancer by the Carton", in 1952. They rebounded, however, following a massive public relations push by industry-supported groups and continued to grow for decades, even as the tobacco firms themselves established (internally) that the hypothesized link between tobacco and cancer was real. Moreover, efforts to regulate tobacco sales, including advertising and sales to minors, stalled for decades.

In this case, industry agents succeeded in propagating the false belief that cigarette smoke is not harmful; and, equally importantly, that the scientific record on the harms of tobacco was not conclusive. Oreskes and Conway (2010) argue that these same methods were used to promote falsehood and manufacture uncertainty regarding a large number of politically salient issues, often with significant public health impacts, including the causes and harms of acid rain; the effects of chlorofluorocarbons on ozone; and the status of anthropogenic climate change.

Several philosophers have used the formal methods discussed in the previous section to better understand the various strategies employed by the tobacco industry and other such groups to manipulate public belief.⁷ In the first work along these lines, Holman and Bruner (2015) consider a modification of the basic model described above in which one agent -- whom they call an "intransigently biased agent", but whom we will call a "propagandist", for consistency in what follows -- shares evidence drawn from a sample biased towards action A, i.e., the worse of the two actions. The idea here is that the propagandist wishes to convince other agents in the

⁷ See O'Connor & Weatherall (2019, Chapter 3) for a more detailed overview of this work.

network that A is, in fact, better; and to do so, they share spurious evidence that supports their preferred hypothesis, but which does not reflect the true success frequencies of the two arms. This agent is epistemically “impure” in the sense that they perform their preferred action, and report biased results, irrespective of what other evidence they have seen.

Holman and Bruner show that with this simple modification, the model will no longer converge to a true consensus in general, because the biased evidence generated by the propagandist will prevent the rest of the community from reaching consensus on whether the better arm is in fact better. They also consider agents in the network who attempt to identify propagandists by comparing the distributions of results shared by other agents with their own beliefs. They find that this strategy for rooting out propagandists can succeed, though it will be more effective in some cases than others--for instance, when the number of “draws” per round of the model is low (corresponding roughly to a situation in which members of a scientific community tend to perform experiments with low statistical power).

Weatherall et al. (2018) consider a different modification to the basic model described above. Whereas in the Holman-Bruner model, the propagandist seeks to influence their “epistemic peers”, i.e., other agents who are gathering and sharing evidence, Weatherall et al. consider what happens when a propagandist seeks to influence outside observers who form beliefs on the basis of evidence shared by others, but do not collect evidence themselves; such agents might represent the public or policymakers who learn from scientists, but are not engaged in research. In this model, the observers receive evidence both from agents who are performing experiments (“scientists”) and sharing (all) of their data as usual in model; and from an additional agent -- the propagandist -- who shares evidence in a manner intended to convince the observers that the incorrect action is, in fact, better.

Weatherall et al. consider two different strategies that the propagandist might adopt. The first strategy is what they call “biased production”. This strategy, which is similar in spirit to the one considered by Holman and Bruner, consists in performing a number of trials of the worse arm, and then reporting only those results that happen to (spuriously) suggest it is better.

On the second strategy, which they call “selective sharing”, the propagandist does not perform any experiments at all. Instead, the propagandist merely searches the results generated each round by the network of scientists, and shares results produced “organically”, that is, by the scientists themselves, that happen to support action A. This strategy works because given the probabilistic nature of evidence in these models, in any given round a certain number of results apparently supporting the worse action may be generated by the scientists themselves--with more such results appearing in large networks of scientists, especially when the number of trials (again, a proxy for the “power” of studies) is small.

In both cases, the effect of the propagandist’s strategy is that the observers see more data supporting action A than they would if the propagandist were not operative. The result is that under a wide range of parameters the observers come to hold the false belief, even as the

scientists converge to the true belief. In this sense, both strategies can succeed. That said, Weatherall et al. also argue that, perhaps contrary to intuition, selective sharing is often more effective and less costly than biased production. This is because this strategy does not require the propagandist to perform their own experiments at all, saving costs and effort.

A key insight from all of these models is this. Real scientific evidence can be noisy, with results sometimes pulling in different directions. Agents come to form the correct belief only by considering a representative sample from their neighbors, which includes both some results supporting the worse action and (generally more, stronger) results supporting the better one. By sharing a disproportionately large sample of results supporting action A, propagandists can shift this balance.

A second, perhaps even more important insight is that a propagandist can influence public belief without manipulating individual results or violating the standard “norms” of science. Under the selective sharing strategy, the propagandist shares only “legitimate” science, in the sense of science produced by unbiased, truth-seeking agents without outside influence. The strategy works by “curating” that evidence. Of course the propagandist is effectively cherry-picking in the model, but they are not sharing fraudulent evidence.

This theme, that propagandists can influence scientific communities and public belief without committing “fraud” or other norm-violations, is further emphasized by another model studied by Holman & Bruner (2017). The set-up of this model is somewhat more complicated, but the basic idea is that within a research community, there may be a diversity of methods, operationalizations, working hypotheses, etc. It is assumed that scientists adopt these methods for their own reasons, free of outside influence. And yet, a propagandist may have reason to believe that some of these methods -- for instance, emphasis on particular forms of evidence or specific research questions -- may be more likely to favor the propagandist’s preferred outcome than others.

In this model, the propagandist adopts the following strategy, which Holman and Bruner dub “industrial selection”: they identify agents who have independently adopted research methods they consider favorable, and they “promote” those agents, essentially by increasing the amount of evidence they can share. For present purposes, one may think of this as lavishing those researchers with funding, enabling them to publish more papers or share their results more widely at conferences or in the press. The net result is, once again, that the propagandist increases the number of industry-favorable results that are shared within the network--now without producing any experiments of their own, or even sharing any results themselves. Instead, they do so by shifting the balance of how much research of different kinds can be produced.

Industrial selection is especially pernicious, because the scientists who are “selected” can honestly claim that they have continued to do precisely the research they would have done anyway, even without industry funding. The propagandist does not ask them to change their

methods, or bias their results. They are, in a certain sense, “independent”. And yet by disproportionately amplifying some methods over others, the propagandist achieves their aim, again, by biasing the total body of evidence seen by other agents in the network.

Once again, we have seen these models used to investigate a number of questions. How do propagandists use social factors in how information spreads to manipulate public belief? What strategies are most effective, and why? But many interesting questions remain. For instance, are their other widely used methods for influencing belief that may be elucidated with these methods? And perhaps most importantly, how can we change the social structure of information and belief to help guard against propagandistic influence?

Conclusion

Formal social epistemology provides a rich toolbox for studying the mechanisms by which false beliefs persist and spread within a community. It can also help identify the conditions under which falsehoods spread most readily--and, perhaps, those conditions under which the truth is more likely to prevail. In this chapter we have reviewed a recent body of work that uses methods from network theory and decision theory to tackle these questions. We have highlighted what we take to be the key insights of these models, but also raised questions that we think should be pursued in future work. This literature has clear implications for those interested in political epistemology, and especially in the successful functioning of democracy.

Bibliography

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

Barrett, J. A., Skyrms, B., & Mohseni, A. (2019). Self-Assembling Networks. *The British Journal for the Philosophy of Science*, 70(1), 301-325.

Brandt, Allan. (2007). *The Cigarette Century: The Rise, Fall, and Deadly Persistence of the Product that Defined America*. New York: Basic Books.

Christensen, David, 2007, “Epistemology of Disagreement: The Good News”, *Philosophical Review*, 116(2): 187–217. doi:10.1215/00318108-2006-035

Coady, C. A. J., 1992, *Testimony: A Philosophical Study*, Oxford: Clarendon Press.

Fricker, Elizabeth, 1994, “Against Gullibility”, in *Knowing from Words*, Bimal Krishna Matilal and Arindam Chakrabarti (eds.), Dordrecht: Springer Netherlands, 125–161. doi:10.1007/978-94-017-2018-2_8

Gilbert, Margaret, 1989, *On Social Facts*, New York: Routledge.

Golub, B., & Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1), 112-49.

Golub, B., & Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3), 1287-1338.

Goldman, Alvin, 2001, "Experts: Which Ones Should You Trust?", *Philosophy and Phenomenological Research*, 63(1): 85–110. Reprinted in Goldman & Whitcomb (eds.) 2011, 109–133. doi:10.1111/j.1933-1592.2001.tb00093.x

Goldman, Alvin and O'Connor, Cailin, "Social Epistemology", *The Stanford Encyclopedia of Philosophy* (Fall 2019 Edition), Edward N. Zalta (ed.)
<<https://plato.stanford.edu/archives/fall2019/entries/epistemology-social/>>.

Granovetter, M. (1978). Threshold models of collective behavior. *American journal of sociology*, 83(6), 1420-1443.

Hegselmann, Rainer and Ulrich Krause, 2002, "Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation", *Journal of Artificial Societies and Social Simulation*, 5(3): 2.

Holman, Bennett, and Justin P. Bruner. 2015. The Problem of Intransigently Biased Agents. *Philosophy of Science* 82 (5):956–68. <https://doi.org/10.1086/683344>.

Holman, B., & Bruner, J. 2017. Experimentation by industrial selection. *Philosophy of Science*, 84(5), 1008-1019.

Jackson, M. O., & Rogers, B. W. 2007. Relating network structure to diffusion properties through stochastic dominance. *The BE Journal of Theoretical Economics*, 7(1).

Kitcher, Philip, 1990, "The Division of Cognitive Labor", *Journal of Philosophy*, 87(1): 5.
doi:10.2307/2026796

Lakoff, G. (2010). *Moral politics: How liberals and conservatives think*. University of Chicago Press.

List, Christian and Philip Pettit, 2011, *Group Agency: The Possibility, Design, and Status of Corporate Agents*, Oxford: Oxford University Press.
doi:10.1093/acprof:oso/9780199591565.001.0001

Mayo-Wilson, Conor, Kevin J. S. Zollman, and David Danks, 2011, "The Independence Thesis: When Individual and Social Epistemology Diverge", *Philosophy of Science*, 78(4): 653–677. doi:10.1086/661777

Michaels, David. 2008. *Doubt Is Their Product: How Industry's Assault on Science Threatens Your Health*. 1 edition. Oxford ; New York: Oxford University Press.

Mohseni, A., & Williams, C. R. (2019). Truth and conformity on networks. *Erkenntnis*, 1-22.

O'Connor, C., & Weatherall, J. O. (2018). "Scientific polarization". *European Journal for Philosophy of Science*, 8(3), 855-875.

Olsson, E. J. (2013). A Bayesian simulation model of group deliberation and polarization. In *Bayesian argumentation* (pp. 113-133). Springer, Dordrecht.

Oreskes, Naomi, and Erik M. Conway. 2010. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. New York, NY: Bloomsbury Press.

Rogers, E. M. (2010). *Diffusion of innovations*. Simon and Schuster.

Rosenstock, S., Bruner, J., & O'Connor, C. (2017). In epistemic networks, is less really more?. *Philosophy of Science*, 84(2), 234-252.

Smith, A. F. (2001). *The tomato in America: early history, culture, and cookery*. University of Illinois Press.

Weatherall, J. O., O'Connor, C., & Bruner, J. (2018). "How to Beat Science and Influence People: Policymakers and Propaganda in Epistemic Networks". *British Journal for Philosophy of Science*. DOI: 10.1093/bjps/axy062

Weatherall, J. O. & O'Connor, C. (2020a). "Conformity in scientific networks". *Synthese*. DOI: 10.1007/s11229-019-02520-2

Weatherall, J. O. & O'Connor, C. (2020b). "Endogenous epistemic factionalization". *Synthese*. DOI : 10.1007/s11229-020-02675-3

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.